NIPS 2018 Workshop book

Generated Thu Jan 31, 2019

Workshop organizers make last-minute changes to their schedule. Download this document again to get the lastest changes, or use the NIPS mobile application.

Schedule Highlights

Dec. 7, 2018

Room 220 C, **Causal Learning** *Arjovsky, Heinze-Deml, Klimovskaia, Oquab, Bottou, Lopez-Paz*

Room 220 D, **Bayesian Deep Learning** *Gal, Hernández-Lobato, Louizos, Wilson, Ghahramani, Murphy, Welling*

Room 220 E, **Deep Reinforcement Learning** *Abbeel, Silver, Singh, Pineau, Achiam, Houthooft, Srinivas*

Room 510 ABCD, **MLSys: Workshop on Systems for ML and Open Source Software** *Lakshmiratan, Bird, Sen, Gonzalez, Crankshaw*

Room 511 ABDE, **Critiquing and Correcting Trends in Machine Learning** *Rainforth, Kusner, Bloem-Reddy, Paige, Caruana, Teh*

Room 511 CF, **Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy** *Veloso, Kallus, Shah, Kumar, Moulinier, Chen, Paisley*

Room 512 ABEF, **Smooth Games Optimization and Machine Learning** *Lacoste-Julien, Mitliagkas, Gidel, Syrgkanis, Tardos, Bottou, Nowozin*

Room 512 CDGH, **Visually grounded interaction and language** *Strub, de Vries, Wijmans, Datta, Perez, Malinowski, Lee, Anderson, Courville, MARY, Batra, Parikh, Pietquin, HORI, Marks, Cherian*

Room 513 ABC, **Modeling and decision-making in the spatiotemporal domain** *Senanayake, Jean, Ramos, Chowdhary*

Room 513DEF, **Workshop on Security in Machine Learning** *Papernot, Tramer, Chaudhuri, Fredrikson, Steinhardt*

Room 514, **2nd Workshop on Machine Learning on the Phone and other Consumer Devices (MLPCD 2)** *Ravi, Chai, Jia, Aradhye, Jain*

Room 515, **Machine Learning for Geophysical & Geochemical Signals** *Pyrak-Nolte, Rustad, Baraniuk*

Room 516 AB, **Workshop on Ethical, Social and Governance Issues in AI** *Bakalar, Bird, Caetano, Felten, Garcia, Kloumann, Lattimore, Mullainathan, Sculley*

Room 516 CDE, **Imitation Learning and its Challenges in Robotics** *Mukadam, Choudhury, Srinivasa*

Room 517 A, **Continual Learning** *Pascanu, Teh, Pickett, Ring*

Room 517 B, **NIPS 2018 workshop on Compact Deep Neural Networks with industrial applications** *Fan, Lin, Welling, Chen, Bailer*

Room 517 C, **Modeling the Physical World: Learning, Perception, and Control** *Wu, Allen, Smith, Hamrick, Dupoux, Toussaint, Tenenbaum*

Room 517 D, **All of Bayesian Nonparametrics (Especially the Useful Bits)** *Cai, Campbell, Hughes, Broderick, Foti, Williamson*

Room 518, **NeurIPS 2018 Competition Track Day 1** *Escalera, Herbrich*

Room 519, **The second Conversational AI workshop – today's practice and tomorrow's potential** *Geramifard, Williams, Boureau, Eskenazi, Gasic, Glass, Hakkani-Tur, Heck, Polymenakos, Young*

Dec. 8, 2018

Room 220 D, **Integration of Deep Learning Theories** *Baraniuk, Anandkumar, Mallat, Patel, H■*

Room 220 E, **NIPS 2018 Workshop on Meta-Learning** *Grant, Hutter, Ravi, Vanschoren, Wang*

Room 510 AC, **Machine Learning for Systems** *Goldie, Mirhoseini, Raiman, Swersky, Hashemi*

Room 510 BD, **Machine Learning for the Developing World (ML4D): Achieving sustainable impact** *De-Arteaga, Herlands, Coston*

Room 511 ABDE, **CiML 2018 - Machine Learning competitions "in the wild": Playing in the real world or in real time** *Guyon, Viegas, Escalera, Abernethy*

Room 511 CF, **NeurIPS 2018 Competition Track Day 2** *Herbrich, Escalera*

Room 512 ABEF, **Wordplay: Reinforcement and Language Learning in Text-based Games** *Trischler, Lazaridou, Bisk, Tay, Kushman, Côté, Sordoni, Ricks, Zahavy, Daumé III*

Room 512 CDGH, **Privacy Preserving Machine Learning** *Bellet, Gascon, Kilbertus, Ohrimenko, Raykova, Weller*

Room 513 ABC, **Medical Imaging meets NIPS** *Konukoglu, Glocker, Lombaert, de Bruijne*

Room 513DEF, **Interpretability and Robustness in Audio, Speech, and Language** *Ravanelli, Serdyuk, Variani, Ramabhadran*

Room 514, **NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018** *Li, Dragan, Niebles, Savarese*

Room 515, **Machine Learning Open Source Software 2018: Sustainable communities** *Strathmann, Gal, Curtin, Lisitsyn, Honkela, Ong*

Room 516 AB, **Learning by Instruction** *Srivastava, Labutov, Yang, Azaria, Mitchell*

Room 516 CDE, **Infer to Control: Probabilistic Reinforcement Learning and Structured Control** *Kaelbling, Riedmiller, Toussaint, Mordatch, Fox, Haarnoja*

Room 517 A, **Relational Representation Learning** *Grover, Varma, Sala, Holtzen, Neville, Ermon, Ré*

Room 517 B, **AI for social good** *Luck, Sylvain, Cohen, Fansi Tchango, Goddard, Helouis, Bengio, Greydanus, Wild, Kucherenko, Farahi,*

Penn, McGregor, Crowley, Gupta, Chen, Côté, Abebe

Room 517 C, **Reinforcement Learning under Partial Observability** *Pajarinen, Amato, Poupart, Hsu*

Room 517 D, **Machine Learning for Health (ML4H): Moving beyond supervised learning in healthcare** *Beam, Naumann, Ghassemi, McDermott, Fiterau, Chen, Beaulieu-Jones, Hughes, Shamout, Chivers, Kandola, Yahi, Finlayson, Jedynak, Schulam, Antropova, Fries, Dalca, Chen*

Room 518, **Second Workshop on Machine Learning for Creativity and Design** *Elliott, Dieleman, Fiebrink, Engel, Roberts, White*

Room 519, **Machine Learning for Molecules and Materials** *Hernández-Lobato, Müller, Paige, Kusner, Chmiela, Schütt*

Room 524, **Emergent Communication Workshop** *Foerster, Lazaridou, Lowe, Mordatch, Kiela, Cho*

Dec. 7, 2018

Causal Learning

Martin Arjovsky, Christina Heinze-Deml, Anna Klimovskaia, Maxime Oquab, Leon Bottou, David Lopez-Paz

Room 220 C, Fri Dec 07, 08:00 AM

Site for the workshop:
https://sites.google.com/view/nips2018causallearning/home

The route from machine learning to artificial intelligence remains uncharted. Recent efforts describe some of the conceptual problems that lie along this route [4, 9, 12]. The goal of this workshop is to investigate how much progress is possible by framing these problems beyond learning correlations, that is, by uncovering and leveraging causal relations:

1. Machine learning algorithms solve statistical problems (e.g. maximum likelihood) as a proxy to solve tasks of interest (e.g. recognizing objects). Unfortunately, spurious correlations and biases are often easier to learn than the task itself [14], leading to unreliable or unfair predictions. This phenomenon can be framed as causal confounding.

2. Machines trained on large pools of i.i.d. data often crash confidently when deployed in different circumstances (e.g., adversarial examples, dataset biases [18]). In contrast, humans seek prediction rules robust across multiple conditions. Allowing machines to learn robust rules from multiple environments can be framed as searching for causal invariances [2, 11, 16, 17].

3. Humans benefit from discrete structures to reason. Such structures seem less useful to learning machines. For instance, neural machine translation systems outperform those that model language structure. However, the purpose of this structure might not be modeling common sentences, but to help us formulate new ones. Modeling new potential sentences rather than observed ones is a form of counterfactual reasoning [8, 9].

4. Intelligent agents do not only observe, but also shape the world with actions. Maintaining plausible causal models of the world allows to build intuitions, as well as to design intelligent experiments and interventions to test them [16, 17]. Is causal understanding necessary for efficient reinforcement learning?

5. Humans learn compositionally; after learning simple skills, we are able to recombine them quickly to solve new tasks. Such abilities have so far eluded our machine learning systems. Causal models are compositional, so they might offer a solution to this puzzle [4].

6. Finally, humans are able to digest large amounts of unsupervised signals into a causal model of the world. Humans can learn causal affordances, that is, imagining how to manipulate new objects to achieve goals, and the outcome of doing so. Humans rely on a simple blueprint for a complex world: models that contain the correct causal structures, but ignore irrelevant details [16, 17].

We cannot address these problems by simply performing inference on known causal graphs. We need to learn from data to discover plausible causal models, and to construct predictors that are robust to distributional shifts. Furthermore, much prior work has focused on estimating explicit causal structures from data, but these methods are often unscalable, rely on untestable assumptions like faithfulness or acyclicity, and are difficult to incorporate into high-dimensional, complex and nonlinear machine learning pipelines. Instead of considering the task of estimating causal graphs as their final goal, learning machines may use notions from causation indirectly to ignore biases, generalize across distributions, leverage structure to reason, design efficient interventions, benefit from compositionality, and build causal models of the world in an unsupervised way.

Call for papers

Submit your anonymous, NIPS-formatted manuscript here[https://easychair.org/cfp/NIPSCL2018]. All accepted submissions will require a poster presentation. A selection of submissions will be awarded a 5-minute spotlight presentation. We welcome conceptual, thought-provoking material, as well as research agendas, open problems, new tasks, and datasets.

Submission deadline: 28 October 2018
Acceptance notifications: 9 November 2018

Schedule:
See https://sites.google.com/view/nips2018causallearning/home for the up-to-date schedule.

Speakers:
Elias Bareinboim
David Blei
Nicolai Meinshausen
Bernhard Schölkopf
Isabelle Guyon
Csaba Szepesvari
Pietro Perona

References

1. Krzysztof Chalupka, Pietro Perona, Frederick Eberhardt (2015): Visual Causal Feature Learning [https://arxiv.org/abs/1412.2309]
2. Christina Heinze-Deml, Nicolai Meinshausen (2018): Conditional Variance Penalties and Domain Shift Robustness [https://arxiv.org/abs/1710.11469]
3. Fredrik D. Johansson, Uri Shalit, David Sontag (2016): Learning Representations for Counterfactual Inference [https://arxiv.org/abs/1605.03661]
4. Brenden Lake (2014): Towards more human-like concept learning in machines: compositionality, causality, and learning-to-learn [https://dspace.mit.edu/handle/1721.1/95856]
5. Brenden M. Lake, Tomer D. Ullman, Joshua B. Tenenbaum, Samuel J. Gershman (2016): Building Machines That Learn and Think Like People [https://arxiv.org/abs/1604.00289]

6. David Lopez-Paz, Krikamol Muandet, Bernhard Schölkopf, Ilya Tolstikhin (2015): Towards a Learning Theory of Cause-Effect Inference [https://arxiv.org/abs/1309.6779]
7. David Lopez-Paz, Robert Nishihara, Soumith Chintala, Bernhard Schölkopf, Léon Bottou (2017): Discovering Causal Signals in Images [https://arxiv.org/abs/1605.08179]
8. Judea Pearl (2009): Causality: Models, Reasoning, and Inference [http://bayes.cs.ucla.edu/BOOK-2K/]
9. Judea Pearl (2018): The Seven Pillars of Causal Reasoning with Reflections on Machine Learning [http://ftp.cs.ucla.edu/pub/stat_ser/r481.pdf]
10. Jonas Peters, Joris Mooij, Dominik Janzing, Bernhard Schölkopf (2014): Causal Discovery with Continuous Additive Noise Models [https://arxiv.org/abs/1309.6779]
11. Jonas Peters, Peter Bühlmann, Nicolai Meinshausen (2016): Causal inference using invariant prediction: identification and confidence intervals [https://arxiv.org/abs/1501.01332]
12. Jonas Peters, Dominik Janzing, Bernhard Schölkopf (2017): Elements of Causal Inference: Foundations and Learning Algorithms [https://mitpress.mit.edu/books/elements-causal-inference]
13. Peter Spirtes, Clark Glymour, Richard Scheines (2001): Causation, Prediction, and Search [http://cognet.mit.edu/book/causation-prediction-and-search]
14. Bob L. Sturm (2016): The HORSE conferences [http://c4dm.eecs.qmul.ac.uk/horse2016/, http://c4dm.eecs.qmul.ac.uk/horse2017/]
15. Dustin Tran, David M. Blei (2017): Implicit Causal Models for Genome-wide Association Studies [https://arxiv.org/abs/1710.10742]
16. Michael Waldmann (2017): The Oxford Handbook of Causal Reasoning [https://global.oup.com/academic/product/the-oxford-handbook-of-causal-reasoning-9780199399550?cc=us&lang=en&]
17. James Woodward (2005): Making Things Happen: A Theory of Causal Explanation [https://global.oup.com/academic/product/making-things-happen-9780195189537?cc=us&lang=en&]
18. Antonio Torralba, Alyosha Efros (2011): Unbiased look at dataset bias. [http://people.csail.mit.edu/torralba/publications/datasets_cvpr11.pdf]

Schedule

09:15 AM	**Opening Remarks**	*Lopez-Paz*
09:30 AM	**Learning Independent Mechanisms**	*Schölkopf*
10:00 AM	**From Micro-Variables to Macro-Causes**	*Perona*
11:00 AM	**Causality and Transfer Learning**	*Bareinboim*
11:30 AM	**Evaluating Causation Coefficients**	*Guyon*
12:00 PM	**The Blessings of Multiple Causes**	*Blei*
02:00 PM	**Causality and Distributional Robustness**	*Meinshausen*
02:30 PM	**Model-free vs. Model-based Learning in a Causal World: Some Stories from Online Learning to Rank**	*Szepesvari*
03:30 PM	**Datasets and Benchmarks for Causal Learning**	*Szepesvari, Guyon, Meinshausen, Blei, Bareinboim, Schölkopf, Perona*
05:00 PM	**Woulda, Coulda, Shoulda: Counterfactually-Guided Policy Search**	
05:05 PM	**Cause-Effect Deep Information Bottleneck For Incomplete Covariates**	*Wieser*
05:10 PM	**NonSENS: Non-Linear SEM Estimation using Non-Stationarity**	
05:15 PM	**Rule-Based Sentence Quality Modeling and Assessment using Deep LSTM Features**	*Pagalla*
05:20 PM	**The Case for Evaluating Causal Models Using Interventional Measures and Empirical Data**	*Gentzel*
05:25 PM	**Marginalized Off-Policy Evaluation for Reinforcement Learning**	*Ma*
05:30 PM	**Domain Adaptation by Using Causal Inference to Predict Invariant Conditional Distributions**	
05:35 PM	**Learning Predictive Models That Transport**	
05:40 PM	**Causal Confusion in Imitation Learning**	
05:45 PM	**Causality in Physics and Effective Theories of Agency**	
06:00 PM	**A Bayesian Solution to the M-Bias Problem**	*Rohde*
06:00 PM	**Real Defect Image Classification through Hierarchical Data Augmentation**	*Kim*
06:00 PM	**Stability of Linear Structural Equation Model of Causal Inference**	
06:00 PM	**Rule Discovery for Exploratory Causal Reasoning**	
06:00 PM	**Quantifying Context-Dependent Causal Influences**	*Coleman*
06:00 PM	**Entropic Latent Variable Discovery**	

06:00 PM	**Intervention Harvesting for Context-Dependent Examination-Bias Estimation**	*Agarwal*
06:00 PM	**Learning Robotic Manipulation through Visual Planning and Acting**	
06:00 PM	**Off-policy learning for causal advertising**	
06:00 PM	**Detecting switching causal interactions using hierarchical segmentation approach**	*Briones*
06:00 PM	**Consistency and Computation for Regularized Maximum Likelihood Estimation of Multivariate Hawkes Processes**	*Hu*
06:00 PM	**Modularization of End-to-End Learning: Case Study in Arcade Games**	*Melnik*
06:00 PM	**Weighted Tensor Completion for Time-Series Causal Inference**	*Mandal*
06:00 PM	**Stochastic Complexity for Testing Conditional Independence on Discrete Data**	*Marx*
06:00 PM	**CAB: Continuous Adaptive Blending Estimator for Policy Evaluation and Learning**	*Su, Wang*

Abstracts (1):

Abstract 1: **Opening Remarks in Causal Learning**, *Lopez-Paz* 09:15 AM

Speaker: David Lopez-Paz

Bayesian Deep Learning

Yarin Gal, Jose Miguel Hernández-Lobato, Christos Louizos, Andrew Wilson, Zoubin Ghahramani, Kevin P Murphy, Max Welling

Room 220 D, Fri Dec 07, 08:00 AM

While deep learning has been revolutionary for machine learning, most modern deep learning models cannot represent their uncertainty nor take advantage of the well studied tools of probability theory. This has started to change following recent developments of tools and techniques combining Bayesian approaches with deep learning. The intersection of the two fields has received great interest from the community over the past few years, with the introduction of new deep learning models that take advantage of Bayesian techniques, as well as Bayesian models that incorporate deep learning elements [1-11]. In fact, the use of Bayesian techniques in deep learning can be traced back to the 1990s', in seminal works by Radford Neal [12], David MacKay [13], and Dayan et al. [14]. These gave us tools to reason about deep models' confidence, and achieved state-of-the-art performance on many tasks. However earlier tools did not adapt when new needs arose (such as scalability to big data), and were consequently forgotten. Such ideas are now being revisited in light of new advances in the field, yielding many exciting new results.

Extending on the workshop's success from the past couple of years, this workshop will again study the advantages and disadvantages of the ideas above, and will be a platform to host the recent flourish of ideas using Bayesian approaches in deep learning and using deep learning tools in Bayesian modelling. The program includes a mix of invited talks, contributed talks, and contributed posters. The main theme this year will be applications of Bayesian deep learning in the real world, highlighting the requirements of practitioners from the research community. Future directions for the field will be debated in a panel discussion.

The BDL workshop was the second largest workshop at NIPS over the past couple of years, with last year's workshop seeing an almost 100% increase in the number of submissions (75 submissions in total), attracting sponsorship from Google, Microsoft Ventures, Uber, and Qualcomm in the form of student travel awards.

Topics:
Probabilistic deep models for classification and regression (such as extensions and application of Bayesian neural networks),
Generative deep models (such as variational autoencoders),
Incorporating explicit prior knowledge in deep learning (such as posterior regularization with logic rules),
Approximate inference for Bayesian deep learning (such as variational Bayes / expectation propagation / etc. in Bayesian neural networks),
Scalable MCMC inference in Bayesian deep models,
Deep recognition models for variational inference (amortized inference),
Model uncertainty in deep learning,
Bayesian deep reinforcement learning,
Deep learning with small data,
Deep learning in Bayesian modelling,
Probabilistic semi-supervised learning techniques,
Active learning and Bayesian optimization for experimental design,
Applying non-parametric methods, one-shot learning, and Bayesian deep learning in general,
Implicit inference,
Kernel methods in Bayesian deep learning.

Call for papers:
A submission should take the form of an extended abstract (3 pages long) in PDF format using the NIPS style. Author names do not need to be anonymized and references (as well as appendices) may extend as far as needed beyond the 3 page upper limit. If research has previously appeared in a journal, workshop, or conference (including NIPS 2017 conference), the workshop submission should extend that previous work. Submissions will be accepted as contributed talks or poster presentations.

Related previous workshops:
Bayesian Deep Learning (NIPS 2017)
Principled Approaches to Deep Learning (ICML 2017)
Bayesian Deep Learning (NIPS 2016)
Data-Efficient Machine Learning (ICML 2016)
Deep Learning Workshop (ICML 2015, 2016)
Deep Learning Symposium (NIPS 2015 symposium)
Advances in Approximate Bayesian Inference (NIPS 2015)
Black box learning and inference (NIPS 2015)
Deep Reinforcement Learning (NIPS 2015)
Deep Learning and Representation Learning (NIPS 2014)
Advances in Variational Inference (NIPS 2014)

Schedule

08:00 AM	**Opening Remarks**	*Gal*
08:05 AM	**TBC 1**	*Wood*
08:25 AM	**TBC 2**	*Vetrov*
08:45 AM	**TBC 3**	
09:00 AM	**TBC 4**	*Marks*
09:20 AM	**TBC 5**	*Valpola*
09:40 AM	**Poster Spotlights**	
09:55 AM	**Poster Session 1**	*Gadatsch, Kuzin, Kumar, Dallaire, Ryder, Pop, Hunt, Kortylewski, Burkhardt, Elnaggar, Lawson, Li, Ryu, Bae, Livne, Pearce, Vladimirova, Ramapuram, Zeng, Hu, He, Maddix, Mittal, Shaw, Le, Sagel, Chen, Gallego, Karami, Zhang, Kachman, Weber, Benatan, Sricharan, Cartillier, Ovinnikov, Phan, Hossam, Liu, Kharitonov, Golikov, Zhang, Kim, Farquhar, Mukhoti, Hu, Gundersen, Tekumalla, Perdikaris, Banijamali, Jain, Liu, Gottwald, Blumer, Yun, Krishnan, Novak, Du, Gong, Gokkaya, Ai, Duckworth, von Oswald, Henning, Morency, Ghodsi, Subedar, Pfister, Lebret, Ma, Wieczorek, Perreault Levasseur*
10:55 AM	**TBC 6**	*Leibig*
11:15 AM	**TBC 7**	
11:30 AM	**TBC 8**	*Lakshminarayanan*
11:50 AM	**Lunch**	
01:20 PM	**TBC 9**	*Levine*
01:40 PM	**TBC 10**	
01:55 PM	**TBC 11**	*Hezaveh*
02:10 PM	**TBC 12**	*Genewein*
02:30 PM	**Poster Session 2**	
03:30 PM	**TBC 13**	*Sontag*
03:50 PM	**TBC 14**	
04:05 PM	**TBC 15**	*Gal*
04:30 PM	**Panel Session**	
05:30 PM	**Poster Session 3**	

Abstracts (1):

Abstract 1: **Opening Remarks in Bayesian Deep Learning**, *Gal* 08:00 AM

Introductory comments by the organisers.

Deep Reinforcement Learning

Pieter Abbeel, David Silver, Satinder Singh, Joelle Pineau, Joshua Achiam, Rein Houthooft, Aravind Srinivas

Room 220 E, Fri Dec 07, 08:00 AM

In recent years, the use of deep neural networks as function approximators has enabled researchers to extend reinforcement learning techniques to solve increasingly complex control tasks. The emerging field of deep reinforcement learning has led to remarkable empirical results in rich and varied domains like robotics, strategy games, and multiagent interaction. This workshop will bring together researchers working at the intersection of deep learning and reinforcement learning, and it will help interested researchers outside of the field gain a high-level view about the current state of the art and potential directions for future contributions.

Schedule

09:00 AM	**Talk by Yann Lecunn**
09:30 AM	**Contrbuted Talks**
10:00 AM	**Talk by Jacob Andreas**
10:30 AM	**Coffee**
11:00 AM	**Talk by Sham Kakade**
12:00 PM	**Talk by Doina Precup**
12:30 PM	**Lunch**
01:30 PM	**Talk by Satinder Singh**
02:00 PM	**Contributed Talks**
02:30 PM	**Talk by Martha White**

03:00 PM	**Poster Session 1 + Coffee**	*Van de Wiele, Zhao, Hernandez-Garcia, Pardo, Lee, Li, Andrychowicz, Tang, Nair, Lee, Colas, Eslami, Wu, McAleer, Julian, Xue, Sabatelli, Shyam, Kalousis, Montana, Pesce, Leibfried, He, Liu, Li, Sawada, Pashevich, Kulkarni, Paster, Rigazio, Vuong, Park, Kwon, Weerasekera, Siriwardhanaa, Wang, Kilinc, Ross, Wang, Schmitt, Anthony, Cater, Agostinelli, Sung, Maruyama, Shmakov, Schwab, Firouzi, Berseth, Osipychev, Farebrother, Luo, Agnew, Vrancx, Heek, Ionescu, Yin, Miyashita, Jay, Rotman, Leroux, Bojja Venkatakrishnan, Schmidt, Terwilliger, Durugkar, Sauder, Kas, Tavakoli, Cohen, Bontrager, Lerer, Paine, Khalifa, Rodriguez, Singh, Zhang*
04:00 PM	**Talk by Jeff Clune**	
04:30 PM	**Contributed Talks**	
05:15 PM	**Poster Session 2**	

Abstracts (1):

Abstract 2: **Contrbuted Talks in Deep Reinforcement Learning**, 09:30 AM

TBD

MLSys: Workshop on Systems for ML and Open Source Software

Aparna Lakshmiratan, Sarah Bird, Siddhartha Sen, Joseph Gonzalez, Dan Crankshaw

Room 510 ABCD, Fri Dec 07, 08:00 AM

This workshop is part two of a two-part series with one day focusing on ML for Systems and the other on Systems for ML. Although the two workshops are being led by different organizers, we are coordinating our call for papers to ensure that the workshops complement each other and that submitted papers are routed to the appropriate venue.

The ML for Systems workshop focuses on developing ML to optimize systems while we focus on designing systems to enable large scale ML with Systems for ML. Both fields are mature enough to warrant a dedicated workshop. Organizers on both sides are open to merging in the future, but this year we plan to run them separately on two different days.

A new area is emerging at the intersection of artificial intelligence, machine learning, and systems design. This has been accelerated by the explosive growth of diverse applications of ML in production, the continued growth in data volume, and the complexity of large-scale learning systems. The goal of this workshop is to bring together experts working at the crossroads of machine learning, system design and software engineering to explore the challenges faced when building large-scale ML systems. In particular, we aim to elicit new connections among these diverse fields, identifying theory, tools and design principles tailored to practical machine learning workflows. We also want to think about best practices for research in this area and how to evaluate it. The workshop will cover state of the art ML and AI platforms and algorithm toolkits (e.g. TensorFlow, PyTorch1.0, MXNet etc.), as well as dive into machine learning-focused developments in distributed learning platforms, programming languages, data structures, GPU processing, and other topics.

This workshop will follow the successful model we have previously run at ICML, NIPS and SOSP 2017.

Our plan is to run this workshop annually co-located with one ML venue and one Systems venue, to help build a strong community which we think will complement newer conferences like SysML targeting research at the intersection of systems and machine learning. We believe this dual approach will help to create a low barrier to participation for both communities.

Schedule

09:00 AM	**Welcome**	*Bird*
09:10 AM	**Invited Talk (Bryan Catanzaro, NVidia)**	
09:40 AM	**Fashionable modeling with Flux**	
10:00 AM	**Model assertions for debugging machine learning**	
10:20 AM	**Poster Intro by OC + Poster Session (Chair: Sid Sen)**	
11:40 AM	**Keynote 2: “Machine Learning at Netflix” (Aish Fenton)**	
12:10 PM	**Parallel training of linear models**	
12:30 PM	**Lunch provided and Open Source ML Systems Showcase (TensorFlow, PyTorch 1.0, MxNET, Keras, CoreML, Ray, Chainer)**	*Monga, Chintala, Moreau, Chollet, Crankshaw, Nishihara, Tokui*

02:55 PM	**Posters (all accepted papers) + Break**	*Wang, Gudovskiy, Jiang, Kaufmann, Anghel, Bradbury, Ioannou, Agrawal, Tosch, Yu, Fischer, Revels, Siracusano, Yang, Johnson, You, Yuen, Ying, Liu, Dryden, Mo, Wang, Juneja, Smith, Yu, gupta, Narayanan, Santhanam, Capes, Dakkak, Mu, Deng, Li, Carreira, Remis, Raghavan, O'Reilly, Singh, Assran, Wu, Bakshy, Wei, Innes, Shah, Lin, Sanderson, Curtin, Edel*
03:40 PM	**Keynote 3: "Infrastructure and Systems for Applied Machine Learning at Facebook" (Kim Hazelwood)**	
04:10 PM	**HiveMind: Accelerating Deep Learning Workloads through Efficient Multi-Model Execution**	
04:30 PM	**Rethinking floating point for deep learning**	
04:50 PM	**A Case for Serverless Machine Learning**	
05:10 PM	**Closing Remarks**	*Lakshmiratan*

Critiquing and Correcting Trends in Machine Learning

Tom Rainforth, Matt Kusner, Ben Bloem-Reddy, Brooks Paige, Rich Caruana, Yee Whye Teh

Room 511 ABDE, Fri Dec 07, 08:00 AM

Workshop Webpage:
https://ml-critique-correct.github.io/

Recently there have been calls to make machine learning more reproducible, less hand-tailored, fair, and generally more thoughtful about how research is conducted and put into practice. These are hallmarks of a mature scientific field and will be crucial for machine learning to have the wide-ranging, positive impact it is expected to have. Without careful consideration, we as a field risk inflating expectations beyond what is possible. To address this, this workshop aims to better understand and to improve all stages of the research process in machine learning.

A number of recent papers have carefully considered trends in machine learning as well as the needs of the field when used in real-world scenarios [1-18]. Each of these works introspectively analyzes what we often take for granted as a field. Further, many propose solutions for moving forward. The goal of this workshop is to bring together researchers from all subfields of machine learning to highlight open problems and widespread dubious practices in the field, and crucially, to propose solutions. We hope to highlight issues and propose solutions in areas such as:

- Common practices [1, 8]

- Implicit technical and empirical assumptions that go unquestioned [2, 3, 5, 7, 11, 12, 13, 17, 18]

- Shortfalls in publication and reviewing setups [15, 16]

- Disconnects between research focus and application requirements [9, 10, 14]

- Surprising observations that make us rethink our research priorities [4, 6]

The workshop program is a collection of invited talks, alongside contributed posters and talks. For some of these talks, we plan a unique open format of 10 minutes of talk + 10 minutes of follow up discussion. Additionally, a separate panel discussion will collect researchers with a diverse set of viewpoints on the current challenges and potential solutions. During the panel, we will also open the conversation to the audience. The discussion will further be open to an online Q&A which will be solicited prior to the workshop.

A key expected outcome of the workshop is a collection of important open problems at all levels of machine learning research, along with a record of various bad practices that we should no longer consider to be acceptable. Further, we hope that the workshop will make inroads in how to address these problems, highlighting promising new frontiers for making machine learning practical, robust, reproducible, and fair when applied to real-world problems.

Call for Papers:

Deadline: October 30rd, 2018, 11:59 UTC

The one day NIPS 2018 Workshop: Critiquing and Correcting Trends in Machine Learning calls for papers that critically examine current common practices and/or trends in methodology, datasets, empirical standards, publication models, or any other aspect of machine learning research. Though we are happy to receive papers that bring attention to problems for which there is no clear immediate remedy, we particularly encourage papers which propose a solution or indicate a way forward. Papers should motivate their arguments by describing gaps in the field. Crucially, this is not a venue for settling scores or character attacks, but for moving machine learning forward as a scientific discipline.

To help guide submissions, we have split up the call for papers into the follows tracks. Please indicate the intended track when making your submission. Papers are welcome from all subfields of machine learning. If you have a paper which you feel falls within the remit of the workshop but does not clearly fit one of these tracks, please contact the organizers at: ml.critique.correct@gmail.com.

Bad Practices (1-4 pages)

Papers that highlight common bad practices or unjustified assumptions at any stage of the research process. These can either be technical shortfalls in a particular machine learning subfield, or more procedural bad practices of the ilk of those discussed in [17].

Flawed Intuitions or Unjustified Assumptions (3-4 pages)

Papers that call into question commonly held intuitions or provide clear evidence either for or against assumptions that are regularly taken for granted without proper justification. For example, we would like to see papers which provide empirical assessments to test out metrics, verify intuitions, or compare popular current approaches with historic baselines that may have unfairly fallen out of favour (see e.g. [2]). We would also like to see work which provides results which makes us rethink our intuitions or the assumptions we typically make.

Negative Results (3-4 pages)

Papers which show failure modes of existing algorithms or suggest new approaches which one might expect to perform well but which do not. The aim of the latter of these is to provide a venue for work which might otherwise go unpublished but which is still of interest to the community, for example by dissuading other researchers from similar ultimately unsuccessful approaches. Though it is inevitably preferable that papers are able to explain why the approach performs poorly, this is not essential if the paper is able to demonstrate why the negative result is of interest to the community in its own right.

Research Process (1-4 pages)

Papers which provide carefully thought through critiques, provide discussion on, or suggest new approaches to areas such as the conference model, the reviewing process, the role of industry in research, open sourcing of code and data, institutional biases and discrimination in the field, research ethics, reproducibility standards, and allocation of conference tickets.

Debates (1-2 pages)

Short proposition papers which discuss issues either affecting all of machine learning or significantly sized subfields (e.g. reinforcement learning, Bayesian methods, etc). Selected papers will be used as the basis for instigating online forum debates before the workshop, leading up to live discussions on the day itself.

Open Problems (1-4 papers/short talks)

Papers that describe either (a) unresolved questions in existing fields that need to be addressed, (b) desirable operating characteristics for ML in particular application areas that have yet to be achieved, or (c) new frontiers of machine learning research that require rethinking current practices (e.g., error diagnosis for when many ML components are interoperating within a system, automating dataset collection/creation).

Submission Instructions
Papers should be submitted as pdfs using the NIPS LaTeX style file. Author names should be anonymized.

All accepted papers will be made available through the workshop website and presented as a poster. Selected papers will also be given contributed talks. We have a small number of complimentary workshop registrations to hand out to students. If you would like to apply for one of these, please email a one paragraph supporting statement. We also have a limited number of reserved tickets slots to assign to authors of accepted papers. If any authors are unable to attend the workshop due to ticketing, visa, or funding issues, they will be allowed to provide a video presentation for their work that will be made available through the workshop website in lieu of a poster presentation.

Please submit papers here:
[https://easychair.org/conferences/?conf=cract2018](https://easychair.org/conferences/

Deadline: October 30rd, 2018, 11:59 UTC

References

[1] Mania, H., Guy, A., & Recht, B. (2018). Simple random search provides a competitive approach to reinforcement learning. arXiv preprint arXiv:1803.07055.
[2] Rainforth, T., Kosiorek, A. R., Le, T. A., Maddison, C. J., Igl, M., Wood, F., & Teh, Y. W. (2018). Tighter variational bounds are not necessarily better. ICML.
[3] Torralba, A., & Efros, A. A. (2011). Unbiased look at dataset bias. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on (pp. 1521-1528). IEEE.
[4] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199.
[5] Mescheder, L., Geiger, A., Nowozin S. (2018) Which Training Methods for GANs do actually Converge? ICML
[6] Daumé III, H. (2009). Frustratingly easy domain adaptation. arXiv preprint arXiv:0907.1815
[7] Urban, G., Geras, K. J., Kahou, S. E., Wang, O. A. S., Caruana, R., Mohamed, A., ... & Richardson, M. (2016). Do deep convolutional nets really need to be deep (or even convolutional)?.
[8] Henderson, P., Islam, R., Bachman, P., Pineau, J., Precup, D., & Meger, D. (2017). Deep reinforcement learning that matters. arXiv preprint arXiv:1709.06560.
[9] Narayanan, M., Chen, E., He, J., Kim, B., Gershman, S., & Doshi-Velez, F. (2018). How do Humans Understand Explanations from Machine Learning Systems? An Evaluation of the Human-Interpretability of Explanation. arXiv preprint arXiv:1802.00682.
[10] Schulam, S., Saria S. (2017). Reliable Decision Support using Counterfactual Models. NIPS.
[11] Rahimi, A. (2017). Let's take machine learning from alchemy to electricity. Test-of-time award presentation, NIPS.
[12] Lucic, M., Kurach, K., Michalski, M., Gelly, S., Bousquet, O. (2018). Are GANs Created Equal? A Large-Scale Study. arXiv preprint arXiv:1711.10337.
[13] Le, T.A., Kosiorek, A.R., Siddharth, N., Teh, Y.W. and Wood, F., (2018). Revisiting Reweighted Wake-Sleep. arXiv preprint arXiv:1805.10469.
[14] Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J. and Mané, D., (2016). Concrete problems in AI safety. arXiv preprint arXiv:1606.06565.
[15] Sutton, C. (2018) Making unblinding manageable: Towards reconciling prepublication and double-blind review. http://www.theexclusive.org/2017/09/arxiv-double-blind.html
[16] Langford, J. (2018) ICML Board and Reviewer profiles. http://hunch.net/?p=8962378

Schedule

08:30 AM	**Opening Remarks**	
08:40 AM	**Zachary Lipton**	*Lipton*

09:05 AM	**Kim Hazelwood**	*Hazelwood*
09:30 AM	**Expanding search in the space of empirical ML**	*Woods*
09:40 AM	**Opportunities for machine learning research to support fairness in industry practice**	*Holstein*
09:50 AM	**Spotlights - Papers 2, 23, 24, 36, 40, 44**	
10:20 AM	**Poster Session 1 (note there are numerous missing names here, all papers appear in all poster sessions)**	*Gotmare, Holstein, Brabec, Uricar, Clary, Rudin, Witty, Ross, O'Brien, Esmaeili, Forde, Caccia, Emami, Jordan, Woods, Sculley, Overdorf, Le Roux, Henderson, Yang, Liu, Jensen, Dalmasso, Liu, TRICHELAIR, Lee, Atrey, Groh, Hechtlinger, Tosch*
11:10 AM	**Finale Doshi-Velez**	*Doshi-Velez*
11:35 AM	**Suchi Saria**	*Saria*
12:00 PM	**Lunch**	
01:30 PM	**Sebastian Nowozin**	*Nowozin*
01:55 PM	**Using Cumulative Distribution Based Performance Analysis to Benchmark Models**	*Jordan*
02:05 PM	**Charles Sutton**	*Sutton*
02:30 PM	**On Avoiding Tragedy of the Commons in the Peer Review Process**	*Sculley*
02:40 PM	**Spotlights - Papers 10, 20, 35, 42**	
03:00 PM	**Coffee Break and Posters**	
03:30 PM	**Panel on research process**	*Lipton, Sutton, Doshi-Velez, Wallach, Saria, Caruana, Rainforth*
04:30 PM	**Poster Session 2**	

Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy

Manuela Veloso, Nathan Kallus, Sameena Shah, Senthil Kumar, Isabelle Moulinier, Jiahao Chen, John Paisley

Room 511 CF, Fri Dec 07, 08:00 AM

The adoption of artificial intelligence in the financial service industry, particularly the adoption of machine learning, presents challenges and opportunities. Challenges include algorithmic fairness, explainability, privacy, and requirements of a very high degree of accuracy. For example, there are ethical and regulatory needs to prove that models used for activities such as credit decisioning and lending are fair and unbiased, or that machine reliance doesn't cause humans to miss critical pieces of data. For some use cases, the operating standards require nothing short of perfect accuracy.

Privacy issues around collection and use of consumer and proprietary data require high levels of scrutiny. Many machine learning models are deemed unusable if they are not supported by appropriate levels of explainability. Some challenges like entity resolution are exacerbated because of scale, highly nuanced data points and missing information. On top of these fundamental requirements, the financial industry is ripe with adversaries who purport fraud and other types of risks.

The aim of this workshop is to bring together researchers and practitioners to discuss challenges for AI in financial services, and the opportunities such challenges represent to the community. The workshop will consist of a series of sessions, including invited talks, panel discussions and short paper presentations, which will showcase ongoing research and novel algorithms.

Schedule

08:35 AM	**Opening Remarks**	*Veloso, Moulinier*
08:50 AM	**Invited Talk 1: Fairness and Causality with Missing Data**	*Udell*
09:10 AM	**Invited Talk 2: Building Augmented Intelligence for a Global Credit Rating Agency**	*Shah*
09:30 AM	**Panel: Explainability, Fairness and Human Aspects in Financial Services**	*Udell, Chen, Mekel-Bobrov, Veloso, Kleinberg, Freeman, Chandarana, Sisk, McBurnett*
10:30 AM	**Coffee Break and Socialization**	
11:00 AM	**Invited Talk 3: Fairness in Allocation Problems**	*Kearns*
11:20 AM	**Paper Presentations (see below for paper titles)**	
12:00 PM	**Lunch**	
01:30 PM	**Invited Talk 4: When Algorithms Trade: Modeling AI in Financial Markets**	*Wellman*
01:50 PM	**Invited Talk 5: ML-Based Evidence that High Frequency Trading Has Made the Market More Efficient**	*Balch*
02:10 PM	**Paper Presentations (see below for paper titles)**	
02:50 PM	**Announcement: FICO XAI Challenge Winners**	
03:00 PM	**Coffee Break**	

03:30 PM	**Invited Talk 6: Is it possible to have interpretable models for AI in Finance?**	*Rudin*
03:50 PM	**Paper Presentations (see below for paper titles)**	
04:30 PM	**Posters and Open Discussions (see below for poster titles)**	*Malur Srinivasan, Perez, Liu, Wood, Philps, Brown, Martin, Pechenizkiy, Costabello, Wang, Sarkar, Yoon, Xiong, Horel, Zhang, Johansson, Kochems, Sidier, Reddy, Cuthbertson, Wambui, Marfaing, Harrison, Unceta Mendieta, Kehler, Weber, Ling, Modarres, Dhall, Nourian, Byrd, Chander, Liu, Yang, Zhai, Lecue, Yao, McGrath, Garcez, Bacoyannis, Garcia, Gonon, Ibrahim, Louie, Ardakanian, Sönströd, Oshiba, Chen, Jin, pareja, Suzumura*
05:40 PM	**Closing Remarks**	

Abstracts (4):

Abstract 7: **Paper Presentations (see below for paper titles) in Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy**, 11:20 AM

11:20 - 11:28 In (Stochastic) Search of a Fairer Alife
11:28 - 11:36 Where's the Bias? Developing Effective Model Goveranance
11:36 - 11:44 Scalable Graph Learning for Anti-Money Laundering: A First Look
11:44 - 11:52 Fair Resource Allocation in a Volatile Marketplace
11:52 - 12:00 Algorithmic Confidence – A Key Criterion for XAI and FAT

Abstract 11: **Paper Presentations (see below for paper titles) in Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy**, 02:10 PM

02:10 - 02:18: Accurate, Data-Efficient Learning from Noisy, Choice-Based Labels for Inherent Risk Scoring
02:18 - 02:26: An AI-based, Multi-stage detection system of banking botnets
02:26 - 02:34: Robust Classification of Financial Risk
02:34 - 02:42: An Empirical Evaluation of Deep Sequential Models for Volatility Prediction
02:42 - 02:50: Computer-Assisted Fraud Detection, from Active Learning to Reward Maximization

Abstract 15: **Paper Presentations (see below for paper titles) in Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy**, 03:50 PM

15:50 - 15:58 Use of Machine Learning Techniques to Create a Credit Score Model for Prepaid Basic Services in East Africa. Case Study: Airtime Loans
15:58 - 16:06 Towards Global Explanations for Credit Risk Scoring
16:06 - 16:14 Interpretable Credit Application Predictions With Counterfactual Explanations
16:14 - 16:22 Interpretable Feature Selection Using Local Information for Credit Assessment
16:22 - 16:30 Towards Explainable Deep Learning for Credit Lending: A Case Study

Abstract 16: **Posters and Open Discussions (see below for poster titles) in Challenges and Opportunities for AI in Financial Services: the Impact of Fairness, Explainability, Accuracy, and Privacy**, *Malur Srinivasan, Perez, Liu, Wood, Philps, Brown, Martin, Pechenizkiy, Costabello, Wang, Sarkar, Yoon, Xiong, Horel, Zhang, Johansson, Kochems, Sidier, Reddy, Cuthbertson, Wambui, Marfaing, Harrison, Unceta Mendieta, Kehler, Weber, Ling, Modarres, Dhall, Nourian, Byrd, Chander, Liu, Yang, Zhai, Lecue, Yao, McGrath, Garcez, Bacoyannis, Garcia, Gonon, Ibrahim, Louie, Ardakanian, Sönströd, Oshiba, Chen, Jin, pareja, Suzumura* 04:30 PM

1. Clustering and Learning from Imbalanced Data
2. Deep Hedging: Hedging Derivatives Under Generic Market Frictions Using Reinforcement Learning
3. Generating User-friendly Explanations for Loan Denials using GANs
4. Practical Deep Reinforcement Learning Approach for Stock Trading
5. Idiosyncrasies and challenges of data driven learning in electronic trading
6. Machine learning-aided modeling of fixed income instruments
7. An Interpretable Model with Globally Consistent Explanations for Credit Risk
8. Continuous learning augmented investment decisions
9. HELOC Applicant Risk Performance Evaluation by Topological Hierarchical Decomposition
10. Looking Deeper into the Deep Learning Models: Attribution-based Explanations of TextCNN
11. Matrix Regression and Its Applications in Cryptocurrency Trading
12. Sensitivity based Neural Networks Explanations
13. On the Need for Fairness in Financial Recommendation Engines
14. Read the News, not the Books: Predicting Firms' Financial Health

Smooth Games Optimization and Machine Learning

Simon Lacoste-Julien, Ioannis Mitliagkas, Gauthier Gidel, Vasilis Syrgkanis, Eva Tardos, Leon Bottou, Sebastian Nowozin

Room 512 ABEF, Fri Dec 07, 08:00 AM

Overview

Advances in generative modeling and adversarial learning gave rise to a recent surge of interest in smooth two-players games, specifically in the context of learning generative adversarial networks (GANs). Solving these games raise intrinsically different challenges than the minimization tasks the machine learning community is used to. The goal of this workshop is to bring together the several communities interested in such smooth games, in order to present what is known on the topic and identify current open questions, such as how to handle the non-convexity appearing in GANs.

Background and objectives

A number of problems and applications in machine learning are formulated as games. A special class of games, smooth games, have come into the spotlight recently with the advent of GANs. In a two-players smooth game, each player attempts to minimize their differentiable cost function which depends also on the action of the other player. The dynamics of such games are distinct from the better understood dynamics of optimization problems. For example, the Jacobian of gradient descent on a smooth two-player game, can be non-symmetric and have complex eigenvalues. Recent work by ML researchers has identified these dynamics as a key challenge for efficiently solving similar problems.

A major hurdle for relevant research in the ML community is the lack of interaction with the mathematical programming and game theory communities where similar problems have been tackled in the past, yielding useful tools. While ML researchers are quite familiar with the convex optimization toolbox from mathematical programming, they are less familiar with the tools for solving games. For example, the extragradient algorithm to solve variational inequalities has been known in the mathematical programming literature for decades, however the ML community has until recently mainly appealed to gradient descent to optimize adversarial objectives.

The aim of this workshop is to provide a platform for both theoretical and applied researchers from the ML, mathematical programming and game theory community to discuss the status of our understanding on the interplay between smooth games, their applications in ML, as well existing tools and methods for dealing with them.
We also encourage, and will devote time during the workshop, on work that identifies and discusses open, forward-looking problems of interest to the NIPS community.

Examples of topics of interest to the workshop are as follow:
* Other examples of smooth games in machine learning (e.g. actor-critic models in RL).
* Standard or novel algorithms to solve smooth games.
* Empirical test of algorithms on GAN applications.
* Existence and unicity results of equilibria in smooth games.
* Can approximate equilibria have better properties than the exact ones ? [Arora 2017, Lipton and Young 1994].
* Variational inequality algorithms [Harker and Pang 1990, Gidel et al. 2018].
* Handling stochasticity [Hazan et al. 2017] or non-convexity [Grnarova et al. 2018] in smooth games.
* Related topics from mathematical programming (e.g. bilevel optimization) [Pfau and Vinyals 2016].

Schedule

08:30 AM	**Opening remarks**	*Lacoste-Julien, Gidel*
08:50 AM	**Improving Generative Adversarial Networks using Game Theory and Statistics**	*Daskalakis*
09:30 AM	**Poster spotlight**	*Yang, Dvurechenskii, Mertikopoulos, Berard*
10:00 AM	**Poster session**	
10:30 AM	**Morning coffee break**	
11:00 AM	**Smooth Games in Machine Learning Beyond GANs**	*He*
11:40 AM	**Finding Mixed Nash Equilibria of Generative Adversarial Networks**	*Cevher*
12:00 PM	**Bounding Inefficiency of Equilibria in Continuous Actions Games using Submodularity and Curvature**	*Sessa*
12:20 PM	**Lunch break**	
02:00 PM	**Building Algorithms by Playing Games**	*Abernethy*
02:40 PM	**Negative Momentum for Improved Game Dynamics**	*Askari Hemmat*
03:00 PM	**Afternoon coffee break**	
03:30 PM	**Regret Decomposition in Sequential Games with Convex Action Spaces and Losses**	*Farina*
03:50 PM	**An interpretation of GANs via online learning and game theory**	*Grnarova*
04:30 PM	**Poster spotlight #2**	*Fusi, Arachie, Monteiro, Wolf*
05:00 PM	**Discussion panel**	
05:30 PM	**Concluding remarks**	
05:40 PM	**Poster session afternoon**	

Abstracts (11):

Abstract 2: **Improving Generative Adversarial Networks using Game Theory and Statistics in Smooth Games Optimization and Machine Learning**, *Daskalakis* 08:50 AM

Generative Adversarial Networks (aka GANs) are a recently proposed approach for learning samplers of high-dimensional distributions with intricate structure, such as distributions over natural images, given samples from these distributions. They are trained by setting up a two-player zero-sum game between two neural networks, which learn statistics of a target distribution by adapting their strategies in the game using gradient descent. Despite their intriguing performance in practice, GANs pose great challenges to both Optimization and Statistics. Their training suffers from oscillations, and they are difficult to scale to high-dimensional settings. We study how game-theoretic and statistical techniques can be brought to bare on these important challenges. We use Game Theory towards improving GAN training, and Statistics towards scaling up the dimensionality of the generated distributions.

Abstract 3: **Poster spotlight in Smooth Games Optimization and Machine Learning**, *Yang, Dvurechenskii, Mertikopoulos, Berard* 09:30 AM

1) Provable Non-Convex Min-Max Optimization; Mingrui Liu, Rafique Hassan, Qihang Lin and Tianbao Yang.

2) Solving differential games by methods for finite-dimensional saddle-point problems; Pavel Dvurechensky, Yurii Nesterov and Vladimir Spokoiny.

3) Generalized Mirror Prox Algorithm for Variational Inequalities; Pavel Dvurechensky, Alexander Gasnikov, Fedor Stonyakin and Alexander Titov.

4) On the convergence of stochastic forward-backward-forward algorithms with variance reduction in pseudo-monotone variational inequalities; Mathias Staudigl, Radu Ioan Bot, Phan Tu Vuong and Panayotis Mertikopoulos.

5) A Variational Inequality Perspective on Generative Adversarial Networks; Gauthier Gidel, Hugo Berard, Gaëtan Vignoud, Pascal Vincent and Simon Lacoste-Julien.

Abstract 6: **Smooth Games in Machine Learning Beyond GANs in Smooth Games Optimization and Machine Learning**, *He* 11:00 AM

This talk will discuss a wide spectrum of recent advances in machine learning using smooth games, beyond the phenomenal GANs. Such showcases include reinforcement learning, robust and adversarial machine learning, approximate Bayesian computation, maximum likelihood estimation in exponential family, and etc. We show that all of these classical machine learning tasks can be reduced to solving (non) convex-concave min-max optimization problems. Hence, it is of paramount importance to developing a good theoretical understanding and principled algorithms for min-max optimization. We will review some of the theory and algorithms for smooth games and variational inequalities in the convex regime and shed some light on their counterparts in the non-convex regime.

Abstract 7: **Finding Mixed Nash Equilibria of Generative Adversarial Networks in Smooth Games Optimization and Machine Learning**, *Cevher* 11:40 AM

We reconsider the training objective of Generative Adversarial Networks (GANs) from the mixed Nash Equilibria (NE) perspective. Inspired by the classical prox methods, we develop a novel algorithmic framework for GANs via an infinite-dimensional two-player game and prove rigorous convergence rates to the mixed NE. We then propose a principled procedure to reduce our novel prox methods to simple sampling routines, leading to practically efficient algorithms. Finally, we provide experimental evidence that our approach outperforms methods that seek pure strategy equilibria, such as SGD, Adam, and RMSProp, both in speed and quality.

Abstract 8: **Bounding Inefficiency of Equilibria in Continuous Actions Games using Submodularity and Curvature in Smooth Games Optimization and Machine Learning**, *Sessa* 12:00 PM

Games with continuous strategy sets arise in several machine learning problems (e.g. adversarial learning). For such games, simple no-regret learning algorithms exist in several cases and ensure convergence to coarse correlated equilibria (CCE). The efficiency of such equilibria with respect to a social function, however, is not well understood. In this paper, we define the class of valid utility games with continuous strategies and provide efficiency bounds for their CCEs. Our bounds rely on the social function satisfying recently introduced notions of submodularity over continuous domains. We further refine our bounds based on the curvature of the social function. Furthermore, we extend our efficiency bounds to a class of non-submodular functions that satisfy approximate submodularity properties. Finally, we show that valid utility games with continuous strategies can be designed to maximize monotone DR-submodular functions subject to disjoint constraints with approximation guarantees. The approximation guarantees we derive are based on the efficiency of the equilibria of such games and can improve the existing ones in the literature. We illustrate and validate our results on a budget allocation game and a sensor coverage problem.

Abstract 10: **Building Algorithms by Playing Games in Smooth Games Optimization and Machine Learning**, *Abernethy* 02:00 PM

A very popular trick for solving certain types of optimization problems is this: write your objective as the solution of a two-player zero-sum game, endow both players with an appropriate learning algorithm, watch how the opponents compete, and extract an (approximate) solution from the actions/decisions taken by the players throughout the process. This approach is very generic and provides a natural template to produce new and interesting algorithms. I will describe this framework and show how it applies in several scenarios, and describe recent work that draws a connection to the Frank-Wolfe algorithm and Nesterov's Accelerated Gradient Descent.

Abstract 11: **Negative Momentum for Improved Game Dynamics in Smooth Games Optimization and Machine Learning**, *Askari Hemmat* 02:40 PM

Games generalize the single-objective optimization paradigm by introducing different objective functions for different players. Differentiable games often proceed by simultaneous or alternating gradient updates. In machine learning, games are gaining new importance through formulations like generative adversarial networks (GANs) and actor-critic systems. However, compared to single-objective optimization, game dynamics are more complex and less understood. In this paper, we analyze gradient-based methods with momentum on simple games. Next, we show empirically that alternating gradient updates with a negative momentum term achieves convergence on the notoriously difficult to train saturating GANs.

Abstract 13: **Regret Decomposition in Sequential Games with Convex Action Spaces and Losses in Smooth Games Optimization and Machine Learning**, *Farina* 03:30 PM

We derive a new framework for regret minimization on sequential decision problems and extensive-form games with general compact convex sets at each decision point and general convex losses, as opposed to prior work which has been for simplex decision points and linear losses. We call our framework laminar regret decomposition. It generalizes the CFR algorithm to this more general setting. Furthermore, our framework enables a new proof of CFR even in the known setting, which is derived from a perspective of decomposing polytope regret, thereby leading to an arguably simpler interpretation of the algorithm. Our generalization to convex compact sets and convex losses allows us to develop new algorithms for several problems: regularized sequential decision making, regularized Nash equilibria in extensive-form games, and computing approximate extensive-form perfect equilibria. Our generalization also leads to the first regret-minimization algorithm for computing reduced-normal-form quantal response equilibria based on minimizing local regrets.

Abstract 14: **An interpretation of GANs via online learning and game theory in Smooth Games Optimization and Machine Learning**, *Grnarova* 03:50 PM

Generative Adversarial Networks (GANs) have become one of the most powerful paradigms in learning real-world distributions. Despite this success, their minimax nature makes them fundamentally different to more classical generative models thus raising novel challenges; most notably in terms of training and evaluation. Indeed, finding a saddle-point is in general a harder task than converging to an extremum. We view the problem of training GANs as finding a mixed strategy in a zero-sum game. Building upon ideas from online learning and game theory, we propose (i) a novel training method with provable convergence to an equilibrium for semi-shallow GAN architectures, i.e. architectures where the discriminator is a one layer network and the generator is an arbitrary network and (ii) a natural metric for detecting non-convergence, namely the duality gap.

Abstract 15: **Poster spotlight #2 in Smooth Games Optimization and Machine Learning**, *Fusi, Arachie, Monteiro, Wolf* 04:30 PM

1) Model Compression with Generative Adversarial Networks; Ruishan Liu, Nicolo Fusi and Lester Mackey.

2) An Adversarial Labeling Game for Learning from Weak Supervision; Chidubem Arachie and Bert Huang.

3) Multi-objective training of Generative Adversarial Networks with multiple discriminators; Isabela Albuquerque, Joao Monteiro, Thang Doan, Breandan Considine, Tiago Falk and Ioannis Mitliagkas.

4) A GAN framework for Instance Segmentation using the Mutex Watershed Algorithm; Mandikal Vikram and Steffen Wolf.

Abstract 16: **Discussion panel in Smooth Games Optimization and Machine Learning**, 05:00 PM

Panel with invited speakers.

Visually grounded interaction and language

Florian Strub, Harm de Vries, Erik Wijmans, Samyak Datta, Ethan Perez, Mateusz Malinowski, Stefan Lee, Peter Anderson, Aaron Courville, Jeremie MARY, Dhruv Batra, Devi Parikh, Olivier Pietquin, Chiori HORI, Tim Marks, Anoop Cherian

Room 512 CDGH, Fri Dec 07, 08:00 AM

The dominant paradigm in modern natural language understanding is learning statistical language models from text-only corpora. This approach is founded on a distributional notion of semantics, i.e. that the "meaning" of a word is based only on its relationship to other words. While effective for many applications, methods in this family suffer from limited semantic understanding, as they miss learning from the multimodal and interactive environment in which communication often takes place - the symbols of language thus are not grounded in anything concrete. The symbol grounding problem first highlighted this limitation, that “meaningless symbols (i.e.) words cannot be grounded in anything but other meaningless symbols” [18].

On the other hand, humans acquire language by communicating about and interacting within a rich, perceptual environment. This behavior provides the necessary grounding for symbols, i.e. to concrete objects or concepts (i.e. physical or psychological). Thus, recent work has aimed to bridge vision, interactive learning, and natural language understanding through language learning tasks based on natural images (ReferIt [1], GuessWhat?! [2], Visual Question Answering [3,4,5,6], Visual Dialog [7], Captioning [8]) or through embodied agents performing interactive tasks [13,14,17,22,23,24,26] in physically simulated environments (DeepMind Lab [9], Baidu XWorld [10], OpenAI Universe [11], House3D [20], Matterport3D [21], GIBSON [24], MINOS [25], AI2-THOR [19], StreetLearn [17]), often drawing on the recent successes of deep learning and reinforcement learning. We believe this line of research poses a promising, long-term solution to the grounding problem faced by current, popular language understanding models.

While machine learning research exploring visually-grounded language learning may be in its earlier stages, it may be possible to draw insights from the rich research literature on human language acquisition. In neuroscience, recent progress in fMRI technology has enabled to better understand the interleave between language, vision and other modalities [15,16] suggesting that the brains shares neural representation of concepts across vision and language. Differently, developmental cognitive scientists have also argued that children acquiring various words is closely linked to them learning the underlying concept in the real world [12].

This workshop thus aims to gather people from various backgrounds - machine learning, computer vision, natural language processing, neuroscience, cognitive science, psychology, and philosophy - to share and debate their perspectives on why grounding may (or may not) be important in building machines that truly understand natural language.

We invite you to submit papers related to the following topics:
- language acquisition or learning through interactions
- visual captioning, dialog, and question-answering
- reasoning in language and vision
- visual synthesis from language
- transfer learning in language and vision tasks
- navigation in virtual worlds via natural-language instructions or multi-agent communication
- machine translation with visual cues
- novel tasks that combine language, vision and actions
- modeling of natural language and visual stimuli representations in the human brain
- position papers on grounded language learning
- audio visual scene-aware dialog
- audio-visual fusion

Submissions should be up to 4 pages excluding references, acknowledgements, and supplementary material, and should be NIPS format and anonymous. The review process is double-blind.

We also welcome published papers that are within the scope of the workshop (without re-formatting). This specific papers do not have to be anonymous. They are not eligible for oral session and will only have a very light review process.

Please submit your paper to the following address: https://cmt3.research.microsoft.com/VIGIL2018

Accepted workshop papers are eligible to the pool of reserved

conference tickets (one ticket per accepted papers).

If you have any question, send an email to:
vigilworkshop2018@gmail.com

[1] Sahar Kazemzadeh et al. "ReferItGame: Referring to Objects in Photographs of Natural Scenes." EMNLP, 2014.
[2] Harm de Vries et al. "GuessWhat?! Visual object discovery through multi-modal dialogue." CVPR, 2017.
[3] Stanislaw Antol et al. "Vqa: Visual question answering." ICCV, 2015.
[4] Mateusz Malinowski et al. “Ask Your Neurons: A Neural-based Approach to Answering Questions about Images.” ICCV, 2015.
[5] Mateusz Malinowski et al. “A Multi-World Approach to Question Answering about Real-World Scenes based on Uncertain Input.” NIPS, 2014.
[6] Geman Donald, et al. “Visual Turing test for computer vision systems.” PNAS, 2015.
[7] Abhishek Das et al. "Visual dialog." CVPR, 2017.
[8] Anna Rohrbach et al. “Generating Descriptions with Grounded and Co-Referenced People.” CVPR, 2017.
[9] Charles Beattie et al. Deepmind lab. arXiv, 2016.
[10] Haonan Yu et al. “Guided Feature Transformation (GFT): A Neural Language Grounding Module for Embodied Agents.” arXiv, 2018.
[11] Openai universe. https://universe.openai.com, 2016.
[12] Alison Gopnik et al. “Semantic and cognitive development in 15- to 21-month-old children.” Journal of Child Language, 1984.
[13] Abhishek Das et al. "Learning Cooperative Visual Dialog Agents with Deep Reinforcement Learning." ICCV, 2017.
[14] Karl Moritz Hermann et al. "Grounded Language Learning in a Simulated 3D World." arXiv, 2017.
[15] Alexander G. Huth et al. "Natural speech reveals the semantic maps that tile human cerebral cortex." Nature, 2016.
[16] Alexander G. Huth, et al. "Decoding the semantic content of natural movies from human brain activity." Frontiers in systems neuroscience, 2016.
[17] Piotr Mirowski et al. “Learning to Navigate in Cities Without a Map.” arXiv, 2018.
[18] Stevan Harnad. “The symbol grounding problem.” CNLS, 1989.
[19] E Kolve, R Mottaghi, D Gordon, Y Zhu, A Gupta, A Farhadi. “AI2-THOR: An Interactive 3D Environment for Visual AI.” arXiv, 2017.
[20] Yi Wu et al. “House3D: A Rich and Realistic 3D Environment.” arXiv, 2017.
[21] Angel Chang et al. “Matterport3D: Learning from RGB-D Data in Indoor Environments.” arXiv, 2017.
[22] Abhishek Das et al. “Embodied Question Answering.” CVPR, 2018.
[23] Peter Anderson et al. “Vision-and-Language Navigation: Interpreting visually-grounded navigation instructions in real environments.” CVPR, 2018.
[24] Fei Xia et al. “Gibson Env: Real-World Perception for Embodied Agents.” CVPR, 2018.
[25] Manolis Savva et al. “MINOS: Multimodal indoor simulator for navigation in complex environments.” arXiv, 2017.
[26] Daniel Gordon, Aniruddha Kembhavi, Mohammad Rastegari, Joseph Redmon, Dieter Fox, Ali Farhadi. “IQA: Visual Question Answering in Interactive Environments.” CVPR, 2018.

Schedule

08:30 AM	**Opening Remarks**	*Strub*
08:40 AM	**Steven Harnad - The symbol grounding problem**	*Harnad*
09:20 AM	**Antonio Torralba - Learning to See and Hear**	*Torralba*
10:00 AM	**Audio Visual Semantic Understanding Challenge**	*HORI, Marks*
10:15 AM	**Spotlights**	
10:30 AM	**Coffee Break**	
10:50 AM	**Douwe Kiela - Learning Multimodal Embeddings**	*Kiela*
11:30 AM	**Roozbehm Mottaghi - Interactive Scene Understanding**	*Mottaghi*
12:10 PM	**Poster Sessions and Lunch (Provided)**	*Utsumi, Suhr, Zhang, Sanabria, Kafle, Chen, Kim, Agrawal, DUMPALA, Murty, Azagra, ROUAT, Ali, , OOTA, Lin, Palaskar, Lai, Aly, Shen, Li, Zhang, Kuznetsova, An, Delbrouck, Kornuta, Javed, Davis, Co-Reyes, Sharma, Lyu, Xie, Kalra, ling, Maksymets, Jain, Chuang, Agarwal, Abdelnour, Feng, albouy, Karamcheti, Doran, Raileanu, Heek*
01:40 PM	**Angeliki Lazaridou - Emergence of (linguistic communication) through multi-agent interactions**	*Lazaridou*
02:20 PM	**Barbara Landau - Learning simple spatial terms: Core and more**	*Landau*
03:00 PM	**Coffee Break and Poster Session**	
03:50 PM	**Joyce Chai - Language Communication with Robots**	*Chai*
04:30 PM	**Christopher Manning - Towards real-world visual reasoning**	*Manning*
05:10 PM	**Panel Discussion**	*Torralba, Kiela, Landau, Lazaridou, Chai, Manning, Harnad, Mottaghi*
06:00 PM	**Closing Remarks**	*Strub*

Abstracts (1):

Abstract 10: **Angeliki Lazaridou - Emergence of (linguistic communication) through multi-agent interactions in Visually grounded interaction and language**, *Lazaridou* 01:40 PM

Distributional models and other supervised models of language focus on the structure of language and are an excellent way to learn general statistical associations between sequences of symbols. However, they do not capture the functional aspects of communication, i.e., that humans have intentions and use words to coordinate with others and make things happen in the real world. In this talk, I will present two studies on multi-agent emergent communication, where agents exist in some grounded environment and have to communicate about objects and their properties. This process requires the negotiation of linguistic meaning in this pragmatic context of achieving their goal. In the first study, I will present experiments in which agents learn to form a common ground that allow them to communicate about disentangled (i.e., feature norm) and entangled (i.e., raw pixels) input. In the second study, I will talk about properties of linguistic communication as arising in the context of self-interested agents.

Modeling and decision-making in the spatiotemporal domain

Ransalu Senanayake, Neal Jean, Fabio Ramos, Girish Chowdhary

Room 513 ABC, Fri Dec 07, 08:00 AM

Friday, December 07, 2018 at Room 513ABC

Abstract: Understanding the evolution of a process over space and time is fundamental to a variety of disciplines. To name a few, such phenomena that exhibit dynamics in both space and time include propagation of diseases, variations in air pollution, dynamics in fluid flows, and patterns in neural activity. In addition to these fields in which modeling the nonlinear evolution of a process is the focus, there is also an emerging interest in decision-making and controlling of autonomous agents in the spatiotemporal domain. That is, in addition to learning what actions to take, when and where to take actions is crucial for an agent to efficiently and safely operate in dynamic environments. Although various modeling techniques and conventions are used in different application domains, the fundamental principles remain unchanged. Automatically capturing the dependencies between spatial and temporal components, making accurate predictions into the future, quantifying the uncertainty associated with predictions, real-time performance, and working in both big data and data scarce regimes are some of the key aspects that deserve our attention. Establishing connections between Machine Learning and Statistics, this workshop aims at;
(1) raising open questions on challenges of spatiotemporal modeling and decision-making,
(2) establishing connections among diverse application domains of spatiotemporal modeling, and
(3) encouraging conversation between theoreticians and practitioners to develop robust predictive models.

Keywords
Theory: deep learning/convolutional LSTM, kernel methods, chaos theory, reinforcement learning for dynamic environments, dynamic policy learning, biostatistics,
epidemiology, geostatistcs, climatology, neuroscience, etc.
Applications:
Natural phenomena: disease propagation and outbreaks, environmental monitoring, climate modeling, etc.
Social and economics: predictive policing, population mapping, poverty mapping, food resources, agriculture, etc.
Engineering/robotics: active data collection, traffic modeling, motion prediction, fluid dynamics, spatiotemporal prediction for safe autonomous driving, etc.

Web: https://sites.google.com/site/nips18spatiotemporal/

Schedule

08:30 AM	**Christopher Wikle (Uni. of Missouri): Introduction to spatiotemporal modeling**	*Wikle*
09:45 AM	**Spotlight talks (session 1)**	*Roberts, Kozak, Owoeye, dahl, Dirie, Chang, Ivashkin*
10:00 AM	**Modeling Rape Reporting Delays Using Spatial, Temporal and Social Features**	*Klemmer*
10:15 AM	**Spotlight talks (session 2)**	*Giffard-Roisin, Rußwurm, Suel, Tang, Maske, Neill, Lee*
10:30 AM	**Coffee break + poster session 1**	
11:00 AM	**Stefano Ermon (Stanford University): Weakly Supervised Spatio-temporal Regression**	*Ermon*
11:30 AM	**Long Range Sequence Generation via Multiresolution Adversarial Training**	*Yu*
11:45 AM	**Modeling Spatiotemporal Multimodal Language with Recurrent Multistage Fusion**	*Liang*
12:00 PM	**Spotlight talks (session 3)**	*Mahdisoltani, Kratzert, OOTA, Motani, Gangopadhyay, Madhusudhan, Rußwurm, Mousavi, Jain*
12:20 PM	**Lunch break**	
01:45 PM	**A Nonparametric Spatio-temporal SDE Model**	*Lahdesmaki*
02:00 PM	**Spotlight talks (session 4)**	*Walid, Li, Pajouheshgar, Hennigh, Kim, Kulkarni, Takeuchi*
02:15 PM	**Ani Hsieh (UPenn): Modeling, Tracking, and Learning Coherent Spatiotemporal Features in Geophysical Flows**	*Hsieh*
02:45 PM	**Spotlight talks (session 5)**	*Asseman, Marchant, TRIVEDI, Narayanaswamy, AMROUCHE, Martin, FERNANDEZ PINTO*
03:00 PM	**Coffee break + poster session 2**	

03:30 PM	**Chelsea Finn (UCBerkeley / Google Brain): Learning Generalizable Behavior through Unsupervised Interaction**	*Finn*
04:00 PM	**Girish Chowdhary (UIUC): Spatiotemporal Learning for Enabling Agricultural Robotics**	*Chowdhary*
04:30 PM	**Quantile Regression Reinforcement Learning with State Aligned Vector Rewards**	*Richter*
04:45 PM	**Path Planning for Mobile Inference of Spatiotemporally Evolving Systems**	*Whitman*
05:00 PM	**Fabio Ramos (Uni. of Sydney): Learning and Planning in Spatial-Temporal Data**	*Ramos*
05:30 PM	**Tomaso Poggio (MIT): Dynamical System Theory for Deep Learning**	*Poggio*
06:00 PM	**Panel Discussion**	

Abstracts (6):

Abstract 1: **Christopher Wikle (Uni. of Missouri): Introduction to spatiotemporal modeling in Modeling and decision-making in the spatiotemporal domain**, *Wikle* 08:30 AM

Christopher K. Wikle is Curators' Distinguished Professor of Statistics at the University of Missouri (MU), with additional appointments in Soil, Environmental and Atmospheric Sciences and the Truman School of Public Affairs. He received a PhD co-major in Statistics and Atmospheric Science in 1996 from Iowa State University. He was research fellow at the National Center for Atmospheric Research from 1996-1998, after which he joined the MU Department of Statistics. His research interests are in spatio-temporal statistics applied to environmental, geophysical, agricultural and federal survey applications, with particular interest in dynamics. Awards include elected Fellow of the American Statistical Association (ASA), Distinguished Alumni Award from the College of Liberal Arts and Sciences at Iowa State University, ASA ENVR Section Distinguished Achievement Award, co-awardee 2017 ASA Statistical Partnership Among Academe, Industry, and Government (SPAIG) Award, the MU Chancellor's Award for Outstanding Research and Creative Activity in the Physical and Mathematical Sciences, the Outstanding Graduate Faculty Award, and Outstanding Undergraduate Research Mentor Award. His book Statistics for Spatio-Temporal Data (co-authored with Noel Cressie) was the 2011 PROSE Award winner for excellence in the Mathematics Category by the Association of American Publishers and the 2013 DeGroot Prize winner from the International Society for Bayesian Analysis. He is Associate Editor for several journals and is one of six inaugural members of the Statistics Board of Reviewing Editors for Science.

Abstract 6: **Stefano Ermon (Stanford University): Weakly Supervised Spatio-temporal Regression in Modeling and decision-making in the spatiotemporal domain**, *Ermon* 11:00 AM

Stefano Ermon is an Assistant Professor in the Department of Computer Science at Stanford University, where he is affiliated with the Artificial Intelligence Laboratory and a fellow of the Woods Institute for the Environment. His research is centered on techniques for scalable and accurate inference in graphical models, statistical modeling of data, large-scale combinatorial optimization, and robust decision making under uncertainty, and is motivated by a range of applications, in particular ones in the emerging field of computational sustainability with applications in poverty mapping and remote sensing.

Abstract 13: **Ani Hsieh (UPenn): Modeling, Tracking, and Learning Coherent Spatiotemporal Features in Geophysical Flows in Modeling and decision-making in the spatiotemporal domain**, *Hsieh* 02:15 PM

M. Ani Hsieh is a Research Associate Professor in the Department of Mechanical Engineering & Applied Mechanics at the University of Pennsylvania. She received a B.S. in Engineering and B.A. in Economics from Swarthmore College in 1999 and a PhD in Mechanical Engineering from the University of Pennsylvania in 2007. Dr. Hsieh has been a Visiting Assistant Professor in the Engineering Department at Swarthmore College (2007-2008) and an Associate Professor in the Mechanical Engineering & Mechanics Department at Drexel University (2008-2017). Her research interests include many robot systems and marine robotics, geophysical fluid dynamics, and dynamical systems. She is a recipient of a 2012 Office of Naval Research (ONR) Young Investigator Award and a 2013 National Science Foundation (NSF) CAREER Award.

Abstract 16: **Chelsea Finn (UCBerkeley / Google Brain): Learning Generalizable Behavior through Unsupervised Interaction in Modeling and decision-making in the spatiotemporal domain**, *Finn* 03:30 PM

Chelsea Finn is a postdoctoral fellow in Computer Science at UC Berkeley, where she works on machine learning and its intersection with robotic perception and control. She is a part of Berkeley AI Research Lab (BAIR). She recently spent time at Google Brain. Before graduate school, she received a Bachelors in EECS at MIT, where she worked on several research projects, including an assistive technology project in CSAIL and an animal biometrics project. She has also spent time at Counsyl, Google, and Sandia National Labs. She will discuss on spatiotemporal aspects of video prediction and deep spatial autoencoders for visuomotor learning.

Abstract 17: **Girish Chowdhary (UIUC): Spatiotemporal Learning for Enabling Agricultural Robotics in Modeling and decision-making in the spatiotemporal domain**, *Chowdhary* 04:00 PM

Girish Chowdhary is the director of DAS laboratory and Assistant Professor. Girish has a Ph.D. degree from Georgia Institute of Technology. He then spent around two years at Massachusetts Institute of Technology's Laboratory for Information and Decision Systems and the School of Aeronautics and Astronautics as a postdoctoral associate. Prior to coming to Georgia Tech, he spent three years working as a research engineer with the German Aerospace Center's (DLR's) Institute for Flight Systems Technology in Braunschweig, Germany. He holds a BE with honors from RMIT University in Melbourne, Australia. Girish is

the author of several peer-reviewed publications spanning the area of adaptive control, spatiotemporal modeling, autonomy and decision making, LIDAR-based perception for Unmanned Aerial Systems (UAS), and GPS denied navigation.

Abstract 20: **Fabio Ramos (Uni. of Sydney): Learning and Planning in Spatial-Temporal Data in Modeling and decision-making in the spatiotemporal domain**, *Ramos* 05:00 PM

Abstract:
Modern sensors provide immense amounts of information that need to be efficiently integrated into probabilistic models representing the environment autonomous systems operate in. In this talk I will show statistical machine learning methods for spatial and spatial-temporal data that are able to fuse information from heterogeneous sources, scaling gracefully to very large datasets. I will demonstrate how Bayesian reasoning and the principle of modelling uncertainty can be used to mitigate risks in decision making, for motion planning with indoor robots, to continental-scale natural resource exploration.

Bio:
Fabio Ramos is an Associate Professor in machine learning and robotics at the School of Information Technologies, University of Sydney, and co-Director of the Centre for Translational Data Science. He received the B.Sc. and the M.Sc. degrees in Mechatronics Engineering at University of Sao Paulo, Brazil, in 2001 and 2003 respectively, and the Ph.D. degree at University of Sydney, Australia, in 2008. He has over 130 peer-reviewed publications and received best paper awards at ECML'18, IROS'05, ACRA'07, and Best Paper Finalist at RSS'17. His research focuses on statistical machine learning techniques for data fusion, with applications in robotics, large-scale autonomous systems, environmental monitoring and healthcare.

Workshop on Security in Machine Learning

Nicolas Papernot, Florian Tramer, Kamalika Chaudhuri, Matt Fredrikson, Jacob Steinhardt

Room 513DEF, Fri Dec 07, 08:00 AM

There is growing recognition that ML exposes new vulnerabilities in software systems. Some of the threat vectors explored so far include training data poisoning, adversarial examples or model extraction. Yet, the technical community's understanding of the nature and extent of the resulting vulnerabilities remains limited. This is due in part to (1) the large attack surface exposed by ML algorithms because they were designed for deployment in benign environments---as exemplified by the IID assumption for training and test data, (2) the limited availability of theoretical tools to analyze generalization, (3) the lack of reliable confidence estimates. In addition, the majority of work so far has focused on a small set of application domains and threat models.

This workshop will bring together experts from the computer security and machine learning communities in an attempt to highlight recent work that contribute to address these challenges. Our agenda will complement contributed papers with invited speakers. The latter will emphasize connections between ML security and other research areas such as accountability or formal verification, as well as stress social aspects of ML misuses. We hope this will help identify fundamental directions for future cross-community collaborations, thus charting a path towards secure and trustworthy ML.

Schedule

09:00 AM	**Sever: A Robust Meta-Algorithm for Stochastic Optimization by Jerry Li**	
09:15 AM	**Semidefinite relaxations for certifying robustness to adversarial examples by Aditi Raghunathan**	
09:45 AM	**On the Effectiveness of Interval Bound Propagation for Training Verifiably Robust Models**	
11:00 AM	**A Sociotechnical Approach to Security in Machine Learning by danah boyd**	*boyd*
11:45 AM	**Law and Adversarial Machine Learning**	
01:30 PM	**Interpretability for when NOT to use machine learning by Been Kim**	*Kim*
02:00 PM	**Rigorous Agent Evaluation: An Adversarial Approach to Uncover Catastrophic Failures**	
02:15 PM	**Semantic Adversarial Examples by Somesh Jha**	*Jha*
04:15 PM	**Safety verification for neural networks with provable guarantees by Marta Kwiatkowska**	*Kwiatkowska*
04:45 PM	**Model Poisoning Attacks in Federated Learning**	

2nd Workshop on Machine Learning on the Phone and other Consumer Devices (MLPCD 2)

Sujith Ravi, Wei Chai, Yangqing Jia, Hrishikesh Aradhye, Prateek Jain

Room 514, Fri Dec 07, 08:00 AM

The 2nd Workshop on Machine Learning on the Phone and other Consumer Devices (MLPCD 2) aims to continue the success of the 1st MLPCD workshop held at NIPS 2017 in Long Beach, CA.

Previously, the first MLPCD workshop edition, held at NIPS 2017 was successful, attracted over 200+ attendees and led to active research & panel discussions as well as follow-up contributions to the open-source community (e.g., release of new inference libraries, tools, models and standardized representations of deep learning models). We believe that

interest in this space is only going to increase, and we hope that the workshop plays the role of an influential catalyst to foster research and collaboration in this nascent community.

After the first workshop where we investigated initial directions and trends, the NIPS 2018 MLPCD workshop focuses on theory and practical applications of on-device machine learning, an area that is highly relevant and specializes in the intersection of multiple topics of interest to NIPS and broader machine learning community -- efficient training & inference for deep learning and other machine learning models; interdisciplinary mobile applications involving vision, language & speech understanding; and emerging topics like Internet of Things.

We plan to incorporate several new additions this year -- inspirational opening Keynote talk on "future of intelligent assistive & wearable experiences"; two panels including a lively closing panel debate discussing pros/cons of two key ML computing paradigms (Cloud vs. On-device); solicited research papers on new & recent hot topics (e.g., theoretical & algorithmic work on low-precision models, compression, sparsity, etc. for training and inference), related challenges, applications and recent trends; demo session showcasing ML in action for real-world apps.

Description & Topics:

Deep learning and machine learning, in general, has changed the computing paradigm. Products of today are built with machine intelligence as a central attribute, and consumers are beginning to expect near-human interaction with the appliances they use. However, much of the Deep Learning revolution has been limited to the cloud, enabled by popular toolkits such as Caffe, TensorFlow, and MxNet, and by specialized hardware such as TPUs. In comparison, mobile devices until recently were just not fast enough, there were limited developer tools, and there were limited use cases that required on-device machine learning. That has recently started to change, with the advances in real-time computer vision and spoken language understanding driving real innovation in intelligent mobile applications. Several mobile-optimized neural network libraries were recently announced (CoreML, Caffe2 for mobile, TensorFlow Lite), which aim to dramatically reduce the barrier to entry for mobile machine learning. Innovation and competition at the silicon layer has enabled new possibilities for hardware acceleration. To make things even better, mobile-optimized versions of several state-of-the-art benchmark models were recently open sourced. Widespread increase in availability of connected “smart” appliances for consumers and IoT platforms for industrial use cases means that there is an ever-expanding surface area for mobile intelligence and ambient devices in homes. All of these advances in combination imply that we are likely at the cusp of a rapid increase in research interest in on-device machine learning, and in particular, on-device neural computing.

Significant research challenges remain, however. Mobile devices are even more personal than “personal computers” were. Enabling machine learning while simultaneously preserving user trust requires ongoing advances in the research of differential privacy and federated learning techniques. On-device ML has to keep model size and power usage low while simultaneously optimizing for accuracy. There are a few exciting novel approaches recently developed in mobile optimization of neural networks. Lastly, the newly prevalent use of camera and voice as interaction models has fueled exciting research towards neural techniques for image and speech/language understanding. This is an area that is highly relevant to multiple topics of interest to NIPS -- e.g., core topics like machine learning & efficient inference and interdisciplinary applications involving vision, language & speech understanding as well as emerging area (namely, Internet of Things).

With this emerging interest as well as the wealth of challenging research problems in mind, we are proposing the second NIPS 2018 workshop dedicated to on-device machine learning for mobile and ambient home consumer devices.

Areas/topics of interest include, but not limited to:

* Model compression for efficient inference with deep networks and other ML models
* Privacy preserving machine learning
* Low-precision training/inference & Hardware acceleration of neural computing on mobile devices
* Real-time mobile computer vision
* Language understanding and conversational assistants on mobile devices
* Speech recognition on mobile and smart home devices
* Machine intelligence for mobile gaming
* ML for mobile health other real-time prediction scenarios
* ML for on-device applications in the automotive industry (e.g., computer vision for self-driving cars)
* Software libraries (including open-source) optimized for on-device ML

Target Audience:

The next wave of ML applications will have significant processing on mobile and ambient devices. Some immediate examples of these are single-image classification, depth estimation, object recognition and segmentation running on-device for creative effects, or on-device recommender and ranking systems for privacy-preserving, low-latency experiences. This workshop will bring ML practitioners up to speed on the latest trends for on-device applications of ML, offer an overview of the latest HW and SW framework developments, and champion active research towards hard technical challenges emerging in this nascent area. The target audience for the workshop is both industrial and academic researchers and practitioners of on-device, native machine learning. The workshop will cover both “informational” and “aspirational” aspects of this emerging research area for delivering ground-breaking experiences on real-world products.

Given the relevance of the topic, target audience (mix of industry + academia & related parties) as well as the timing (confluence of research ideas + practical implementations both in industry as well as through publicly available toolkits), we feel that NIPS 2018 would continue to be a great venue for this workshop.

Schedule

08:15 AM	**Opening (Chairs)**	
08:30 AM	**Aurélien Bellet**	*Bellet*
08:45 AM	**Neel Guha**	*Guha*
09:00 AM	**Prof. Kurt Keutzer**	*Keutzer*

09:30 AM	**Ting-Wu Chin**	*Chin*
09:45 AM	**Prof. Thad Starner**	*Starner*
10:30 AM	**Coffee break (morning)**	
11:00 AM	**Prof. Max Welling**	*Welling*
11:30 AM	**Zornitsa Kozareva**	*Kozareva*
11:50 AM	**Spotlight (poster, demo), Lunch & Poster Session**	*Singh, Dow, Dürichen, Whatmough, Feng, Patra, Patil, Jeong, Lin, Izumi, Leang, Xu, zhang, Witteveen*
01:30 PM	**Brendan McMahan**	*McMahan*
02:00 PM	**Prof. Virginia Smith**	*Smith*
02:30 PM	**Meghan Cowan**	*Cowan*
02:45 PM	**Kuan Wang**	*Wang*
03:00 PM	**Coffee break (afternoon)**	
03:30 PM	**Jan Kautz**	*Kautz*
04:00 PM	**Prof. Song Han**	*Han*
04:30 PM	**Demo session**	*Damani, Dow, Izumi, Patil, Leang, Xu, zhang*

Machine Learning for Geophysical & Geochemical Signals

Laura Pyrak-Nolte, Jim Rustad, Richard Baraniuk

Room 515, Fri Dec 07, 08:00 AM

Motivation
The interpretation of Earth's subsurface evolution from full waveform analysis requires a method to identify the key signal components related to the evolution in physical properties from changes in stress, fluids, geochemical interactions and other natural and anthropogenic processes. The analysis of seismic waves and other geophysical/geochemical signals remains for the most part a tedious task that geoscientists may perform by visual inspection of the available seismograms. The complexity and noisy nature of a broad array of geoscience signals combined with sparse and irregular sampling make this analysis difficult and imprecise. In addition, many signal components are ignored in tomographic imaging and continuous signal analysis that may prevent discovery of previously unrevealed signals that may point to new physics.

Ideally a detailed interpretation of the geometric contents of these data sets would provide valuable prior information for the solution of corresponding inverse problems. This unsatisfactory state of affairs is indicative of a lack of effective and robust algorithms for the computational parsing and interpretation of seismograms (and other geoscience data sets). Indeed, the limited frequency content, strong nonlinearity, temporally scattered nature of these signals make their analysis with standard signal processing techniques difficult and insufficient.

Once important seismic phases are identified, the next challenge is determining the link between a remotely-measured geophysical response and a characteristic property (or properties) of the fractures and fracture system. While a strong laboratory-based foundation has established a link between the mechanical properties of simple fracture systems (i.e. single fractures, parallel sets of fractures) and elastic wave scattering, bridging to the field scale faces additional complexity and a range of length scales that cannot be achieved from laboratory insight alone. This fundamental knowledge gap at the critical scale for long-term monitoring and risk assessment can only be narrowed or closed with the development of appropriate mathematical and numerical representations at each scale and across scales using multiphysics models that traverse spatial and temporal scales.

Topic
Major breakthroughs in bridging the knowledge gaps in geophysical sensing are anticipated as more researchers turn to machine learning (ML) techniques; however, owing to the inherent complexity of machine learning methods, they are prone to misapplication, may produce uninterpretable models, and are often insufficiently documented. This combination of attributes hinders both reliable assessment of model validity and consistent interpretation of model outputs. By providing documented datasets and challenging teams to apply fully documented workflows for ML approaches, we expect to accelerate progress in the application of data science to longstanding research issues in geophysics.

The goals of this workshop are to:
(1) bring together experts from different fields of ML and geophysics to explore the use of ML techniques related to the identification of the physics contained in geophysical and chemical signals, as well as from images of geologic materials (minerals, fracture patterns, etc.); and
(2) announce a set of geophysics machine learning challenges to the community that address earthquake detection and the physics of rupture and the timing of earthquakes.

Target Audience
We aim to elicit new connections among these diverse fields, identify novel tools and models that can be transferred from one to the other, and explore novel ML applications that will benefit from ML algorithms paradigm. We believe that a successful workshop will lead to new research directions in a variety of areas and will also inspire the development of novel theories and tools.

Schedule

08:30 AM	**Introduction**	*Pyrak-Nolte, Rustad, Baraniuk*
08:40 AM	**Paul Johnson**	*Johnson*
09:05 AM	**Greg Beroza, Mostafa Mousavi, and Weiqiang Zhu.**	*Beroza*
09:30 AM	**Maarten de Hoop**	*de Hoop*
09:55 AM	**Karianne Jodine Bergen**	*Bergen*
10:20 AM	**Coffee Break**	
10:40 AM	**Ping Lu**	*Lu*
10:40 AM	**Mauricio Araya-Polo**	*Araya*
10:45 AM	**Jorge Guevara**	*Guevara Diaz*
10:45 AM	**Ben Yuxing**	*Ben*

10:50 AM	**Timothy Draelos**	*Draelos*
10:50 AM	**Zachary Ross**	*Ross*
10:55 AM	**Ben Moseley**	*Moseley*
10:55 AM	**Men-Andrin Meier**	*Meier*
11:00 AM	**Mathieu Chambefort**	*Chambefort*
11:00 AM	**Xiaojin Tan**	*Tan*
11:05 AM	**Zheng Zhou**	*Lin*
11:05 AM	**Isabell Leang**	*Leang*
11:10 AM	**Cheng Zhan**	*Zhan*
11:10 AM	**Tan Nguyen**	
11:15 AM	**Laura Pyrak-Nolte**	*Pyrak-Nolte*
11:15 AM	**Poster Session**	
12:00 PM	**Lunch**	
02:00 PM	**Bertrand Rouet-Leduc**	*Rouet-Leduc*
02:20 PM	**Joan Bruna**	*Bruna*
02:40 PM	**Claudia Hulbert**	*Hulbert*
03:00 PM	**Coffee Break**	
03:30 PM	**Ivan Dokmanic**	*Dokmanic*
03:50 PM	**Joe Morris**	*Morris*
04:10 PM	**Youzou Lin**	*Lin*
04:30 PM	**Panel Discussion**	*Baraniuk, de Hoop, Johnson*

Abstracts (26):

Abstract 1: **Introduction in Machine Learning for Geophysical & Geochemical Signals**, *Pyrak-Nolte, Rustad, Baraniuk* 08:30 AM

Introductory comments by organizers

Abstract 2: **Paul Johnson in Machine Learning for Geophysical & Geochemical Signals**, *Johnson* 08:40 AM

Probing Earthquake Fault Slip using Machine Learning

Earthquakes take place when two juxtaposed fault blocks are stressed sufficiently to overcome the frictional force holding them in place and they abruptly slip relative to each other. Earthquake faults exhibit a continuum of behaviors ranging from stick slip associated with strong shaking, to slow slip which is primarily aseismic, to very slow slip that is both aseismic and can take place over hours to months. We are characterizing faulting physics by analyzing with machine learning continuous acoustic data streams in the laboratory and continuous seismic data streams in Earth. We use as labels characteristics of the measured fault slip behavior in the laboratory such as the fault friction, shear displacement and fault thickness. In Earth, we use surface displacement as determined by Global Positioning Systems (GPS). Other data data such as INSAR can be used as well. We find that the laboratory acoustic data and the Earth seismic data are a type of Rosetta Stone revealing fault characteristics at all times and fault displacements. This is a surprising observation because previously we believed most or much of the signal was noise. Here we describe an overview of recent work in this area and also describe recent efforts on parallel problems such as volcanoes and geysers.

Abstract 3: **Greg Beroza, Mostafa Mousavi, and Weiqiang Zhu. in Machine Learning for Geophysical & Geochemical Signals**, *Beroza* 09:05 AM

Deep Learning of Earthquake Signals

Gregory C. Beroza, S. Mostafa Mousavi, and Weiqiang Zhu

Diverse algorithms have been developed for efficient earthquake signal detection and processing. These algorithms are becoming increasingly important as seismologists strive to extract as much insight as possible from exponentially increasing volumes of continuous seismic data. Waveform similarity search, based on the premise that adjacent earthquakes generate similar waveforms, is now widely and effectively used to detect earthquakes too small to appear routinely in earthquake catalogs. Machine learning has the potential to generalize this similarity search from strict waveform similarity to waveforms that have similar characteristics. Convolutional and recurrent networks have each been shown to be promising tools for earthquake signal detection, and we have developed a deep convolutional-recurrent network to combine the advantages of each. This architecture is well-suited to learn both the spectral and temporal characteristics of earthquake signals. We have applied it to different, but inter-related tasks in earthquake analysis, including: earthquake detection, classification of continuous seismic data into P-waves, S-waves, and noise, and the problem of de-noising of earthquake signals. In our presentation we demonstrate the performance of deep learning applied to seismic signals for each of these tasks.

Abstract 5: **Karianne Jodine Bergen in Machine Learning for Geophysical & Geochemical Signals**, *Bergen* 09:55 AM

Towards data-driven earthquake detection: Extracting weak seismic signals with locality-sensitive hashing

Extracting weak earthquake signals from continuous waveform data recorded by sensors in a seismic network is a fundamental and challenging task in seismology. In this talk, I will present Fingerprint and Similarity Thresholding (FAST; Yoon et al, 2015), a computationally efficient method for large-scale earthquake detection. FAST adapts technology used for rapid audio identification to the problem of extracting weak earthquake signals in continuous seismic data. FAST uses locality-sensitive hashing, a data mining technique for efficiently identifying similar items in large data sets, to detect similar waveforms (candidate earthquakes) in continuous seismic data. A distinguishing feature of our approach is that FAST is an unsupervised detector; FAST can discover new sources without any template waveforms or waveform characteristics available as training data – a common situation for seismic data sets. In our recent work, we have extended FAST to enable earthquake detection using data from multiple sensors spaced tens or hundreds of kilometers apart (Bergen and Beroza, 2018), and optimized the FAST software for detection at scale (Rong et al., 2018). FAST can now detect earthquakes with previously unknown sources in 10-year, multi-sensor seismic data sets without training data – a capability that was not previously available for seismic data analysis.

Abstract 6: **Coffee Break in Machine Learning for Geophysical & Geochemical Signals**, 10:20 AM

Poster Spotlight

*Mauricio Araya-Polo, Stuart Farris and Manuel Florez, Combining Unsupervised and Supervised Deep Learning approaches for Seismic Tomography
Signals from inner earth, seismic waveforms, are heavily manipulated before human interpreters have a chance of figuring the subsurface structures. That manipulation adds modeling biases and it is limited by methodological shortcomings. Alternatively, using waveforms directly is becoming possible thanks to current Deep Learning (DL) advances such as (Araya-Polo et al., 2017 and 2018; Lin et al., 2017). Further extending that work, we present a DL approach that takes realistic raw seismic waveforms as inputs and produces subsurface velocity models as output. When insufficient data is used for training, DL algorithms tend to either over-fit or fail completely. Gathering large amounts of labeled and standardized seismic data sets is not straight forward. We address this shortage of quality data by building a Generative Adversarial Network (GAN) to augment our original training data set, which then is used by the DL seismic tomography as input.

*Yuxing Ben, Chris James, Dingzhou Can, Drilling State Classification with Machine Learning
The sensors on drilling rigs and production sites are leading oil and gas companies to mine so-called big data. Leveraging historical time series data and real-time drilling data can help drilling engineers improve rig and well delivery efficiencies; however, it can also help geoscientists understand the geophysical properties of the reservoir. In this case study, we describe how to use machine learning to classify drilling states. We investigated several machine learning methods and architectures including Random Forest tree models, Convolutional Neural Networks, and Recurrent Neural Networks which were then tested against 15 million rows of real, labeled drilling time-series data. We found that machine learning models were superior to rule based models. For wells drilled in two different onshore basins, the accuracies of our in-house rule based models were 70% and 90% respectively, while the accuracies of machine learning models were over 99%. The best identified machine learning model has been deployed in a drilling analytics platform and used to automatically detect the drilling state in realtime for use by Drilling Engineers to evaluate and analyze well performance.

*Jorge Guevara, Blanca Zadrozny, Alvaro Buoro, Ligang Lu, John Tolle, Jan Limbeck, Mingqi Wu, Defletf Hohl, An Interpretable Machine Learning Methodology for Well Data Integration and Sweet Spotting Identification.
The huge amount of heterogeneous data provided by the petroleum industry brings opportunities and challenges for applying machine learning methodologies. For instance, petrophysical data recorded in well logs, completions datasets and well production data also constitute good examples of data for training machine learning models with the aim of automating procedures and giving data-driven solutions to problems arisen in the petroleum industry. In this work, we present a machine learning methodology for oil exploration that 1) opens the possibility of integration of heterogeneous data such as completion, engineering, and well production data, as well as, petrophysical feature estimation from petrophysical data from horizontal and vertical wells; 2) it enables the discovery of new locations with high potential for production by using predictive modeling for sweet spotting identification; 3) it facilitates the analysis of the effect, role, and impact of some engineering decisions on production by means of interpretable Machine learning modeling, allowing the model validation; 4) it allows the incorporation of prior/expert knowledge by using Shape Constraint Additive Models and; 5) it enables the construction of hypothetical "what-if" scenarios for production prediction. Among the results, it is important to highlight that 1) performance improves by including prior knowledge via SCAMs, for example, we have a percentage change of 24% between the best RMSE result from black-box ML models vs a model that incorporates prior knowledge. 2) we were able to construct hypothetical what-if scenarios based on actual petrophysical data and hypothetical completion and engineering values, 3) we were able to assess the validity of ML models through effect analysis via conditonal plots.

*Ping Lu, Hunter Danque, Jianxiong Chen, Seth Brazell, and Mostafa Karimi, Enhanced Seismic Imaging with Predictive Neural Networks for Geophysics
Full-waveform inversion (FWI) has become a popular method to estimate elastic earth properties from seismic data, and it has great utility in seismic velocity model building and seismic reflectivity imaging in areas of complex salt. FWI is a non-linear data-fitting procedure that matches the predicted to observed waveform data given an initial guess of the subsurface parameters. The velocity model parameters are updated to reduce the misfit between the observed and predicted data until the misfit is sufficiently small. Sharp velocity boundaries such as between salt and sediment are often updated manually for each iteration based on the seismic reflectivity images. Here, we propose a predictive neural network architecture as a potential alternative to the complex FWI workflow. An unsupervised learning model of predicting of future frames in a video sequence is explored to simulate direct inversion procedures for seismic data. Such neural network architectures are comprised of two main components: an encoder based on convolutional neural networks (CNNs), and a recurrent neural networks (RNNs) for iteratively predicting geophysical velocity models. Both the proposed networks are able to robustly train individual layers and make a layer-specific prediction, which is compared with a target to produce an error term. It is then propagated to the subsequent network layers. With a few iterative training steps, the networks are capable of learning internal representations decoded from latent parameters of seismic wave propagation which controls how FWI velocity modelling converges. These representations learned from one dataset could be transferred to predict the future velocity model of a brand-new area where the shape of salt body is not well imaged or known. Altogether, experimental results generated from a real Gulf of Mexico seismic data suggest that the prediction represents a powerful framework for unsupervised learning, which provides an alternative approach to the FWI procedure to generate a high resolution velocity model including an accurate salt model and ultimately a sharp subsalt image.

*Zachary Ross, PhaseLink: A Deep Learning Approach to Seismic Phase Association
We present PhaseLink, a deep learning approach to seismic phase association. Seismic phase association is a fundamental task in seismology that pertains to linking together phase detections on different sensors that originate from a common earthquake. This task can be challenging because the number of sources is unknown, events frequently overlap in time, or can occur simultaneously in different parts of a network. Our PhaseLink approach has many desirable properties. First, it is trained entirely on synthetic simulated data (i.e., "sim-to-real"), and is thus easily portable to any tectonic regime. Second, it is straightforward to tune PhaseLink by simply adding examples of problematic cases to the training dataset -- whereas conventional

approaches require laborious adjusting of ad hoc hyperparameters. Third, we empirically demonstrate state-of-the-art performance in a wide range of settings. For instance, PhaseLink can precisely associate P- and S-picks to events that are separated by ~12 seconds in origin time. We expect PhaseLink to substantially improve many aspects of seismic analysis, including the resolution of seismicity catalogs, real-time seismic monitoring, and streamlined processing of large seismic datasets.

*Timothy Draelos, Stephen Heck, Jennifer Galasso, and Ronald Brogan, Seismic Phase Identification with a Merged Deep Neural Network
Seismic signals are composed of the seismic waves (phases) that reach a sensor, similar to the way speech signals are composed of phonemes that reach a listener's ear. We leverage ideas from speech recognition for the classification of seismic phases at a seismic sensor. Seismic Phase ID is challenging due to the varying paths and distances an event takes to reach a sensor, but there is consistent structure and ordering of the different phases arriving at the sensor. Together with scalar value measurements of seismic signal detections (horizontal slowness, amplitude, Signal-to-Noise Ratio (SNR), and the time since the previous signal detection), we use the seismogram and its spectrogram of detection waveforms as inputs to a merged deep neural network (DNN) with convolutional (CNN) and recurrent (LSTM) layers to learn the frequency structure over time of different phases. The binary classification performance of First-P phases versus non-First-P (95.6% class average accuracy) suggests a potentially significant impact on the reduction of false and missed events in seismic signal processing pipelines. Other applications include discrimination between noise and non-noise detections for induced seismicity networks and for early warning of large hazards

*Ben Moseley, Andrew Markham, and Tarje Nissen-Meyer, Fast Approximate Simulation of Seismic Waves with Deep Learning
The simulation of seismic waves is a core task in many geophysical applications, yet it is computationally expensive. As an alternative approach, we simulate acoustic waves in horizontally layered media using a deep neural network. In contrast to traditional finite-difference (FD) modelling, our network is able to directly approximate the recorded seismic response at multiple receiver locations in a single inference step, without needing to iteratively model the seismic wavefield through time. This results in an order of magnitude reduction in simulation time, from the order of 1 s for FD modelling to the order of 0.1 s using our approach. Such a speed improvement could lead to real-time seismic simulation applications and benefit seismic inversion algorithms based on forward modelling, such as full waveform inversion. Our network design is inspired by the WaveNet network originally used for speech synthesis. We train our network using 50,000 synthetic examples of seismic waves propagating through different horizontally layered velocity models. We are also able to alter our WaveNet architecture to carry out seismic inversion directly on the dataset, which offers a fast inversion algorithm.

* Men-Andrin Meier, Zachary Ross, Anshul Ramachandran, Ashwin Balakrishna, Suraj Nair, Peter Kundzicz, Zefeng Li, Egill Hauksson, Jennifer Andrews, Reliable Real-Time Signal/Noise Discrimination with Deep and Shallow Machine Learning Classifiers
In Earthquake Early Warning (EEW), every sufficiently impulsive signal is potentially the first evidence for an unfolding large earthquake. More often than not, however, impulsive signals are mere nuisance signals. One of the most fundamental - and difficult - tasks in EEW is to rapidly and reliably discriminate between real local earthquake signals, and any kind of other signal. Current EEW systems struggle to avoid discrimination errors, and suffer from false and missed alerts. In this study we show how machine learning classifiers can strongly improve real-time signal/noise discrimination. We develop and compare a series of non-linear classifiers with variable architecture depths, including random forests, fully connected, convolutional (CNN, Figure 1) and recurrent neural networks, and a generative adversarial network (GAN). We train all classifiers on the same waveform data set that includes 374k 3-component local earthquake records with magnitudes M3.0-9.1, and 946k impulsive noise signals. We find that the deep architectures significantly outperform the more simple ones. Using 3s long waveform snippets, the CNN and the GAN classifiers both reach 99.5% precision and 99.3% recall on an independent validation data set. Our results suggest that machine learning classifiers can strongly improve the reliability and speed of EEW alerts.

*Mathieu Chambefort, Nicolas Salaun, Emillie Chautru, Stephan Clemencon, Guillaume Poulain, Signal and Noise Detection using Recurrent Autoencoders on Seismic Marine Data
In order to meet the industrial constraints in the Big Data era, i.e. processing more and more seismic data (more than 106 shot points per marine seismic survey
from [Belz and Dolymnyj, 2018]) in a more timely, reliable and efficient manner (i.e. with a better signal enhancement, [Martin et al., 2015]), we develop a deep
learning approach based on recurrent LSTM ([Wong and Luo, 2018]) to the processing of seismic time series, so as to separate the signal from the noise based
on the encoded information. This contribution provides empirical evidence that the representation provided by the internal layers of the autoencoder deployed encodes
well the original information. More precisely, focus is here on the linear noise possibly blurring marine seismic data, which is mainly due to the tug and motor of the boat but can also be caused by bad weather or other elements, rig and other boats in the area ([Elboth et al., 2009]). The data under study are composed of massive synthetic shot points.
The goal pursued is to design an autoencoder capable of detecting the possible occurrence of linear noise in the data. The encoded information is next classified
and the results obtained are compared with those of a traditional technique, that essentially consists in applying directly a K -NN algorithm on the envelope of the
analytical signal, as if all the dataset comes from the same area.

*Xiaojin Tan and Eldad Haber■, Semantic Segmentation for Geophysical Data
Segmentation of geophysical data is the process of dividing a geophysical image into multiple geological units. This process is typically done manually by experts, it is time consuming and inefficient. In recent years, machine learning techniques such as Convolutional Neural Networks (CNNs) have been used for semantic segmentation. Semantic segmentation is the process that associates each pixel in a natural image with a labeled class. When attempting to use similar technology to automatically segment geophysical data there are a number of challenges to consider, in particular, data inconsistency, scarcity and complexity. To overcome these challenges, we develop a new process that we call geophysical semantic segmentation (GSS). This process addresses the pre-processing of geophysical data in order to enable learning, the enrichment of the data set (data augmentation) by using a geo-statistical technique, referred to as Multiple-Point Simulations (MPS) and finally, the training of such a data set based on a new neural network architecture called inverse Convolution Neural Networks (iCNN) that is

specifically developed to identify patterns. As demonstrated by the results on a field magnetic data set, this approach shows its competitiveness with human segmentation and indicates promising results.

*B Ravi Kiran and Stefan Milz, Aerial LiDAR reconstruction using Conditional GANS
Recently, aerial LiDAR data opened lots of new opportunities for many research disciplines like macroscopic geophysical analysis or archaeological investigations.
However, LiDAR measurements are expensive and the data is not widely distributed or accessible. We propose a novel method for image to image translation
performing HD-LiDAR reconstruction using RGB input images based on conditional GANs. The conditional mapping function of the generator G : [c; z] -> y is
transformed to G : [x; z] -> y , whereas y represents the reconstructed LiDAR map and c represents the condition. c is replaced by the aligned aerial camera image x .
z represents the noise. Our approach is able to reconstruct LiDAR data as elevation maps based on small scaled training data, which includes RGB and LiDAR sample
pairs based on 256 ■ 256 image matrices. The model offers the opportunity to complete geophysical LiDAR databases, where measurements are missing. The
method is validated on the ISPRS dataset with an overall rRMSE of 14.53% .

Zheng Zhou, Youzuo Lin, Zhongping Zhang, Zan Wang, Robert Dilmore and George Guthrie, CO2 and Brine Leakage Detection Using Multi-Physics-Informed Convolutional Neural Networks
In carbon capture and sequestration, it is crucial to build effective monitoring techniques to detect both brine and CO2 leakage from legacy wells into underground sources of drinking water. The CO2 and brine leakage detection methods rely on geophysical observations from different physical domains. Most of the current detection methods are built on physical models, and the leakage mass of CO2 and brine are detected separately. However, those physics-driven methods can be computationally demanding and yields low detection accuracy. In this paper, we developed a novel end-to-end data-driven detection method, called multi-physics-informed convolutional neural network (Multi-physics CNN), which directly learns a mapping relationship between physical measurements and leakage mass. Our Multi-physical CNN takes simulated reflection seismic and pressure data as inputs, and captures different patterns in leakage process. In particular, we capture two types of multi-physical features from seismic and pressure data, respectively. With those features, we can further detect the CO2 and brine leakage mass, simultaneously. We evaluate our novel method for CO2 and brine leakage mass detection task on simulated multi-physical datasets generated using Kimberlina 1.2 model. Our results show that our Multi-physics CNN yields promising results in detecting both leakage mass of CO2 and brine.

Abstract 7: **Ping Lu in Machine Learning for Geophysical & Geochemical Signals**, *Lu* 10:40 AM

Enhanced Seismic Imaging with Predictive Neural Networks for Geophysics
Ping Lu, Yanyan Zhang, Jianxiong Chen, Seth Brazell, Mostafa Karimi
Anadarko Petroleum Corporation, Houston, and Texas A&M University--College Station

We propose a predictive neural network architecture that can be utilized to update reference velocity models as inputs to full waveform inversion (FWI). Deep learning models are explored to augment velocity model building workflows during 3D seismic volume reprocessing in salt-prone environments. Specifically, a neural network architecture, with 3D convolutional, de-convolutional layers, and 3D max-pooling, is designed to take standard amplitude 3D seismic volumes as an input. Enhanced data augmentations through generative adversarial networks and a weighted loss function enable the network to train with few sparsely annotated slices. Batch normalization is also applied for faster convergence. Moreover, a 3D probability cube for salt bodies is generated through ensembles of predictions from multiple models in order to reduce variance. Velocity models inferred from the proposed networks provide opportunities for FWI forward models to converge faster with an initial condition closer to the true model. In each iteration step, the probability cubes of salt bodies inferred from the proposed networks can be used as a regularization term in FWI forward modelling, which may result in an improved velocity model estimation while the output of seismic migration can be utilized as an input of the 3D neural network for subsequent iterations.

Abstract 8: **Mauricio Araya-Polo in Machine Learning for Geophysical & Geochemical Signals**, *Araya* 10:40 AM

Mauricio Araya-Polo, Stuart Farris and Manuel Florez
Standford University and Shell International Exploration & Production Inc.

Combining Unsupervised and Supervised Deep Learning approaches for Seismic Tomography

Signals from inner earth, seismic waveforms, are heavily manipulated before human interpreters have a chance of figuring the subsurface structures. That manipulation adds modeling biases and it is limited by methodological shortcomings. Alternatively, using waveforms directly is becoming possible thanks to current Deep Learning (DL) advances such as (Araya-Polo et al., 2017 and 2018; Lin et al., 2017). Further extending that work, we present a DL approach that takes realistic raw seismic waveforms as inputs and produces subsurface velocity models as output. When insufficient data is used for training, DL algorithms tend to either over-fit or fail completely. Gathering large amounts of labeled and standardized seismic data sets is not straight forward. We address this shortage of quality data by building a Generative Adversarial Network (GAN) to augment our original training data set, which then is used by the DL seismic tomography as input.

Abstract 9: **Jorge Guevara in Machine Learning for Geophysical & Geochemical Signals**, *Guevara Diaz* 10:45 AM

Jorge Guevara, Blanca Zadrozny, Alvaro Buoro, Ligang Lu, John Tolle, Jan Limbeck, Mingqi Wu, Defletf Hohl
IBM Research and Shell Inc.

An Interpretable Machine Learning Methodology for Well Data Integration and Sweet Spotting Identification.
The huge amount of heterogeneous data provided by the petroleum industry brings opportunities and challenges for applying machine learning methodologies aimed to optimize and automate process and procedures in this area. For instance, petrophysical data recorded in well logs, completions datasets and well production data also constitute good

examples of data for training machine learning models with the aim of automating procedures and giving data-driven solutions to problems arisen in the petroleum industry. In this work, we present a machine learning methodology for oil exploration that 1) integrates heterogeneous well data such as: completions, engineering values, well production data and petrophysical data; 2) performs feature engineering of petrophysical data from horizontal and vertical wells using Gaussian Process Regression (Kriging); 3) it enables the discovery of new locations with high potential for production by using machine learning modeling for sweet spotting identification; 4) it facilitates the analysis of the effect, role, and impact of some engineering decisions on production by means of interpretable Machine learning modeling; 5) it allows the incorporation of prior/expert knowledge by using Shape Constraint Additive Models and; 6) it enables the construction of hypothetical "what-if" scenarios for production prediction, by means of conditional plots based on residual plots analysis. We validated this methodology using real well production data. We used nested leave-one-out cross-validation for assessing the generalization power of models. Among the results, it is important to highlight that 1) performance improves by including prior knowledge via SCAMs, for example, we have a percentage change of 24\% between the best RMSE result from black-box ML models vs a model that incorporates prior knowledge. 2) we were able to construct hypothetical what-if scenarios based on actual petrophysical data and hypothetical completion and engineering values, 3) we were able to assess the validity of ML models through effect analysis via conditional plots.

Abstract 10: **Ben Yuxing in Machine Learning for Geophysical & Geochemical Signals**, *Ben* 10:45 AM

Drilling State Classification with Machine Learning
Yuxing Ben, Chris James, Dingzhou Cao
Advanced Analytics and Emerging Technology, Anadarko Petroleum Corporation

The sensors on drilling rigs and production sites are leading oil and gas companies to mine so-called big data. Leveraging historical time series data and real-time drilling data can help drilling engineers improve rig and well delivery efficiencies; however, it can also help geoscientists understand the geophysical properties of the reservoir. In this case study, we describe how to use machine learning to classify drilling states. We investigated several machine learning methods and architectures including Random Forest tree models, Convolutional Neural Networks, and Recurrent Neural Networks which were then tested against 15 million rows of real, labeled drilling time-series data. We found that machine learning models were superior to rule based models. For wells drilled in two different onshore basins, the accuracies of our in-house rule based models were 70% and 90% respectively, while the accuracies of machine learning models were over 99%. The best identified machine learning model has been deployed in a drilling analytics platform and used to automatically detect the drilling state in realtime for use by Drilling Engineers to evaluate and analyze well performance.

Abstract 11: **Timothy Draelos in Machine Learning for Geophysical & Geochemical Signals**, *Draelos* 10:50 AM

Seismic Phase Identification with a Merged Deep Neural Network

Timothy J. Draelos, Stephen Heck, Jennifer Galasso, Ronald Brogan
Sandia National Laboratories & ENSCO, Inc.

Seismic signals are composed of the seismic waves (phases) that reach a sensor, similar to the way speech signals are composed of phonemes that reach a listener's ear. We leverage ideas from speech recognition for the classification of seismic phases at a seismic sensor. Seismic Phase ID is challenging due to the varying paths and distances an event takes to reach a sensor, but there is consistent structure and ordering of the different phases arriving at the sensor. Together with scalar value measurements of seismic signal detections (horizontal slowness, amplitude, Signal-to-Noise Ratio (SNR), and the time since the previous signal detection), we use the seismogram and its spectrogram of detection waveforms as inputs to a merged deep neural network (DNN) with convolutional (CNN) and recurrent (LSTM) layers to learn the frequency structure over time of different phases. The binary classification performance of First-P phases versus non-First-P (95.6% class average accuracy) suggests a potentially significant impact on the reduction of false and missed events in seismic signal processing pipelines. Other applications include discrimination between noise and non-noise detections for induced seismicity networks and for early warning of large hazards.

Abstract 12: **Zachary Ross in Machine Learning for Geophysical & Geochemical Signals**, *Ross* 10:50 AM

PhaseLink: A Deep Learning Approach to Seismic Phase Association
Zachary Ross
California Institute of Technology

We present PhaseLink, a deep learning approach to seismic phase association. Seismic phase association is a fundamental task in seismology that pertains to linking together phase detections on different sensors that originate from a common earthquake. This task can be challenging because the number of sources is unknown, events frequently overlap in time, or can occur simultaneously in different parts of a network. Our PhaseLink approach has many desirable properties. First, it is trained entirely on synthetic simulated data (i.e., "sim-to-real"), and is thus easily portable to any tectonic regime. Second, it is straightforward to tune PhaseLink by simply adding examples of problematic cases to the training dataset -- whereas conventional approaches require laborious adjusting of ad hoc hyperparameters. Third, we empirically demonstrate state-of-the-art performance in a wide range of settings. For instance, PhaseLink can precisely associate P- and S-picks to events that are separated by ~12 seconds in origin time. We expect PhaseLink to substantially improve many aspects of seismic analysis, including the resolution of seismicity catalogs, real-time seismic monitoring, and streamlined processing of large seismic

Abstract 13: **Ben Moseley in Machine Learning for Geophysical & Geochemical Signals**, *Moseley* 10:55 AM

Fast approximate simulation of seismic waves with deep learning
Ben Moseley, Andrew Markham, and Tarje Nissen-Meyer
Centre for Doctoral Training in Autonomous Intelligent Machines and Systems, University of Oxford, UK & Department of Earth Sciences, University of Oxford, UK

The simulation of seismic waves is a core task in many geophysical applications, yet it is computationally expensive. As an alternative approach, we simulate acoustic waves in horizontally layered media using a deep neural network. In contrast to traditional finite-difference (FD) modelling, our network is able to directly approximate the recorded seismic response at multiple receiver locations in a single inference step, without needing to iteratively model the seismic wavefield through time.

This results in an order of magnitude reduction in simulation time, from the order of 1 s for FD modelling to the order of 0.1 s using our approach. Such a speed improvement could lead to real-time seismic simulation applications and benefit seismic inversion algorithms based on forward modelling, such as full waveform inversion. Our network design is inspired by the WaveNet network originally used for speech synthesis. We train our network using 50,000 synthetic examples of seismic waves propagating through different horizontally layered velocity models. We are also able to alter our WaveNet architecture to carry out seismic inversion directly on the dataset, which offers a fast inversion algorithm.

Abstract 14: **Men-Andrin Meier in Machine Learning for Geophysical & Geochemical Signals**, *Meier* 10:55 AM

Reliable Real-Time Signal/Noise Discrimination with Deep and Shallow Machine Learning Classifiers
Men-Andrin Meier, Zachary Ross, Anshul Ramachandran, Ashwin Balakrishna, Suraj Nair, Peter Kundzicz, Zefeng Li, Egill Hauksson, Jennifer Andrews
California Institute of Technology

In Earthquake Early Warning (EEW), every sufficiently impulsive signal is potentially the first evidence for an unfolding large earthquake. More often than not, however, impulsive signals are mere nuisance signals. One of the most fundamental - and difficult - tasks in EEW is to rapidly and reliably discriminate between real local earthquake signals, and any kind of other signal. Current EEW systems struggle to avoid discrimination errors, and suffer from false and missed alerts. In this study we show how machine learning classifiers can strongly improve real-time signal/noise discrimination. We develop and compare a series of non-linear classifiers with variable architecture depths, including random forests, fully connected, convolutional (CNN, Figure 1) and recurrent neural networks, and a generative adversarial network (GAN). We train all classifiers on the same waveform data set that includes 374k 3-component local earthquake records with magnitudes M3.0-9.1, and 946k impulsive noise signals. We find that the deep architectures significantly outperform the more simple ones. Using 3s long waveform snippets, the CNN and the GAN classifiers both reach 99.5% precision and 99.3% recall on an independent validation data set. Our results suggest that machine learning classifiers can strongly improve the reliability and speed of EEW alerts. Figure

Abstract 15: **Mathieu Chambefort in Machine Learning for Geophysical & Geochemical Signals**, *Chambefort* 11:00 AM

Signal and Noise Detection using Recurrent Autoencoders on Seismic Marine Data

Mathieu Chambefort, Nicolas Salaun, Emilie Chautru, Stephan Clémençon and Guillaume Poulain

MINES ParisTech - PSL University Centre de Géosciences, CGG, and Telecom ParisTech, LTCI, Université Paris Saclay

In the Big Data era, geophysics are faced with new industrial contrains like processing more and more seismic data (more than 106 shot points per marine seismic survey [Belz and Dolymnyj, 2018]) in a more timely, reliable and ef■cient manner (improving signal enhancement, [Martin et al., 2015]). To deal with these challenges, we develop a deep learning approach based on recurrent LSTM ([Wong and Luo, 2018]) to the processing of seismic time series; this separates the signal from the noise based on the encoded information. This contribution provides empirical evidence that the representation provided by the internal layers of the deployed autoencoder encodes the original information well. More precisely, focus is here on the linear noise that possibly blurs marine seismic data ([Elboth et al., 2009]). The data under study is composed of massive synthetic shot points. The goal pursued is to design an autoencoder capable of detecting the possible occurrence of linear noise in the data. Next, the encoded information is classi■ed. The obtained results are compared with those of a traditional technique, which essentially consists in applying directly a K-NN algorithm on the envelope of the analytical signal, as if all the dataset came from the same area.

Abstract 16: **Xiaojin Tan in Machine Learning for Geophysical & Geochemical Signals**, *Tan* 11:00 AM

Semantic Segmentation for Geophysical Data

Xiaojin Tan and Eldad Haber
The University of British Columbia, Vancouver, BC, Canada

Segmentation of geophysical data is the process of dividing a geophysical image into multiple geological units. This process is typically done manually by experts, it is time consuming and inefficient. In recent years, machine learning techniques such as Convolutional Neural Networks (CNNs) have been used for semantic segmentation. Semantic segmentation is the process that associates each pixel in a natural image with a labeled class. When attempting to use similar technology to automatically segment geophysical data there are a number of challenges to consider, in particular, data inconsistency, scarcity and complexity. To overcome these challenges, we develop a new process that we call geophysical semantic segmentation (GSS). This process addresses the pre-processing of geophysical data in order to enable learning, the enrichment of the data set (data augmentation) by using a geo-statistical technique, referred to as Multiple-Point Simulations (MPS) and finally, the training of such a data set based on a new neural network architecture called inverse Convolution Neural Networks (iCNN) that is specifically developed to identify patterns. As demonstrated by the results on a field magnetic data set, this approach shows its competitiveness with human segmentation and indicates promising results.

Abstract 17: **Zheng Zhou in Machine Learning for Geophysical & Geochemical Signals**, *Lin* 11:05 AM

CO2 and Brine Leakage Detection Using
Multi-Physics-Informed Convolutional Neural Networks

Zheng Zhou, Youzuo Lin, Zhongping Zhang, Zan Wang, Robert Dilmore and George Guthrie

Electrical Engineering Department at State university of New York at Buffalo, Los Alamos National Laboratory, and National Energy Technology Laboratory, United States Department of Energy, Pittsburgh, PA 15236.

In carbon capture and sequestration, it is crucial to build effective monitoring techniques to detect both brine and CO2 leakage from legacy wells into underground sources of drinking water. The CO2 and brine leakage detection methods rely on geophysical observations from

different physical domains. Most of the current detection methods are built on physical models, and the leakage mass of CO2 and brine are detected separately. However, those physics-driven methods can be computationally demanding and yields low detection accuracy. In this paper, we developed a novel end-to-end data-driven detection method, called multi-physics-informed convolutional neural network (Multi-physics CNN), which directly learns a mapping relationship between physical measurements and leakage mass. Our Multi-physical CNN takes simulated reflection seismic and pressure data as inputs, and captures different patterns in leakage process. In particular, we capture two types of multi-physical features from seismic and pressure data, respectively. With those features, we can further detect the CO2 and brine leakage mass, simultaneously. We evaluate our novel method for CO2 and brine leakage mass detection task on simulated multi-physical datasets generated using Kimberlina 1.2 model. Our results show that our Multi-physics CNN yields promising results in detecting both leakage mass of CO2 and brine.

Abstract 18: **Isabell Leang in Machine Learning for Geophysical & Geochemical Signals**, *Leang* 11:05 AM

Aerial LiDAR reconstruction using conditional GANs
Isabelle Leang, B Ravi Kiran and Stefan Milz
Recently, aerial LiDAR data opened lots of new opportunities for many research disciplines like macroscopic geophysical analysis or archaeological investigations. However, LiDAR measurements are expensive and the data is not widely distributed or accessible. We propose a novel method for image to image translation performing HD-LiDAR reconstruction using RGB input images based on conditional GANs. The conditional mapping function of the generator G : [c; z] -> y is transformed to G : [x; z] -> y , whereas y represents the reconstructed LiDAR map and c represents the condition. c is replaced by the aligned aerial camera image x . z represents the noise. Our approach is able to reconstruct LiDAR data as elevation maps based on small scaled training data, which includes RGB and LiDAR sample pairs based on 256 x 256 image matrices. The model offers the opportunity to complete geophysical LiDAR databases, where measurements are missing. The method is validated on the ISPRS dataset with an overall rRMSE of 14.53%.

Abstract 19: **Cheng Zhan in Machine Learning for Geophysical & Geochemical Signals**, *Zhan* 11:10 AM

Deep Semi-Supervised Learning Approach in Characterizing Salt on Seismic Images

Licheng Zhang, Zhenzhen Zhong, Meng Zhang, Tianxia Zhao, Varun Tyagi, Cheng Zhan

The salt body characterization is crucial in exploration and drilling. Due to its mobility, salt can move extensively to create diapirs, which generate significant traps for hydrocarbons, meanwhile, they present drilling hazards, as salt intrusion distorts the stress field making wellbore stability challenging in the geomechanical models. Here we utilized deep learning to identify salt body based on seismic images. Many techniques from the domains of geophysics and data science, have been successfully incorporated into the work-flow. The seismic images are produced from various locations. Here we use convolutional neural network that is the main methodology to process images segmentations. The underlying architecture is dedicated to restoring pixel position. In addition, the highlight here is Semi-Supervised learning, and we utilized the large unlabeled test set to gain more understanding of the data distribution, and the pseudo labeling of unlabeled test set comes from prediction. The metric implemented is “IOU”, Intersection over Union, which fundamentally measures how much area the predicted salt body overlay with the true answer. Our IOU score is 0.849, equivalent to 95% of the predicted salt body is correct. Challenges still exist as geology varies across locations, and the corresponding features might not share similar statistical properties.

Abstract 20: **Tan Nguyen in Machine Learning for Geophysical & Geochemical Signals**, 11:10 AM

Tremor Generative Adversarial Networks: A Deep Generative Model Approach for Geophysical Signal Generation

Inspired by the recent success of the Generative Adversarial Networks (GANs) for images, we
propose to employ GANs to generate realistic geophysical signals from labeled data. Signals, here, include seismicity, sedimentary sequences, geological models etc. We present a preliminary application of a GAN to generate tremors: Synthetic tremors generated by one of
our GANs, trained with data collected in Mexico. Studying the trained GANs facilitates our understanding of the data generating process. These GANs can also be inverted into
inference algorithms that capture intrinsic properties of the generating process. GAN-generated tremors can be used as templates to help detect additional tremors and potentially result in better generalization to new sensor signals.

Abstract 21: **Laura Pyrak-Nolte in Machine Learning for Geophysical & Geochemical Signals**, *Pyrak-Nolte* 11:15 AM

Data Challenge: Machine Learning for Earthquake Detection and Rupture Timing

Laura Pyrak-Nolte, Richard Baraniuk, Greg Beroza, Maarten de Hoop, Brad Hager, Eugene Ilton, Paul Johnson, Steve Laubach, Alan Levander, Semechah Lui, Joe Morris, Beatrice Rivera, James Rustad

Affiliations: Purdue University, Rice University, Stanford University, MIT, PNNL, LANL, Bureau of Economic Geology, University of Toronto, LLNL, Department of Energy-Basic Energy Sciences

Major breakthroughs and discoveries in geophysics are anticipated because of increases in computational power, massive sensor deployments that yield massive datasets, and advancements in machine learning algorithms. However, owing to the inherent complexity, machine learning methods are prone to misapplication, lack of transparency, and often do not attempt to produce interpretable models. Moreover, due to the flexibility in specifying machine learning models, results are often insufficiently documented in research articles, hindering both reliable assessment of model validity and consistent interpretation of model outputs. By providing documented datasets and challenging teams to apply fully documented workflows for machine learning approaches, we expect to accelerate progress in the application of data science to longstanding research issues in geophysics.

In this poster presentation, the guidelines for a challenge problem will be given. Challenge 1 will address the physics of rupture and timing of earthquakes (from laboratory data collected during shearing of gouge-filled faults). While using the data set in the challenge, the

expected reported information pertains to supervised and unsupervised machine learning components:
● the architecture of the machine-learning approach and why it was chosen;
● the loss function and learning rule;
● preprocessing designed and applied as appropriate;
● description of and choice of the set of hyperparameters;
● description of featurization or feature learning.
We expect the design of the machine learning approach to be an iterative process and seek a description of this. In view of the lack of ground truth data in general, while being a physics-based challenge, we invite proposed metrics to validate and compare the performance of the different results. Information will be provided on how to obtain the data, timelines for completion of the challenge, and reporting of results.

Acknowledgment: US Department of Energy, Office of Basic Energy Sciences, Chemical Sciences, Geosciences and Biosciences Division.

Abstract 24: **Bertrand Rouet-Leduc in Machine Learning for Geophysical & Geochemical Signals**, *Rouet-Leduc* 02:00 PM

Estimating the State of Faults from the Full Continuous Seismic Data Using Machine Learning

Nearly all aspects of earthquake rupture are controlled by the friction along the fault that progressively increases with tectonic forcing, but in general cannot be directly measured. Using machine learning, we show that instantaneous statistical characteristics of the seismic data are a fingerprint of the fault zone frictional state in laboratory experiments. Using a similar methodology in Earth, where we rely on other geophysical datasets as labels in order to extract informative signals from raw seismic waves, we show that subduction zones are continuously broadcasting a tremor-like signal that precisely informs of fault displacement rate throughout their slow earthquake slip cycle. We posit that this signal provides indirect, real-time access to frictional properties of megathrusts and may ultimately reveal a connection between slow slip and megaquakes

Abstract 25: **Joan Bruna in Machine Learning for Geophysical & Geochemical Signals**, *Bruna* 02:20 PM

Geometric Deep Learning for Many-Particle and non Euclidean Systems

Across many areas of science, one is required to process data defined on irregular and non-Euclidean domains. For example, in particle physics, measurements in the LHC are highly variable particle collisions with cylindrical calorimeters, whereas the IceCube detector looks for neutrinos using an irregular 3d array of sensors. Despite such non-Euclidean structure, many of these tasks satisfy essential geometric priors, such as stability to deformations. In this talk, I will describe a broad family of neural architectures that leverage such geometric priors to learn efficient models with provable stability. I will also describe recent and current progress on several applications including particle physics and inverse problems.

Abstract 26: **Claudia Hulbert in Machine Learning for Geophysical & Geochemical Signals**, *Hulbert* 02:40 PM

Machine Learning Reveals the Coupling Between Slow Slips and Major Earthquakes

The potential connection between slow slips and earthquakes of large magnitude in subduction zones remains an open question in seismology. Slow slips (earthquakes releasing energy over long periods of times, up to several months) have been observed preceding major earthquake ruptures, suggesting that they may couple to or evolve into a megaquake.

We rely on supervised machine learning algorithms to analyze vast amounts of continuous seismic data, with the goal of identifying hidden signals preceding earthquakes. We find that continuous seismic signals identified in our previous studies of slow slip events carry information about the timing of impending earthquakes of large magnitude. Our results suggest that large earthquakes occur almost systematically in the same phase of the slow slip cycle, and point to a systematic, large-scale coupling between slow slip events and major earthquakes.

Abstract 28: **Ivan Dokmanic in Machine Learning for Geophysical & Geochemical Signals**, *Dokmanic* 03:30 PM

I will present a new learning-based approach to ill-posed inverse problems. Instead of directly learning the ill-posed inverse mapping, we learn an ensemble of simpler mappings from the data to the projections of the unknown model into random low-dimensional subspaces. We choose structured subspaces of piecewise-constant images on random Delaunay triangulations. With this choice, the projected inverse maps are simpler to learn in terms of robustness and generalization error. We form the reconstruction by combining the estimated subspace projections. This allow us to address inverse problems with extremely sparse data and still get good reconstructions of the unknown geometry; it also makes our method robust against arbitrary data corruptions not seen during training. Further, it marginalizes the role of the training dataset which is essential for applications in geophysics where ground-truth datasets are exceptionally scarce.

Abstract 29: **Joe Morris in Machine Learning for Geophysical & Geochemical Signals**, *Morris* 03:50 PM

Towards Realtime Hydraulic Fracture Monitoring using Machine Learning and Distributed Fiber Sensing

Joseph Morris, Christopher Sherman, Robert Mellors, Frederick Ryerson, Charles Yu, Michael Messerly

Abstract: Hydraulic fracturing operations (“pumping jobs”) are typically planned well in advance and do not allow for on-the-fly modification of control parameters, such as pumping rate and viscosity enhancement, that can be used to optimize the efficacy of the operation. Monitoring technologies, such as microseismic, have enabled an iterative cycle where observations of one pumping job may influence the selection of parameters of subsequent jobs. However, the significant time lag introduced by data processing and interpretation means that the iterative cycle may take weeks. We seek to enable a future where data collected during a job enables actionable, realtime decision making. Recent advances in distributed acoustic sensor (DAS) technology have produced a source of abundant new data for monitoring processes in the subsurface. Because of the massive dataset size (TB per day), developing a machine learning approach for interpreting DAS data is essential for effective use, such as in operational situations, which require near-realtime results. In our work, we use the massively parallel multi-physics code GEOS to generate a catalog of synthetic DAS measurements that are typical of those recorded during the stimulation of a hydraulic fracture. We then relate physical observables in the model

such as the extents of the generated fractures, fluid flow, and interactions with pre-existing rock fractures to the DAS. These data quantify the potential of DAS measurements for revealing subsurface processes in realtime. Determining how best to construct and train a neural network is challenging. We will present our specific approach to building a deep neural network, including the nature of the training data and subsequent success of the network in identifying features.
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

Abstract 30: **Youzou Lin in Machine Learning for Geophysical & Geochemical Signals**, *Lin* 04:10 PM

Accurate and Efficient Seismic Waveform-Inversion with Convolutional Neural Networks

Seismic full-waveform inversion has become a promising tool for velocity estimation in complex geological structures. The traditional seismic full-waveform inversion problems are usually posed as nonlinear optimization problems. Solving full-waveform inversion can be computationally challenging for two major reasons. One is the expensive computational cost and the other is the issue of local minima. In this work, we develop an end-to-end data-driven inversion technique, called “InversionNet”, to learn a regression relationship from seismic waveform datasets to subsurface models. Specifically, we build a novel deep convolutional neural network with an encoder-decoder structure, where the encoder learns an abstract representation of the seismic data, which is then used by the decoder to produce a subsurface model. We further incorporate atrous convolutions in our network structure to account for contextual information from the subsurface model. We evaluate the performance of our InversionNet with synthetic seismic waveform data. The experiment results demonstrate that our InversionNet not only yields accurate inversion results but also produces almost real-time inversion.

Workshop on Ethical, Social and Governance Issues in AI

Chloe Bakalar, Sarah Bird, Tiberio Caetano, Edward W Felten, Dario Garcia, Isabel Kloumann, Finnian Lattimore, Sendhil Mullainathan, D. Sculley

Room 516 AB, Fri Dec 07, 08:00 AM

Abstract

Ethics is the philosophy of human conduct: It addresses the question “how should we act?” Throughout most of history the repertoire of actions available to us was limited and their consequences constrained in scope and impact through dispersed power structures and slow trade. Today, in our globalised and networked world, a decision can affect billions of people instantaneously and have tremendously complex repercussions. Machine learning algorithms are replacing humans in making many of the decisions that affect our everyday lives. How can we decide how machine learning algorithms and their designers should act? What is the ethics of today and what will it be in the future?

In this one day workshop we will explore the interaction of AI, society, and ethics through three general themes.

Advancing and Connecting Theory: How do different fairness metrics relate to one another? What are the trade-offs between them? How do fairness, accountability, transparency, interpretability and causality relate to ethical decision making? What principles can we use to guide us in selecting fairness metrics within a given context? Can we connect these principles back to ethics in philosophy? Are these principles still relevant today?

Tools and Applications: Real-world examples of how ethical considerations are affecting the design of ML systems and pipelines. Applications of algorithmic fairness, transparency or interpretability to produce better outcomes. Tools that aid identifying and or alleviating issues such as bias, discrimination, filter bubbles, feedback loops etc. and enable actionable exploration of the resulting trade-offs.

Regulation: With the GDPR coming into force in May 2018 it is the perfect time to examine how regulation can help (or hinder) our efforts to deploy AI for the benefit of society. How are companies and organisations responding to the GDPR? What aspects are working and what are the challenges? How can regulatory or legal frameworks be designed to continue to encourage innovation, so society as a whole can benefit from AI, whilst still providing protection against its harms.

This workshop is designed to be focused on some of the larger ethical issues related to AI and can be seen as a complement to the FATML proposal, which is focused more on fairness, transparency and accountability. We would be happy to link or cluster the workshops together, but we (us and the FATML organizers) think that there is more than 2 day worth of material that the community needs to discuss in the area of AI and ethics, so it would be great to have both workshops if possible.

Schedule

08:20 AM	**Welcome and organisers comments**	*Bakalar, Lattimore, Bird, Mullainathan*
08:30 AM	**Jon Kleinberg - Fairness, Simplicity, and Ranking**	*Kleinberg*
09:00 AM	**Rich Caruna - Justice May Be Blind But It Shouldn’t Be Opaque: The Risk of Using Black-Box Models in Healthcare & Criminal Justice**	*Caruana*
09:30 AM	**Hoda Heidari - What Can Fair ML Learn from Economic Theories of Distributive Justice?**	
10:00 AM	**Poster Spotlights 1**	
10:20 AM	**Posters 1**	*Wei, Calmon, Dick, Gilpin, Lévesque, Ben Salem, Wang, Fitzsimons, Semenovich, Gu, Fruchter*
11:30 AM	**BriarPatches: Pixel-Space Interventions for Inducing Demographic Parity**	
11:50 AM	**Temporal Aspects of Individual Fairness**	

12:10 PM	**Explaining Explanations to Society**	
12:30 PM	**Lunch**	
02:00 PM	**Hanna Wallach - Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?**	*Wallach*
02:30 PM	**Rolle Dobbe - Ethics & Accountability in AI and Algorithmic Decision Making Systems - There's No Such Thing As A Free Lunch**	
03:00 PM	**Poster Spotlights 2**	
03:20 PM	**Posters 2**	
04:30 PM	**Manuel Gomez Rodriguez - Enhancing the Accuracy and Fairness of Human Decision Making**	*Rodriguez*
05:00 PM	**Discussion Panel**	

Abstracts (9):

Abstract 2: **Jon Kleinberg - Fairness, Simplicity, and Ranking in Workshop on Ethical, Social and Governance Issues in AI**, *Kleinberg* 08:30 AM

Recent discussion in the public sphere about classification by algorithms has involved tension between competing notions of what it means for such a classification to be fair to different groups. We consider several of the key fairness conditions that lie at the heart of these debates. In particular, we study how these properties operate when the goal is to rank-order a set of applicants by some criterion of interest, and then to select the top-ranking applicants. Among other results, we show that imposing a constraint to favor "simple" rules -- for example, to promote interpretability -- can have consequences for the equity of the ranking toward disadvantaged groups.

Abstract 3: **Rich Caruna - Justice May Be Blind But It Shouldn’t Be Opaque: The Risk of Using Black-Box Models in Healthcare & Criminal Justice in Workshop on Ethical, Social and Governance Issues in AI**, *Caruana* 09:00 AM

In machine learning often a tradeoff must be made between accuracy and intelligibility. This tradeoff sometimes limits the accuracy of models that can be safely deployed in mission-critical applications such as healthcare and criminal justice where being able to understand, validate, edit, and ultimately trust a learned model is important. In this talk I’ll present a case study where intelligibility is critical to uncover surprising patterns in the data that would have made deploying a black-box model dangerous. I’ll also show how distillation with intelligible models can be used to detect bias inside black-box models.

Abstract 4: **Hoda Heidari - What Can Fair ML Learn from Economic Theories of Distributive Justice? in Workshop on Ethical, Social and Governance Issues in AI**, 09:30 AM

Recently, a number of technical solutions have been proposed for tackling algorithmic unfairness and discrimination. I will talk about some of the connections between these proposals and to the long-established economic theories of fairness and distributive justice. In particular, I will overview the axiomatic characterization of measures of (income) inequality, and present them as a unifying framework for quantifying individual- and group-level unfairness; I will propose the use of cardinal social welfare functions as an an effective method for bounding individual-level inequality; and last but not least, I will cast existing notions of algorithmic (un)fairness as special cases of economic models of equality of opportunity---through this lens, I hope to offer a better understanding of the moral assumptions underlying technical definitions of fairness.

Abstract 7: **BriarPatches: Pixel-Space Interventions for Inducing Demographic Parity in Workshop on Ethical, Social and Governance Issues in AI**, 11:30 AM

We introduce the BriarPatch, a pixel-space intervention that obscures sensitive attributes from representations encoded in pre-trained classifiers. The patches encourage internal model representations not to encode sensitive information, which has the effect of pushing downstream predictors towards exhibiting demographic parity with respect to the sensitive information. The net result is that these BriarPatches provide an intervention mechanism available at user level, and complements prior research on fair representations that were previously only applicable by model developers and ML experts.

Abstract 8: **Temporal Aspects of Individual Fairness in Workshop on Ethical, Social and Governance Issues in AI**, 11:50 AM

The concept of individual fairness advocates similar treatment of similar individuals to ensure equality in treatment \cite{Dwork2012}. In this paper, we extend this notion to account for the time at which a decision is made, in settings where there exists a notion of "conduciveness" of decisions as perceived by individuals. We introduce two definitions: (i) fairness-across-time and (ii) fairness-in-hindsight. In the former, treatments of individuals are required to be individually fair relative to the past as well as future, while in the latter we only require individual fairness relative to the past. We show that these two definitions can have drastically different implications in the setting where the principal needs to learn the utility model: one can achieve a vanishing asymptotic loss in long-run average utility relative to the full-information optimum under the fairness-in-hindsight constraint, whereas this asymptotic loss can be bounded away from zero under the fairness-across-time constraint.

Abstract 9: **Explaining Explanations to Society in Workshop on Ethical, Social and Governance Issues in AI**, 12:10 PM

There is a disconnect between explanatory artificial intelligence (XAI) methods for
deep neural networks and the types of explanations that are useful for and demanded
by society (policy makers, government officials, etc.) Questions that experts in
artificial intelligence (AI) ask opaque systems provide inside explanations, focused
on debugging, reliability, and validation. These are different from those that society
will ask of these systems to build trust and confidence in their decisions.

Although
explanatory AI systems can answer many questions that experts desire, they often
don’t explain why they made decisions in a way that is precise (true to the model)
and understandable to humans. These outside explanations can be used to build
trust, comply with regulatory and policy changes, and act as external validation. In
this paper, we explore the types of questions that explanatory deep neural network
(DNN) systems can answer and discuss challenges inherent in building explanatory
systems that provide outside explanations of systems for societal requirements and
benefit.

Abstract 11: **Hanna Wallach - Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need? in Workshop on Ethical, Social and Governance Issues in AI**, *Wallach* 02:00 PM

The potential for machine learning systems to amplify social inequities and unfairness is receiving increasing popular and academic attention. A surge of recent research has focused on the development of algorithmic tools to detect and mitigate such unfairness. However, if these tools are to have a positive impact on industry practice, it is crucial that their design be informed by an understanding of industry teams’ actual needs. Through semi-structured interviews with 35 machine learning practitioners, spanning 19 teams and 10 companies,
and an anonymous survey of 267 practitioners, we conducted the first systematic investigation of industry teams' challenges and needs for support in developing fairer machine learning systems. I will describe this work and summarize areas of alignment and disconnect between the challenges faced by industry practitioners and solutions proposed in the academic literature. Based on these findings, I will highlight directions for future research that will better address practitioners' needs.

Abstract 12: **Rolle Dobbe - Ethics & Accountability in AI and Algorithmic Decision Making Systems - There's No Such Thing As A Free Lunch in Workshop on Ethical, Social and Governance Issues in AI**, 02:30 PM

Addressing a rapidly growing public awareness about bias and fairness issues in algorithmic decision-making systems (ADS), the tech industry is now championing a set of tools to assess and mitigate these. Such tools, broadly categorized as algorithmic fairness definitions, metrics and mitigation strategies find their roots in recent research from the community on Fairness, Accountability and Transparency in Machine Learning (FAT/ML), which started convening in 2014 at popular machine learning conferences, and has since been succeeded by a broader conference on Fairness, Accountability and Transparency in Sociotechnical Systems (FAT*). Whereas there is value in this research to assist diagnosis and informed debate about the inherent trade-offs and ethical choices that come with data-driven approaches to policy and decision-making, marketing poorly validated tools as quick fix strategies to eliminate bias is problematic and threatens to deepen an already growing sense of distrust among companies and institutions procuring data analysis software and enterprise platforms. This trend is coinciding with efforts by the IEEE and others to develop certification and marking processes that "advance transparency, accountability and reduction in algorithmic bias in Autonomous and Intelligent Systems". These efforts combined suggest a checkbox recipe for improving accountability and resolving the many ethical issues that have surfaced in the rapid deployment of ADS. In this talk, we nuance this timely debate by pointing at the inherent technical limitations of fairness metrics as a go-to tool for fixing bias. We discuss earlier attempts of certification to clarify pitfalls. We refer to developments in governments adopting ADS systems and how a lack of accountability and existing power structures are leading to new forms of harm that question the very efficacy of ADS. We end with discussing productive uses of diagnostic tools and the concept of Algorithmic Impact Assessment as a new framework for identifying the value, limitations and challenges of integrating algorithms in real world contexts.

Abstract 15: **Manuel Gomez Rodriguez - Enhancing the Accuracy and Fairness of Human Decision Making in Workshop on Ethical, Social and Governance Issues in AI**, *Rodriguez* 04:30 PM

Societies often rely on human experts to take a wide variety of decisions affecting their members, from jail-or-release decisions taken by judges and stop-and-frisk decisions taken by police officers to accept-or-reject decisions taken by academics. In this context, each decision is taken by an expert who is typically chosen uniformly at random from a pool of experts. However, these decisions may be imperfect due to limited experience, implicit biases, or faulty probabilistic reasoning. Can we improve the accuracy and fairness of the overall decision making process by optimizing the assignment between experts and decisions?

In this talk, we address the above problem from the perspective of sequential decision making and show that, for different fairness notions from the literature, it reduces to a sequence of (constrained) weighted bipartite matchings, which can be solved efficiently using algorithms with approximation guarantees. Moreover, these algorithms also benefit from posterior sampling to actively trade off exploitation---selecting expert assignments which lead to accurate and fair decisions---and exploration---selecting expert assignments to learn about the experts' preferences and biases. We demonstrate the effectiveness of our algorithms on both synthetic and real-world data and show that they can significantly improve both the accuracy and fairness of the decisions taken by pools of experts.

Imitation Learning and its Challenges in Robotics

Mustafa Mukadam, Sanjiban Choudhury, Siddhartha Srinivasa

Room 516 CDE, Fri Dec 07, 08:00 AM

Many animals including humans have the ability to acquire skills, knowledge, and social cues from a very young age. This ability to imitate by learning from demonstrations has inspired research across many disciplines like anthropology, neuroscience, psychology, and artificial intelligence. In AI, imitation learning (IL) serves as an essential tool for learning skills that are difficult to program by hand. The applicability of IL to robotics in particular, is useful when learning by trial and error (reinforcement learning) can be hazardous in the real world. Despite the many recent breakthroughs in IL, in the context of robotics there are several challenges to be addressed if robots are to operate freely and interact with humans in the real world.

Some important challenges include: 1) achieving good generalization and sample efficiency when the user can only provide a limited number of demonstrations with little to no feedback; 2) learning safe behaviors in human environments that require the least user intervention in terms of safety overrides without being overly conservative; and 3) leveraging data from multiple sources, including non-human sources, since limitations in hardware interfaces can often lead to poor quality demonstrations.

In this workshop, we aim to bring together researchers and experts in robotics, imitation and reinforcement learning, deep learning, and human robot interaction to
- Formalize the representations and primary challenges in IL as they pertain to robotics
- Delineate the key strengths and limitations of existing approaches with respect to these challenges
- Establish common baselines, metrics, and benchmarks, and identify open questions

Schedule

08:50 AM	**Introduction**	*Mukadam, Choudhury, Srinivasa*
09:00 AM	**Peter Stone**	*Stone*
09:30 AM	**Sonia Chernova**	*Chernova*
10:00 AM	**Contributed Spotlights**	
10:15 AM	**Coffee Break and Poster Session I**	*de Haan, Wang, Wang, Hayat, Sobh, Rana, Buhet, Rhinehart, Sharma, Bewley, Kelly, Blondé, Oguz, Viswanathan, Vanbaar, ■o■na, Rostamzadeh, McAllister, Thakur, Kalousis, Sidrane, Paul, Chen, Garmulewicz, Michalewski, Devin, Ren, Song, Sun, Hu, Liu, Wirbel*
11:00 AM	**Ingmar Posner**	*Posner*
11:30 AM	**Dorsa Sadigh**	*Sadigh*
12:00 PM	**Lunch Break**	
02:00 PM	**Byron Boots**	*Boots*
02:30 PM	**Dileep George**	*George*
02:45 PM	**Coffee Break and Poster Session II**	
03:30 PM	**Yisong Yue**	*Yue*
04:00 PM	**Anca Dragan**	*Dragan*
04:30 PM	**Drew Bagnell / Wen Sun**	*Bagnell, Sun*
05:00 PM	**Panel Discussion**	

Continual Learning

Razvan Pascanu, Yee Teh, Marc Pickett, Mark Ring

Room 517 A, Fri Dec 07, 08:00 AM

Continual learning (CL) is the ability of a model to learn continually from a stream of data, building on what was learnt previously, hence exhibiting positive transfer, as well as being able to remember previously seen tasks. CL is a fundamental step towards artificial intelligence, as it allows the agent to adapt to a continuously changing environment, a hallmark of natural intelligence. It also has implications for supervised or unsupervised learning. For example, when the dataset is not properly shuffled or there exists a drift in the input distribution, the model overfits the recently seen data, forgetting the rest -- phenomena referred to as catastrophic forgetting, which is part of CL and is something CL systems aim to address.

Continual learning is defined in practice through a series of desiderata. A non-complete lists includes:
* Online learning -- learning occurs at every moment, with no fixed tasks or data sets and no clear boundaries between tasks;
* Presence of transfer (forward/backward) -- the model should be able to transfer from previously seen data or tasks to new ones, as well as possibly new task should help improve performance on older ones;
* Resistance to catastrophic forgetting -- new learning does not destroy performance on previously seen data;
* Bounded system size -- the model capacity should be fixed, forcing the system use its capacity intelligently as well as gracefully forgetting information such to ensure maximising future reward;
* No direct access to previous experience -- while the model can remember a limited amount of experience, a continual learning algorithm should not have direct access to past tasks or be able to rewind the environment;

In the previous edition of the workshop the focus has been on defining a complete list of desiderata, of what a continual learning (CL) enabled system should be able to do. We believe that in this edition we should further constrain the discussion with a focus on how to evaluate CL and how it relates to other existing topics (e.g. life-long learning, transfer learning, meta-learning) and how ideas from these topics could be useful for continual learning.

Different aspects of continual learning are in opposition of each other (e.g. fixed model capacity and not-forgetting), which also raises the question of how to evaluate continual learning systems. One one hand, what are the right trade-offs between these different opposing forces? How do we compare existing algorithms given these different dimensions along which we should evaluate them (e.g. forgetting, positive transfer)? What are the right metrics we should report? On the other hand, optimal or meaningful trade-offs will be tightly defined by the data or at least type of tasks we use to test the algorithms. One prevalent task used by many recent papers is PermutedMNIST. But as MNIST is not a reliable dataset for classification, so PermutedMNIST might be extremely misleading for continual learning. What would be the right benchmarks, datasets or tasks for fruitfully exploiting this topic?

Finally, we will also encourage presentation of both novel approaches to CL and implemented systems, which will help concretize the discussion of what CL is and how to evaluate CL systems.

Schedule

08:30 AM	**Introduction of the workshop**	*Pascanu, Teh, Ring, Pickett*
09:15 AM	**Spotlight #1**	
09:30 AM	**Spotlight #2**	
09:45 AM	**Spotlight #3**	
10:00 AM	**Invited Speaker #1 Chelsea Finn**	*Finn*
11:00 AM	**Invited Speaker #2 Raia Hadsell**	*Hadsell*
11:30 AM	**Invited Speaker #3 Marc'Aurelio Ranzato**	*Ranzato*
12:00 PM	**Lunch & Posters**	*Fayek, Parisi, Xu, Mudrakarta, Cerf, Wassermann, Soselia, Aljundi, Elhoseiny, Lavda, Liang, Chaudhry, Narvekar, Lomonaco, Chung, Chang, Zhao, Kira, Bashivan, Rafiee, Ostapenko, Jones, Kaplanis, Kalkan, Teng, He, Liu, Nath, Ahn, Chen, Huang, Chandak, Sprague, Schrimpf, Kendall, Schwarz, Li, Du, Hsu, Abnar, Wang*
02:00 PM	**Invited Speaker #4 Juergen Schmidhuber**	*Schmidhuber*
02:30 PM	**Invited Speaker #5 Yarin Gal**	*Gal*
03:00 PM	**Coffee Break & Posters**	
03:30 PM	**Spotlight #4**	
03:45 PM	**Spotlight #5**	
04:00 PM	**Spotlight #6**	
04:15 PM	**Overview of Darpa's Lifelong learning program (Hava Siegelmann)**	*Siegelmann*
04:30 PM	**Invited Speaker #6 Martha White**	*White*
05:00 PM	**Panel Discussion**	

Abstracts (6):

Abstract 1: **Introduction of the workshop in Continual Learning**, *Pascanu, Teh, Ring, Pickett* 08:30 AM

Introduction of the Continual Learning workshop from the organizers, expressing their opinion of the goal of the workshop.

Abstract 2: **Spotlight #1 in Continual Learning**, 09:15 AM

TBD

Abstract 3: **Spotlight #2 in Continual Learning**, 09:30 AM

TBD

Abstract 4: **Spotlight #3 in Continual Learning**, 09:45 AM

TBD

Abstract 12: **Spotlight #4 in Continual Learning**, 03:30 PM

TBD

Abstract 14: **Spotlight #6 in Continual Learning**, 04:00 PM

TBD

NIPS 2018 workshop on Compact Deep Neural Networks with industrial applications

Lixin Fan, Zhouchen Lin, Max Welling, Yurong Chen, Werner Bailer

Room 517 B, Fri Dec 07, 08:00 AM

This workshop aims to bring together researchers, educators, practitioners who are interested in techniques as well as applications of making compact and efficient neural network representations. One main theme of the workshop discussion is to build up consensus in this rapidly developed field, and in particular, to establish close connection between researchers in Machine Learning community and engineers in industry. We believe the workshop is beneficial to both academic researchers as well as industrial practitioners.

===
News and announcements:

. For authors of spotlight posters, please send your one-minute slides (preferably with recorded narrative) to lixin.fan01@gmail.com, or copy it to a UPS stick. See you at the workshop then.

. Please note the change of workshop schedule. Due to visa issues, some speakers are unfortunately unable to attend the workshop.

. There are some reserve NIPS/NeurIPS tickets available now, on a first come first serve basis, for co-authors of workshop accepted papers! Please create NIPS acocunts, and inform us the email addresses if reserve tickets are needed.

. For authors included in the spot light session, please prepare short slides with presentation time stictly within 1 minute. It is preferably to record your presentation with audio & video (as instructed e.g. at https://support.office.com/en-us/article/record-a-slide-show-with-narration-and-slide-tim

. For authors included in the spot light session, please also prepare a poster for your paper, and make sure either yourself or your co-authors will present the poster after the coffee break.

. Please make your poster 36W x 48H inches or 90 x 122 cm. Make sure your poster is in *portrait* orientation and does not exceed the maximal size, since we have limited space for the poster session.

===
For authors of following accepted papers, please revise your submission as per reviewers comments to address raised issues. If there are too much contents to be included in 3 page limit, you may use appendix for supporting contents such as proofs or detailed experimental results. The

camera ready abstract should be prepared with authors information (name, email address, affiliation) using the NIPS camera ready template.

Please submit the camera ready abstract through OpenReview (https://openreview.net/group?id=NIPS.cc/2018/Workshop/CDNNRIA) by Nov. 12th. Use your previous submission page to update the abstract. In case you have to postpone the submission, please inform us immeidately. Otherwise, the abstract will be removed from the workshop schedule.

===
We invite you to submit original work in, but not limited to, following areas:

Neural network compression techniques:
. Binarization, quantization, pruning, thresholding and coding of neural networks
. Efficient computation and acceleration of deep convolutional neural networks
. Deep neural network computation in low power consumption applications (e.g., mobile or IoT devices)
. Differentiable sparsification and quantization of deep neural networks
. Benchmarking of deep neural network compression techniques

Neural network representation and exchange:
. Exchange formats for (trained) neural networks
. Efficient deployment strategies for neural networks
. Industrial standardization of deep neural network representations
. Performance evaluation methods of compressed networks in application context (e.g., multimedia encoding and processing)

Video & media compression methods using DNNs such as those developed in MPEG group:
. To improve video coding standard development by using deep neural networks
. To increase practical applicability of network compression methods

An extended abstract (3 pages long using NIPS style, see https://nips.cc/Conferences/2018/PaperInformation/StyleFiles) in PDF format should be submitted for evaluation of the originality and quality of the work. The evaluation is double-blind and the abstract must be anonymous. References may extend beyond the 3 page limit, and parallel submissions to a journal or conferences (e.g. AAAI or ICLR) are permitted.

Submissions will be accepted as contributed talks (oral) or poster presentations. Extended abstract should be submitted through OpenReview (https://openreview.net/group?id=NIPS.cc/2018/Workshop/CDNNRIA) by 20 Oct 2018. All accepted abstracts will be posted on the workshop website and archived.

Selection policy: all submitted abstracts will be evaluted based on their novelty, soundness and impacts. At the workshop we encourage DISCUSSION about NEW IDEAS, each submitter is thus expected to actively respond on OpenReview webpage and answer any questions about his/her ideas. The willingness to respond in OpenReview Q/A disucssions will be an important factor for the selection of accepted oral or poster presentations.

Important dates:
. Extended abstract submission deadline: 20 Oct 2018,
. Acceptance notification: 29 Oct. 2018,
. Camera ready submission: 12 November 2018,
. Workshop: 7 December 2018

Submission:
Please submit ■your extended abstract through OpenReivew system (https://openreview.net/group?id=NIPS.cc/2018/Workshop/CDNNRIA).
For prospective authors: please send author information to workshop chairs (lixin.fan@nokia.com), so that you submission can be assigned to reviewers without conflict of interests.
. Reviewers comments will be released by Oct. 24th, then authors have to reply by Oct. 27th, which leaving us two days for decision-making.
. It is highly recommended for authors submit abstracts early, in case you need more time to address reviewers' comments.

NIPS Complimentary workshop registration
We will help authors of accepted submissions to get access to a reserve pool of NIPS tickets. So please register to the workshop early.

===
Accepted papers & authors:

1. Minimal Random Code Learning: Getting Bits Back from Compressed Model Parameters,
Marton Havasi, Robert Peharz, José Miguel Hernández-Lobato

2. Neural Network Compression using Transform Coding and Clustering,
Thorsten Laude, Jörn Ostermann

3. Pruning neural networks: is it time to nip it in the bud?,
Elliot J. Crowley, Jack Turner, Amos Storkey, Michael O'Boyle

4. Compressing Recurrent Neural Networks with Tensor Ring for Action Recognition,
Yu Pan, Jing Xu, Maolin Wang, Fei Wang, Kun Bai, Zenglin Xu

5. Efficient Inference on Deep Neural Networks by Dynamic Representations and Decision Gates,
Mohammad Saeed Shafiee, Mohammad Javad Shafiee, Alexander Wong

6. Iteratively Training Look-Up Tables for Network Quantization,
Fabien Cardinaux, Stefan Uhlich, Kazuki Yoshiyama, Javier Alonso García, Stephen Tiedemann, Thomas Kemp, Akira Nakamura

7. Hybrid Pruning: Thinner Sparse Networks for Fast Inference on Edge Devices,
Xiaofan Xu, Mi Sun Park, Cormac Brick

8. Compression of Acoustic Event Detection Models with Low-rank Matrix Factorization and Quantization Training,
Bowen Shi, Ming Sun, Chieh-Chi Kao, Viktor Rozgic, Spyros Matsoukas, Chao Wang

9. On Learning Wire-Length Efficient Neural Networks,
Christopher Blake, Luyu Wang, Giuseppe Castiglione, Christopher Srinavasa, Marcus Brubaker

10. FLOPs as a Direct Optimization Objective for Learning Sparse

Neural Networks,
Raphael Tang, Ashutosh Adhikari, Jimmy Lin

11. Three Dimensional Convolutional Neural Network Pruning with Regularization-Based Method,
Yuxin Zhang, Huan Wang, Yang Luo, Roland Hu

12. Differentiable Training for Hardware Efficient LightNNs,
Ruizhou Ding, Zeye Liu, Ting-Wu Chin, Diana Marculescu, R.D. (Shawn) Blanton

13. Structured Pruning for Efficient ConvNets via Incremental Regularization,
Huan Wang, Qiming Zhang, Yuehai Wang, Haoji Hu

14. Block-wise Intermediate Representation Training for Model Compression,
Animesh Koratana*, Daniel Kang*, Peter Bailis, Matei Zahaira

15. Targeted Dropout,
Aidan N. Gomez, Ivan Zhang, Kevin Swersky, Yarin Gal, Geoffrey E. Hinton

16. Adaptive Mixture of Low-Rank Factorizations for Compact Neural Modeling,
Ting Chen, Ji Lin, Tian Lin, Song Han, Chong Wang, Denny Zhou

17. Differentiable Fine-grained Quantization for Deep Neural Network Compression,
Hsin-Pai Cheng, Yuanjun Huang, Xuyang Guo, Yifei Huang, Feng Yan, Hai Li, Yiran Chen

18. Transformer to CNN: Label-scarce distillation for efficient text classification,
Yew Ken Chia, Sam Witteveen, Martin Andrews

19. EnergyNet: Energy-Efficient Dynamic Inference,
Yue Wang, Tan Nguyen, Yang Zhao, Zhangyang Wang, Yingyan Lin, Richard Baraniuk

20. Recurrent Convolutions: A Model Compression Point of View,
Zhendong Zhang, Cheolkon Jung

21, Rethinking the Value of Network Pruning,
Zhuang Liu, Mingjie Sun, Tinghui Zhou, Gao Huang, Trevor Darrell

22. Linear Backprop in non-linear networks,
Mehrdad Yazdani

23. Bayesian Sparsification of Gated Recurrent Neural Networks,
Ekaterina Lobacheva, Nadezhda Chirkova, Dmitry Vetrov

24. Demystifying Neural Network Filter Pruning,
Zhuwei Qin, Fuxun Yu, Chenchen Liu, Xiang Chen

25. Learning Compact Networks via Adaptive Network Regularization,
Sivaramakrishnan Sankarapandian, Anil Kag, Rachel Manzelli, Brian Kulis

26. Pruning at a Glance: A Structured Class-Blind Pruning Technique for Model Compression
Abdullah Salama, Oleksiy Ostapenko, Moin Nabi, Tassilo Klein

27. Succinct Source Coding of Deep Neural Networks
Sourya Basu, Lav R. Varshney

28. Fast On-the-fly Retraining-free Sparsification of Convolutional Neural Networks
Amir H. Ashouri, Tarek Abdelrahman, Alwyn Dos Remedios

29. PocketFlow: An Automated Framework for Compressing and Accelerating Deep Neural Networks
Jiaxiang Wu, Yao Zhang, Haoli Bai, Huasong Zhong, Jinlong Hou, Wei Liu, Junzhou Huang

30. Universal Deep Neural Network Compression
Yoojin Choi, Mostafa El-Khamy, Jungwon Lee

31. Compact and Computationally Efficient Representations of Deep Neural Networks
Simon Wiedemann, Klaus-Robert Mueller, Wojciech Samek

32. Dynamic parameter reallocation improves trainability of deep convolutional networks
Hesham Mostafa, Xin Wang

33. Compact Neural Network Solutions to Laplace's Equation in a Nanofluidic Device
Martin Magill, Faisal Z. Qureshi, Hendrick W. de Haan

34. Distilling Critical Paths in Convolutional Neural Networks
Fuxun Yu, Zhuwei Qin, Xiang Chen

35. SeCSeq: Semantic Coding for Sequence-to-Sequence based Extreme Multi-label Classification
Wei-Cheng Chang, Hsiang-Fu Yu, Inderjit S. Dhillon, Yiming Yang

===
A best paper award will be presented to the contribution selected by reviewers, who will also take into account active disucssions on OpenReview. One FREE NIPS ticket will be awarded to the best paper presenter.

The best paper award is given to the authors of "Rethinking the Value of Network Pruning",
Zhuang Liu, Mingjie Sun, Tinghui Zhou, Gao Huang, Trevor Darrell

=====
Acknowledgement to reviewers

The workshop organizers gratefully acknowledge the assistance of the following people, who reviewed submissions and actively disucssed with authors:

Zhuang Liu, Ting-Wu Chin, Fuxun Yu, Huan Wang, Mehrdad Yazdani, Qigong Sun, Tim Genewein, Abdullah Salama, Anbang Yao, Chen Xu, Hao Li, Jiaxiang Wu, Zhisheng Zhong, Haoji Hu, Hesham Mostafa, Seunghyeon Kim, Xin Wang, Yiwen Guo, Yu Pan, Fereshteh Lagzi, Martin Magill, Wei-Cheng Chang, Yue Wang, Caglar Aytekin, Hannes Fassold, Martin Winter, Yunhe Wang, Faisal Qureshi, Filip Korzeniowski, jianguo Li, Jiashi Feng, Mingjie Sun, Shiqi Wang, Tinghuai Wang, Xiangyu Zhang, Yibo Yang, Ziqian Chen, Francesco Cricri, Jan Schlüter,

Jing Xu, Lingyu Duan, Maoin Wang, Naiyan Wang, Stephen Tyree, Tianshui Chen, Vasileios Mezaris, Christopher Blake, Chris Srinivasa, Giuseppe Castiglione, Amir Khoshamam, Kevin Luk, Luyu Wang, Jian Cheng, Pavlo Molchanov, Yihui He, Sam Witteveen, Peng Wang,

with special thanks to Ting-Wu Chin who contributed 7 reviewer comments.

=====

Workshop meeting room: 517B
Workshop schedule on December 7th, 2018:

Schedule

09:00 AM	**Opening and Introduction**	
09:05 AM	**Rethinking the Value of Network Pruning**	*Liu*
09:30 AM	**Bandwidth efficient deep learning by model compression**	*Han*
09:55 AM	**Neural network compression in the wild: why aiming for high compression factors is not enough**	*Genewein*
10:20 AM	**Linear Backprop in non-linear networks**	*Yazdani*
10:45 AM	**Coffee break (morning)**	
11:00 AM	**Challenges and lessons learned in DNN portability in production**	*Lee*
11:30 AM	**Bayesian Sparsification of Gated Recurrent Neural Networks**	*Chirkova*
12:00 PM	**Lunch break (on your own)**	
02:00 PM	**Efficient Computation of Deep Convolutional Neural Networks: A Quantization Perspective**	*Welling*
02:25 PM	**Deep neural network compression and acceleration**	*Yao*
02:50 PM	**Poster spotlight session.**	*Salama, Chang, Gomez, Tang, YU, Zhang, Zhang, Lin, Tiedemann, Bai, Sankarapandian, Havasi, Turner, Cheng, Wang, Xu, Ding, Hu, Shafiee, Blake, Kao, Kang, Chia, Ashouri, Basu, Wiedemann, Laude*
03:20 PM	**Coffee break (afternoon)**	
03:30 PM	**Poster presentations**	*Wiedemann, Wang, Zhang, Wang, Shafiee, Manzelli, Huang, Klein, Zhang, Adhikari, Qureshi, Castiglione*
04:45 PM	**Panel disucssion**	*Welling, Genewein, Park, Han*
05:45 PM	**Closing**	

Abstracts (5):

Abstract 3: **Bandwidth efficient deep learning by model compression in NIPS 2018 workshop on Compact Deep Neural Networks with industrial applications**, *Han* 09:30 AM

n the post-ImageNet era, computer vision and machine learning researchers are solving more complicated AI problems using larger datasets driving the demand for more computation. However, we are in the post-Moore's Law world where the amount of computation per unit cost and power is no longer increasing at its historic rate. This mismatch between supply and demand for computation highlights the need for co-designing efficient algorithms and hardware. In this talk, I will talk about bandwidth efficient deep learning by model compression, together with efficient hardware architecture support, saving memory bandwidth, networking bandwidth, and engineer bandwidth.

Abstract 4: **Neural network compression in the wild: why aiming for high compression factors is not enough in NIPS 2018 workshop on Compact Deep Neural Networks with industrial applications**, *Genewein* 09:55 AM

Abstract: the widespread use of state-of-the-art deep neural network models in the mobile, automotive and embedded domains is often hindered by the steep computational resources that are required for running such models. However, the recent scientific literature proposes a plethora of of ways to alleviate the problem, either on the level of efficient network architectures, efficiency-optimized hardware or via network compression methods. Unfortunately, the usefulness of a network compression method strongly depends on the other aspects (network architecture and target hardware) as well as the task itself (classification, regression, detection, etc.), but very few publications consider this interplay. This talk highlights some of the issues that arise from the strong interplay between network architecture, target hardware, compression algorithm and target task. Additionally some shortcomings in the current literature on network compression methods are pointed-out, such as incomparability of results (different base-line networks, different training-/data-augmentation schemes, etc.), lack of results on tasks other than classification, or use of very different (and perhaps not very informative) quantitative performance indicators such as naive compression rate, operations-per-second, size of stored weight matrices, etc. The talk concludes by proposing some guidelines and best-practices for increasing practical applicability of network compression methods and a call for standardizing network compression benchmarks.

Abstract 7: **Challenges and lessons learned in DNN portability in production in NIPS 2018 workshop on Compact Deep Neural Networks with industrial applications**, *Lee* 11:00 AM

Deploying state-of-the-art deep neural networks into high-performance production system comes with many challenges. There is a plethora of

deep learning frameworks with different operator designs and model format. As a deployment platform developer, having a portable model format to parse, instead of developing parsers for every single framework seems very attractive.

As a pioneer in the deep learning inference platform, NVIDIA TensorRT introduced UFF as a proposed solution last year, and now there are more exchange format available such as ONNX and NNEF.

In this talk, we will share the lessons learned from the TensorRT use cases in various production environment working with a portable format, with consideration of optimizations such as pruning, quantization, and auto-tuning on different target accelerators. We will also discuss some of the open challenges.

Abstract 10: **Efficient Computation of Deep Convolutional Neural Networks: A Quantization Perspective in NIPS 2018 workshop on Compact Deep Neural Networks with industrial applications**, *Welling* 02:00 PM

Abstract: neural network compression has become an important research area due to its great impact on deployment of large models on resource constrained devices. In this talk, we will introduce two novel techniques that allow for differentiable sparsification and quantization of deep neural networks; both of these are achieved via appropriate smoothing of the overall objective. As a result, we can directly train architectures to be highly compressed and hardware-friendly via off-the-self stochastic gradient descent optimizers.

Abstract 11: **Deep neural network compression and acceleration in NIPS 2018 workshop on Compact Deep Neural Networks with industrial applications**, *Yao* 02:25 PM

In the past several years, Deep Neural Networks (DNNs) have demonstrated record-breaking accuracy on a variety of artificial intelligence tasks. However, the intensive storage and computational costs of DNN models make it difficult to deploy them on the mobile and embedded systems for real-time applications. In this technical talk, Dr. Yao will introduce their recent works on deep neural network compression and acceleration, showing how they achieve impressive compression performance without noticeable loss of model prediction accuracy, from the perspective of pruning and quantization.

Modeling the Physical World: Learning, Perception, and Control

Jiajun Wu, Kelsey Allen, Kevin Smith, Jessica Hamrick, Emmanuel Dupoux, Marc Toussaint, Josh Tenenbaum

Room 517 C, Fri Dec 07, 08:00 AM

Despite recent progress, AI is still far from achieving common-sense scene understanding and reasoning. A core component of this common sense is a useful representation of the physical world and its dynamics that can be used to predict and plan based on how objects interact. This capability is universal in adults, and is found to a certain extent even in infants. Yet despite increasing interest in the phenomenon in recent years, there are currently no models that exhibit the robustness and flexibility of human physical reasoning.

There have been many ways of conceptualizing models of physics, each with their complementary strengths and weaknesses. For instance, traditional physical simulation engines have typically used symbolic or analytic systems with “built-in” knowledge of physics, while recent connectionist methods have demonstrated the capability to learn approximate, differentiable system dynamics. While more precise, symbolic models of physics might be useful for long-term prediction and physical inference; approximate, differentiable models might be more practical for inverse dynamics and system identification. The design of a physical dynamics model fundamentally affects the ways in which that model can, and should, be used.

This workshop will bring together researchers in machine learning, computer vision, robotics, computational neuroscience, and cognitive psychology to discuss artificial systems that capture or model the physical world. It will also explore the cognitive foundations of physical representations, their interaction with perception, and their applications in planning and control. There will be invited talks from world leaders in the fields, presentations and poster sessions based on contributed papers, and a panel discussion.

Topics of discussion will include
- Building and learning physical models (deep networks, structured probabilistic generative models, physics engines)
- How to combine model-based and model-free approaches to physical prediction
- How to use physics models in higher-level tasks such as navigation, video prediction, robotics, etc.
- How perception and action interact with physical representations
- How cognitive science and computational neuroscience may inform the design of artificial systems for physical prediction
- Methodology for comparing models of infant learning with artificial systems
- Development of new datasets or platforms for physics and visual common sense

Schedule

08:40 AM	**Opening Remarks: Josh Tenenbaum**	*Tenenbaum*
09:00 AM	**Talk 1: Zico Kolter - Differentiable Physics and Control**	*Kolter*
09:30 AM	**Talk 2: Emo Todorov - Physics-Based Control**	*Todorov*
10:00 AM	**Contributed Talk 1: ChainQueen: A Real-Time Differentiable Physical Simulator for Soft Robotics**	*Spielberg*
10:10 AM	**Contributed Talk 2: To Stir or Not to Stir: Online Estimation of Liquid Properties for Pouring Actions**	*López-Guevara*
10:20 AM	**Contributed Talk 3: Learning Robotic Manipulation through Visual Planning and Acting**	*Wang*

10:30 AM	**Coffee Break 1 (Posters)**	*Kumar, Huang, Xu, Janner, Chadha, Thuerey, Lu, Bauza, Tompkins, Shi, Baumeister, Ofner, Cheng, Luo, Bablani, Vanbaar, Subr, López-Guevara, Jha, Fuchs, Rosa, Pouplin, Ray, Liu, Crawford*
11:00 AM	**Talk 3: Jitendra Malik - Linking Perception and Action**	*Malik*
11:30 AM	**Talk 4: Chelsea Finn - An agent that can do many things (by modeling the world)**	*Finn*
12:00 PM	**Lunch Break**	
02:00 PM	**Talk 5: Peter Battaglia - Structure in Physical Intelligence**	*Battaglia*
02:30 PM	**Talk 6: Dan Yamins - The Objects of Our Curiosity: Intrinsic Motivation, Intuitive Physics and Self-Supervised Learning**	*Yamins*
03:00 PM	**Coffee Break 2 (Posters)**	
03:30 PM	**Talk 7: Jeannette Bohg - On perceptual representations and how they interact with actions and physical models**	*Bohg*
04:00 PM	**Talk 8: Leslie Kaelbling - Learning models of very large hybrid domains**	*Kaelbling*
04:30 PM	**Talk 9: Marc Toussaint - Models & Abstractions for Physical Reasoning**	*Toussaint*
05:00 PM	**Panel Discussion**	

All of Bayesian Nonparametrics (Especially the Useful Bits)

Diana Cai, Trevor Campbell, Mike Hughes, Tamara Broderick, Nick Foti, Sinead Williamson

Room 517 D, Fri Dec 07, 08:00 AM

Bayesian nonparametric (BNP) methods are well suited to the large data sets that arise in a wide variety of applied fields. By making use of infinite-dimensional mathematical structures, BNP methods allow the complexity of a learned model to grow as the size of a data set grows, exhibiting desirable Bayesian regularization properties for small data sets and allowing the practitioner to learn ever more from larger data sets. These properties have resulted in the adoption of BNP methods across a diverse set of application areas---including, but not limited to, biology, neuroscience, the humanities, social sciences, economics, and finance.

This workshop aims to highlight recent advances in modeling and computation through the lens of applied, domain-driven problems that require the infinite flexibility and interpretability of BNP. In this workshop, we will explore new BNP methods for diverse applied problems, including cutting-edge models being developed by application domain experts. We will also discuss the limitations of existing methods and discuss key problems that need to be solved. A major focus of the workshop will be to expose participants to practical software tools for performing Bayesian nonparametric analyses. In particular, we plan to host hands-on tutorials to introduce workshop participants to some of the software packages that can be used to easily perform posterior inference for BNP models. On the software panel, we will have researchers who have experience with BNP and development experience with popular software systems, such as TensorFlow, Edward, Stan, and Autograd.

We expect workshop participants to come from a variety of fields, including but not limited to machine learning, statistics, engineering, the social sciences, and biological sciences. The workshop will be relevant both to BNP experts as well as those interested in learning how to apply BNP models. There will be a special emphasis on novel application areas and computational developments that make BNP more accessible to the broader machine learning audience. Participants will leave the workshop with (i) exposure to recent advances in the field, (ii) hands-on experience with software implementing BNP methods, and (iii) an idea of the current major challenges in the field. These goals will be accomplished through a series of invited and contributed talks, a poster session, and at least one hands-on tutorial session where participants can get their hands dirty with BNP methods.

This workshop builds off of:
1. NIPS 2015: “Bayesian Nonparametrics: The Next Generation”: https://sites.google.com/site/nipsbnp2015/, and
2. NIPS 2016: “Practical Bayesian Nonparametrics”: https://sites.google.com/site/nipsbnp2016/,
which have spanned various areas of BNP, such as theory, applications and computation. This year’s workshop will have a fresh take on recent developments in BNP in connection to the broader range of research in statistics, machine learning, and application domains.

The 2018 workshop has received an endorsement from the International Society of Bayesian Analysis (ISBA) and sponsorship from Google.

Organizing Committee:
Diana Cai (Princeton)
Trevor Campbell (MIT/UBC)
Mike Hughes (Harvard/Tufts)
Tamara Broderick (MIT)
Nick Foti (U Washington)
Sinead Williamson (UT Austin)

Advisory Committee:
Emily Fox (U Washington)
Antonio Lijoi (Bocconi U)
Sonia Petrone (Bocconi U)
Igor Prünster (Bocconi U)
Erik Sudderth (UC Irvine)

Schedule

08:20 AM	**Opening remarks**	
08:30 AM	**Invited Talk 1**	
09:00 AM	**Invited Talk 2**	*Bloem-Reddy*
09:30 AM	**Contributed Talk 1**	
09:45 AM	**Poster Spotlights**	
11:00 AM	**Invited Talk 3**	*Chaney*
11:45 AM	**Poster Session**	*Masoero, Rukat, Liu, Ray Chowdhury, Coelho de Castro, Wehrhahn, Saad, Verma, Hsu, Cabreros, Prabhakaran, Sun, Rischard, Liu, Farooq, Liu, F. Pradier, Romeres, Campbell, Xu, dundar, Keuter, Gyawali, Sennesh, De Palma, Flam-Shepherd, Kubo*
12:00 PM	**TensorFlow Probability Tutorial**	
02:00 PM	**Invited Talk 4**	*Futoma*
02:30 PM	**Software Panel**	*Letham, Duvenaud, Tran, Vehtari*
03:30 PM	**Invited Talk 5**	*Balakrishnan*
04:00 PM	**Contributed Talk 3**	
04:30 PM	**Research Panel**	*Williamson, Engelhardt, Griffiths, Lawrence, Wallach*
05:30 PM	**Poster Session + Individual Discussion**	

NeurIPS 2018 Competition Track Day 1

Sergio Escalera, Ralf Herbrich

Room 518, Fri Dec 07, 08:00 AM

coming soon.

Schedule

07:50 AM	**Opening and Best Demo Award announcement**	*Escalera, Herbrich*
08:00 AM	**InclusiveImages: Competition Overview**	
08:20 AM	**InclusiveImages: Competitor Presentations**	*Halpern, Baljekar, Sculley, Ostyakov, Mohammed, Wang, Austin*
09:00 AM	**InclusiveImages: Wrap-Up**	
09:15 AM	**Adversarial Vision Challenge: Shooting ML Models in the Dark: The Landscape of Blackbox Attacks**	*Madry*
09:30 AM	**Adversarial Vision Challenge: Results of the Adversarial Vision Challenge**	*Brendel, Rauber, Salathé, Kurakin, Papernot, Mohanty, Bethge*
09:40 AM	**Adversarial Vision Challenge: Theory-inspired Approaches for Adversarial Machine Learning**	*xu, Yu*
09:50 AM	**Adversarial Vision Challenge: Improving attack transferability with adversarially trained surrogate models**	*Rony, Hafemann*
10:00 AM	**Adversarial Vision Challenge: Guessing Smart, New Directions for Sampling-Based Attacks**	*Brunner, Diehl*
10:10 AM	**Adversarial Vision Challenge: Black-box Attacks with Small Search Subspace and Covariance Matrix Adaptation of Perturbation Noise**	*Cheng, Dong*
10:20 AM	**Adversarial Vision Challenge: Towards More Effective Black-Box Adversarial Training**	*Ning, Li, Wang*
10:30 AM	**Adversarial Vision Challenge: Poster Session**	*Sharma, Holdijk, Saralajew, Yan, Rashchenko, Rashchenko, Jang, Lee, Yoon, KIM, Laurent, Schott*
10:30 AM	**Coffee break**	
11:00 AM	**The Conversational Intelligence Challenge 2 (ConvAI2) : Setup, Opening Words**	*Weston*
11:10 AM	**The Conversational Intelligence Challenge 2 (ConvAI2) : Winners, Results & Analysis**	*Dinan*
11:20 AM	**The Conversational Intelligence Challenge 2 (ConvAI2) : Winners talks & spotlights**	*Wolf, peng, Saam, Elder, Kurbanov, Alam*
12:05 PM	**The Conversational Intelligence Challenge 2 (ConvAI2) : Conclusion, Discussion of Future Steps**	

12:15 PM	**Lunch break**	
01:15 PM	**Xprize competition: introduction**	*Dadkhahnikoo*
01:20 PM	**Xprize competition: MachineGenes**	*Greenwood*
01:25 PM	**Xprize competition: DeepDrug**	*Mukhopadhyay*
01:30 PM	**AutoML3 - LifeLong ML with concept drift Challenge: Overview and award ceremony**	*Escalante, Guyon, Silver, Viegas, Tu*
01:50 PM	**AutoML3 - LifeLong ML with concept drift Challenge. First place winner: Automatically Optimized Gradient Boosting Trees for Classifying Large Volume High Cardinality Data Streams under Concept Drift**	*Wilson, Meher, Vinodkumar Bindu, Sharma, pareek*
02:10 PM	**AutoML3 - LifeLong ML with concept drift Challenge. Second place winner. A Boosting Tree Based AutoML System for High Cardinality Streaming Data Classification with Concept Drift**	*Xiong, Jiang, Zhang*
02:20 PM	**AutoML3 - LifeLong ML with concept drift Challenge. Thrid place winner: GrandMasters team at AutoML3 challenge**	*Chang, Zhao, Liu, chai*
02:40 PM	**AutoML3 - LifeLong ML with concept drift Challenge: MIT HAN LAB experience in the AutoML3 challenge**	*Chen*
02:40 PM	**AutoML3 - LifeLong ML with concept drift Challenge: Experiences from participants**	*Jin*
03:00 PM	**Coffee break**	
03:30 PM	**TrackML challenge - Introduction : tracking particles in LHC detector**	*Salzburger*
03:45 PM	**TrackML Analyzing some top solutions**	*vlimant*
04:00 PM	**TrackML a LSTM attempt**	*Finnie, Finnie*
04:15 PM	**TrackML when speed matter**	*Gorbunov*
04:30 PM	**TrackML Conclusion and Outlook on the on-going Throughput phase**	*Rousseau*
04:45 PM	**Break**	
05:15 PM	**AI for prosthetics: competition review**	*Kidzi■ski*
05:20 PM	**AI for prosthetics: crowdAI**	*Mohanty*
05:25 PM	**AI for prosthetics: top ranked teams talks**	*Watson, Huynh, Gamble, Wang, Bhatt, Rane, Shpilman, Wang, Ljungström, Kolesnikov, Hrinchuk, Ja■kowski, Shyam, Wang, Zeng, Zhou*
06:25 PM	**AI for prosthetics: OpenAI dexterous in-hand manipulation invited talk**	*Andrychowicz*

The second Conversational AI workshop – today's practice and tomorrow's potential

Alborz Geramifard, Jason Williams, Y-Lan Boureau, Maxine Eskenazi, Milica Gasic, Jim Glass, Dilek Hakkani-Tur, Larry Heck, Lazaros Polymenakos, Steve Young

Room 519, Fri Dec 07, 08:00 AM

In the span of only a few years, conversational systems have become commonplace. Every day, millions of people use natural-language interfaces such as Siri, Google Now, Cortana, Alexa and others via in-home devices, phones, or messaging channels such as Messenger, Slack, Skype, among others. At the same time, interest among the research community in conversational systems has blossomed: for supervised and reinforcement learning, conversational systems often serve as both a benchmark task and an inspiration for new ML methods at conferences which don't focus on speech and language per se, such as NIPS, ICML, IJCAI, and others. Research community challenge tasks are proliferating, including the seventh Dialog Systems Technology Challenge (DSTC7), the Amazon Alexa prize, and the Conversational Intelligence Challenge live competitions at NIPS (2017, 2018).

Following the overwhelming participation in our last year NIPS workshop (9 invited talks, 26 submissions, 3 orals papers, 13 accepted papers, 37 PC members, and couple of hundreds of participants), we are excited to continue promoting cross-pollination of ideas between academic research centers and industry. The goal of this workshop is to bring together researchers and practitioners in this area, to clarify impactful research problems, share findings from large-scale real-world deployments, and generate new ideas for future lines of research.

This workshop will include invited talks from academia and industry, contributed work, and open discussion. In these talks, senior technical leaders from many of the most popular conversational services will give insights into real usage and challenges at scale. An open call for papers will be issued, and we will prioritize forward-looking papers that propose interesting and impactful contributions. We will end the day with an open discussion, including a panel consisting of academic and industrial researchers.

Dec. 8, 2018

Integration of Deep Learning Theories

Richard Baraniuk, Anima Anandkumar, Stephane Mallat, Ankit Patel, nh■t H■

Room 220 D, Sat Dec 08, 08:00 AM

Deep learning has driven dramatic performance advances on numerous difficult machine learning tasks in a wide range of applications. Yet, its theoretical foundations remain poorly understood, with many more questions than answers. For example: What are the modeling assumptions underlying deep networks? How well can we expect deep networks to perform? When a certain network succeeds or fails, can we determine why and how? How can we adapt deep learning to new domains in a principled way?

While some progress has been made recently towards a foundational understanding of deep learning, most theory work has been disjointed, and a coherent picture has yet to emerge. Indeed, the current state of deep learning theory is like the fable “The Blind Men and the Elephant”.

The goal of this workshop is to provide a forum where theoretical researchers of all stripes can come together not only to share reports on their individual progress but also to find new ways to join forces towards the goal of a coherent theory of deep learning. Topics to be discussed include:

- Statistical guarantees for deep learning models
- Expressive power and capacity of neural networks
- New probabilistic models from which various deep architectures can be derived
- Optimization landscapes of deep networks
- Deep representations and invariance to latent factors
- Tensor analysis of deep learning
- Deep learning from an approximation theory perspective
- Sparse coding and deep learning
- Mixture models, the EM algorithm, and deep learning

In addition to invited and contributed talks by leading researchers from diverse backgrounds, the workshop will feature an extended poster/discussion session and panel discussion on which combinations of ideas are most likely to move theory of deep learning forward and which might lead to blind alleys.

Accepted Papers and Authors

1. A Convergence Analysis of Gradient Descent for Deep Linear Neural Networks. Sanjeev Arora, Nadav Cohen, Noah Golowich and Wei Hu.

2. On the convergence of SGD on neural nets and other over-parameterized problems. Karthik Abinav Sankararaman, Soham De, Zheng Xu, W. Ronny Huang and Tom Goldstein.

3. Optimal SGD Hyperparameters for Fully Connected Networks. Daniel Park, Samuel Smith, Jascha Sohl-Dickstein and Quoc Le.

4. Invariant representation learning for robust deep networks. Julian Salazar, Davis Liang, Zhiheng Huang and Zachary Lipton.

5. Characterizing & Exploring Deep CNN Representations Using Factorization. Uday Singh Saini and Evangelos Papalexakis.

6. On the Weak Neural Dependence Phenomenon in Deep Learning. Jiayao Zhang, Ruoxi Jia, Bo Li and Dawn Song.

7. DNN or k-NN: That is the Generalize vs. Memorize Question. Gilad Cohen, Guillermo Sapiro and Raja Giryes.

8. On the Margin Theory of Feedforward Neural Networks. Colin Wei, Jason Lee, Qiang Liu and Tengyu Ma.

9. A Differential Topological View of Challenges in Learning with Deep Neural Networks. Hao Shen.

10. Theoretical Analysis of Auto Rate-tuning by Batch Normalization. Sanjeev Arora, Zhiyuan Li and Kaifeng Lyu.

11. Topological Constraints onHomeomorphic Auto-Encoding. Pim de Haan and Luca Falorsi.

12. Deterministic PAC-Bayesian generalization bounds for deep networks via generalizing noise-resilience. Vaishnavh Nagarajan and J. Zico Kolter.

13. Directional Analysis of Stochastic Gradient Descent via von Mises-Fisher Distributions in Deep Learning. Cheolhyoung Lee, Kyunghyun Cho and Wanmo Kang.

14. Multi-dimensional Count Sketch: Dimension Reduction That Retains Efficient Tensor Operations. Yang Shi and Anima Anandkumar.

15. Gradient Descent Provably Optimizes Over-parameterized Neural Networks. Simon Du, Xiyu Zhai, Aarti Singh and Barnabas Poczos.

16. The Dynamic Distance Between Learning Tasks. Alessandro Achille, Glen Bigan Mbeng and Stefano Soatto.

17. Stochastic Gradient/Mirror Descent: Minimax Optimality and Implicit Regularization. Navid Azizan and Babak Hassibi.

18. Shared Representation Across Neural Networks. Qihong Lu, Po-Hsuan Chen, Jonathan Pillow, Peter Ramadge, Kenneth Norman and Uri Hasson.

19. Learning in gated neural networks. Ashok Makkuva, Sewoong Oh, Sreeram Kannan and Pramod Viswanath.

20. Gradient descent aligns the layers of deep linear networks. Ziwei Ji and Matus Telgarsky.

21. Fluctuation-dissipation relation for stochastic gradient descent. Sho Yaida.

22. Identifying Generalization Properties in Neural Networks. Huan

Wang, Nitish Shirish Keskar, Caiming Xiong and Richard Socher.

23. A Theoretical Framework for Deep and Locally Connected ReLU Network. Yuandong Tian.

24. Minimum norm solutions do not always generalize well for over-parameterized problems. Vatsal Shah, Anastasios Kyrillidis and Sujay Sanghavi.

25. An Empirical Exploration of Gradient Correlations in Deep Learning. Daniel Rothchild, Roy Fox, Noah Golmant, Joseph Gonzalez, Michael Mahoney, Kai Rothauge, Ion Stoica and Zhewei Yao.

26. Geometric Scattering on Manifolds. Michael Perlmutter, Guy Wolf and Matthew Hirn.

27. Theoretical Insights into Memorization in GANs. Vaishnavh Nagarajan, Colin Raffel and Ian Goodfellow.

28. A jamming transition from under- to over-parametrization affects loss landscape and generalization. Stefano Spigler, Mario Geiger, Stéphane d'Ascoli, Levent Sagun, Giulio Biroli and Matthieu Wyart.

29. A Mean Field Theory of Multi-Layer RNNs. David Anderson, Jeffrey Pennington and Satyen Kale.

30. Generalization and regularization in deep learning for nonlinear inverse problems. Christopher Wong, Maarten de Hoop and Matti Lassas.

31. On the Spectral Bias of Neural Networks. Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio and Aaron Courville.

32. On Generalization Bounds for a Family of Recurrent Neural Networks. Minshuo Chen, Xingguo Li and Tuo Zhao.

33. SGD Implicitly Regularizes Generalization Error. Dan Roberts.

34. Iteratively Learning from the Best. Yanyao Shen and Sujay Sanghavi.

35. Towards Understanding the Role of Over-Parametrization in Generalization of Neural Networks. Behnam Neyshabur, Zhiyuan Li, Srinadh Bhojanapalli, Yann LeCun and Nathan Srebro.

36. An Escape-Time Analysis of SGD. Philippe Casgrain, Mufan Li, Gintare Karolina Dziugaite and Daniel Roy.

37. Information Regularized Neural Networks. Tianchen Zhao, Dejiao Zhang, Zeyu Sun and Honglak Lee.

38. Generalization Bounds for Unsupervised Cross-Domain Mapping with WGANs. Tomer Galanti, Sagie Benaim and Lior Wolf.

39. Degeneracy, Trainability, and Generalization in Deep Neural Networks. Emin Orhan and Xaq Pitkow.

40. A Max-Affine Spline View of Deep Network Nonlinearities. Randall Balestriero and Richard Baraniuk.

Schedule

Schedule

08:30 AM	**Opening Remarks**	
08:40 AM	**Contributed Talk 1**	*Lee*
09:00 AM	**Contributed Talk 2**	*Perlmutter*
09:20 AM	**Plenary Talk 1**	*Arora*
10:00 AM	**Invited Talk 1**	*Lipton*
10:30 AM	**Coffee Break**	
10:50 AM	**Plenary Talk 2**	*Chaudhuri*
11:30 AM	**Invited Talk 2**	*Hoffman*
12:00 PM	**Lunch Break**	
01:30 PM	**Plenary Talk 3**	*Soatto*
02:10 PM	**Invited Talk 3**	*Higgins*
02:40 PM	**Contributed Talk 3**	*Nguyen*
03:00 PM	**Poster Session**	*Sanghavi, Shah, Shen, Zhao, Tian, Galanti, Li, Cohen, Rothchild, Baratin, Arpit, Papalexakis, Perlmutter, Makkuva, de Haan, Lin, Kang, Lee, Shen, Yaida, Roberts, Cohen, Casgrain, Zhang, Ma, Ravichandran, Salazar, Li, Liang, Wong, Mbeng, Garg*
03:50 PM	**Plenary Talk 4**	*Fox*
04:30 PM	**Invited Talk 4**	*Sulam*
05:00 PM	**Panel Discussion**	
05:55 PM	**Closing Remarks**	

NIPS 2018 Workshop on Meta-Learning

Erin Grant, Frank Hutter, Sachin Ravi, Joaquin Vanschoren, Jane Wang

Room 220 E, Sat Dec 08, 08:00 AM

Recent years have seen rapid progress in meta-learning methods, which learn (and optimize) the performance of learning methods based on data, generate new learning methods from scratch, and learn to transfer knowledge across tasks and domains. Meta-learning can be seen as the logical conclusion of the arc that machine learning has undergone in the last decade, from learning classifiers, to learning representations, and finally to learning algorithms that themselves acquire representations and classifiers. The ability to improve one’s own learning capabilities through experience can also be viewed as a hallmark of intelligent beings, and there are strong connections with work on human learning in neuroscience.

Meta-learning methods are also of substantial practical interest, since they have, e.g., been shown to yield new state-of-the-art automated machine learning methods, novel deep learning architectures, and

substantially improved one-shot learning systems.

Some of the fundamental questions that this workshop aims to address are:
- What are the fundamental differences in the learning “task” compared to traditional “non-meta” learners?
- Is there a practical limit to the number of meta-learning layers (e.g., would a meta-meta-meta-learning algorithm be of practical use)?
- How can we design more sample-efficient meta-learning methods?
- How can we exploit our domain knowledge to effectively guide the meta-learning process?
- What are the meta-learning processes in nature (e.g, in humans), and how can we take inspiration from them?
- Which ML approaches are best suited for meta-learning, in which circumstances, and why?
- What principles can we learn from meta-learning to help us design the next generation of learning systems?

The goal of this workshop is to bring together researchers from all the different communities and topics that fall under the umbrella of meta-learning. We expect that the presence of these different communities will result in a fruitful exchange of ideas and stimulate an open discussion about the current challenges in meta-learning, as well as possible solutions.

In terms of prospective participants, our main targets are machine learning researchers interested in the processes related to understanding and improving current meta-learning algorithms. Specific target communities within machine learning include, but are not limited to: meta-learning, AutoML, reinforcement learning, deep learning, optimization, evolutionary computation, and Bayesian optimization. Our invited speakers also include researchers who study human learning, to provide a broad perspective to the attendees.

Schedule

09:10 AM	**TBA**	*Getoor*
11:00 AM	**TBA 2**	*Levine*
01:30 PM	**TBA 3**	*Larochelle*
02:00 PM	**TBA 4**	*Sebag*
04:00 PM	**TBA 5**	*de Freitas*

Machine Learning for Systems

Anna Goldie, Azalia Mirhoseini, Jonathan Raiman, Kevin Swersky, Milad Hashemi

Room 510 AC, Sat Dec 08, 08:00 AM

This workshop is part two of a two-part series with one day focusing on Machine Learning for Systems and the other on Systems for Machine Learning. Although the two workshops are being led by different organizers, we are coordinating our call for papers to ensure that the workshops complement each other and that submitted papers are routed to the appropriate venue.

The Systems for Machine Learning workshop focuses on designing systems to enable ML, whereas we focus on developing ML to optimize systems. Both fields are mature enough to warrant a dedicated workshop. Organizers on both sides are open to merging in the future, but this year we plan to run them separately on two different days.

Designing specialized hardware and systems for deep learning is a topic that has received significant research attention, both in industrial and academic settings, leading to exponential increases in compute capability in GPUs and accelerators. However, using machine learning to optimize and accelerate software and hardware systems is a lightly explored but promising field, with broad implications for computing as a whole. Very recent work has outlined a broad scope where deep learning vastly outperforms traditional heuristics, including topics such as: scheduling [1], data structure design [2], microarchitecture [3], compilers [4], and control of warehouse scale computing systems [5].

The focus of this workshop is to expand upon this recent work and build a community focused on using machine learning in computer systems problems. We seek to improve the state of the art in the areas where learning has already proven to perform better than traditional heuristics, as well as expand to new areas throughout the system stack such as hardware/circuit design and operating/runtime systems.

By forming a community of academic and industrial researchers who are excited about this area, we seek to build towards intelligent, self optimizing systems and answer questions such as: How do we generate and share high quality datasets that span the layers of the system stack? Which learned representations best represent code performance and runtime? Which simulators and simulation methodologies provide a tractable proving ground for techniques like reinforcement learning?

To this end, the target audience for this workshop includes a wide variety of attendees from state-of-the-art researchers in machine learning to domain experts in computer systems design. We have invited a broad set of expert speakers to present the potential for impact of combining machine learning research with computer systems. We hope that providing a formal venue for researchers from both fields to meet and interact will push forward both fundamental research in ML as well as real-world impact to computer systems design and implementation.

The workshop will host 6 speakers/panelists (all confirmed) and we will put out a call for researchers to submit relevant papers, up to 4 pages in the default NIPS style, that will undergo a peer review process. Selected works will be presented as spotlights, contributed talks and/or posters. Speakers will be invited to participate in an interactive panel discussion to conclude the workshop.

The organizers of this workshop span core research in machine learning, computer systems and architecture, as well as their intersection. Jointly, they have published in top-tier systems and machine learning conferences including: NIPS, ICML, ICLR, ISCA, MICRO, DAC, and SIGMETRICS.

References:
[1] Device Placement Optimization with Reinforcement Learning, https://arxiv.org/pdf/1706.04972.pdf
[2] The Case for Learned Index Structures, https://arxiv.org/abs/1712.01208
[3] Learning Memory Access Patterns, https://arxiv.org/pdf/1803.02329.pdf

[4] End to End Deep Learning of Optimization Heuristics: https://ieeexplore.ieee.org/document/8091247/?reload=true
[5] https://deepmind.com/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-40/
[6] Bayesian optimization for tuning the JVM, https://www.youtube.com/watch?v=YhNl468S8CI
[7] Safe Exploration for Identifying Linear Systems via Robust Optimization: https://arxiv.org/abs/1711.11165

Schedule

09:00 AM	**Opening**	
09:10 AM	**Invited Speaker 1: Eric Schkufza**	*schkufza*
09:35 AM	**Invited Speaker 2: Song Han**	
10:00 AM	**Poster Session (All Posters)**	*Margaritov, addanki, Mahyar, ZHANG, wildani, Esmaeilzadeh, Ustiugov, Bojja Venkatakrishnan, Ruffy Varga, bhardwaj, Shpeisman*
11:00 AM	**Neural Inference of API Functions from Input–Output Examples**	
11:15 AM	**Placeto: Efficient Progressive Device Placement Optimization**	
11:30 AM	**Iroko: A Framework to Prototype Reinforcement Learning for Data Center Traffic Control**	
11:45 AM	**Invited Speaker 3: Partha Ranganathan**	*Ranganathan*
01:45 PM	**Invited Speaker 4: Neeraja J. Yadwadkar**	*Yadwadkar*
02:10 PM	**Learning to Optimize Tensor Programs**	
02:25 PM	**Learning to Design Circuits**	
02:40 PM	**ReLeQ: A Reinforcement Learning Approach for Deep Quantization of Neural Networks**	
03:00 PM	**Poster Session (All Posters)**	*Macke, Mao, Lemieux, Salman, Jha, Wang, Palkar, Chen, Pumir, Janardhan, bhardwaj, Chi*
04:00 PM	**Invited Speaker 6: Sanjay Krishnan**	
04:25 PM	**Invited Speaker 7: Jeff Dean**	
04:50 PM	**Panel**	
05:50 PM	**Closing**	

Abstracts (11):

Abstract 2: **Invited Speaker 1: Eric Schkufza in Machine Learning for Systems**, *schkufza* 09:10 AM

Traditional compiler use expert-written rules to prove the correctness of program transformations, and hope for the best in terms of performance. Stochastic program optimizers turn that model on its head. They use machine learning techniques to search for aggressive performance-improving transformations, and state-of-the-art verification techniques to prove correctness after the fact. The results are novel, often inscrutable, and in many cases outperform expertly tuned code. In this talk I'll present an overview of the core technique, describe current work, and discuss directions for future research.

Abstract 3: **Invited Speaker 2: Song Han in Machine Learning for Systems**, 09:35 AM

In the post-Moore’s Law era, the amount of computation per unit cost and power is no longer increasing at its historic rate. In the post-ImageNet era, researchers are solving more complicated AI problems using larger data sets which drives the demand for more computation. This mismatch between supply and demand for computation highlights the need for co-designing efficient machine learning algorithms and domain-specific hardware architectures. Such algorithm-hardware co-design opens up a much larger design space, which requires domain experts on both sides (ML+systems), and human heuristics might be sub-optimal to explore the vast design space. We introduce three of our recent work of using machine learning to optimize the machine learning system: learning the optimal pruning strategy (AMC) and quantization strategy (HAQ) on the target hardware, rather than relying on rule-based strategies; learning the optimal neural network architecture that is specialized for a target hardware architecture, optimizing both accuracy and latency (ProxylessNAS), rather than using a generic neural network architecture across all hardware architectures; learning to optimize analog circuit parameters, rather than relying on experienced analog engineers to tune those transistors. On the other side of the loop (design hardware-friendly machine learning algorithms), I'll introduce the temporal shift module (TSM) that offers 8x lower latency, 12x higher throughput than 3D convolution-based methods, while ranking the first on both Something-Something V1 and V2 leaderboards. I’ll conclude the talk by giving an outlook of the design automation for efficient machine learning system.

Abstract 5: **Neural Inference of API Functions from Input–Output Examples in Machine Learning for Systems**, 11:00 AM

Because of the prevalence of APIs in modern software development, an automated interactive code discovery system to help developers use these APIs would be extremely valuable. Program synthesis is a promising method to build such a system, but existing approaches focus on programs in domain-specific languages with much fewer functions than typically provided by an API. In this paper we focus on 112 functions from the Python library for DataFrame manipulation, an order of magnitude more than considered in prior approaches. To assess the viability of program synthesis in this domain, our first goal is a system that reliably synthesizes programs with a single library function. We introduce an encoding of structured input--output examples as graphs that can be fed to existing graph-based neural networks to infer the library function. We evaluate the effectiveness of this approach on synthesized and real-world I/O examples, finding programs matching the I/O examples for 97% of both our validation set and cleaned test set.

Abstract 6: **Placeto: Efficient Progressive Device Placement Optimization in Machine Learning for Systems**, 11:15 AM

We present Placeto, a reinforcement learning (RL) approach to efficiently find device placements for distributed neural network training. Unlike prior approaches that only find a device placement for a specific computational graph, Placeto can learn generalizable device placement policies that can be applied to any graph. We propose two key ideas in our approach: (1) we represent the policy as performing iterative placement improvements, rather than outputting a placement in one shot (2) we use graph embeddings to capture the structural information of the computational graph, without relying on node labels for indexing. These ideas allow Placeto to train efficiently and generalize to unseen graphs. Our experiments show that Placeto can take up to 20x fewer training steps to find placements that are on par with or better than the best placements found by prior approaches.

Abstract 7: **Iroko: A Framework to Prototype Reinforcement Learning for Data Center Traffic Control in Machine Learning for Systems**, 11:30 AM

Recent networking research has identified that data-driven congestion control (CC) can be more efficient than traditional CC in TCP. Deep reinforcement learning (RL), in particular, has the potential to learn optimal network policies. However, RL suffers from instability and over-fitting, deficiencies which so far render it unacceptable for use in datacenter networks. In this paper, we analyze the requirements for RL to succeed in the datacenter context. We present a new emulator, Iroko, which we developed to support different network topologies, congestion control algorithms, and deployment scenarios. Iroko interfaces with the OpenAI gym toolkit, which allows for fast and fair evaluation of different RL and traditional CC algorithms under the same conditions. We present initial benchmarks on three deep RL algorithms compared to TCP New Vegas and DCTCP. Our results show that these algorithms are able to learn a CC policy which exceeds the performance of TCP New Vegas on a dumbbell and fat-tree topology. We make our emulator open-source and publicly available: https://github.com/dcgym/iroko.

Abstract 8: **Invited Speaker 3: Partha Ranganathan in Machine Learning for Systems**, *Ranganathan* 11:45 AM

The computer architecture is facing an important and exciting challenge. The slowing of Moore's law (at the same time demand continues to grow) has led to new approaches to thinking about future system design including accelerators and software-defined hardware. In this talk we will discuss how machine learning has the potential to amplify these opportunities. We will discuss some specific case studies and end with some key insights specific to applying machine learning to improve computer architecture.

Abstract 9: **Invited Speaker 4: Neeraja J. Yadwadkar in Machine Learning for Systems**, *Yadwadkar* 01:45 PM

Traditional resource management techniques that rely on simple heuristics often fail to achieve predictable performance in contemporary complex systems that span physical servers, virtual servers, private and/or public clouds. My research aims to bring the benefits of Machine Learning (ML) models to optimize and manage such complex systems by deriving actionable insights from the performance and utilization data these systems generate. To realize this vision of model-based resource management, we need to deal with the following key challenges data-driven ML models raise: uncertainty in predictions, cost of training, generalizability from benchmark datasets to real-world systems datasets, and interpretability of the models.

In this talk, I will present our the ML formulations to demonstrate how to handle these challenges for two main problem domains in distributed systems: (I) Scheduling in parallel data-intensive computational frameworks for improved tail latencies, and (II) Performance-aware resource allocation in the public cloud environments for meeting user-specified performance and cost goals. Along the way, I will also share a list of guidelines for leveraging ML for solving problems in systems, based on my experience.

Abstract 10: **Learning to Optimize Tensor Programs in Machine Learning for Systems**, 02:10 PM

We introduce a learning-based framework to optimize tensor programs for deep learning workloads. Efficient implementations of tensor operators, such as matrix multiplication and high dimensional convolution, are key enablers of effective deep learning systems. However, current systems rely on manually optimized libraries, e.g., cuDNN, that support only a narrow range of server class GPUs. Such reliance limits the applicability of high-level graph optimizations and incurs significant engineering costs when deploying to new hardware targets. We use learning to remove this engineering burden. We learn domain-specific statistical cost models to guide the search of tensor operator implementations over billions of possible program variants. We further accelerate the search using effective model transfer across workloads. Experimental results show that our framework delivers performance that is competitive with state-of-the-art hand-tuned libraries for low-power CPUs, mobile GPUs, and server-class GPUs.

Abstract 11: **Learning to Design Circuits in Machine Learning for Systems**, 02:25 PM

Analog IC design relies on human experts to search for parameters that satisfy circuit specifications with their experience and intuitions, which is highly labor intensive, time consuming and suboptimal.
Machine learning is a promising tool to automate this process. However, supervised learning is difficult for this task due to the low availability of training data: 1) Circuit simulation is slow, thus generating large-scale dataset is time-consuming; 2) Most circuit designs are propitiatory IPs within individual IC companies, making it expensive to collect large-scale datasets.
We propose Learning to Design Circuits (L2DC) to leverage reinforcement learning that learns to efficiently generate new circuits data and to optimize circuits. We fix the schematic, and optimize the parameters of the transistors automatically by training an RL agent with no prior knowledge about optimizing circuits. After iteratively getting observations, generating a new set of transistor parameters, getting a reward, and adjusting the model, L2DC is able to optimize circuits.
We evaluate L2DC on two transimpedance amplifiers. Trained for a day, our RL agent can achieve comparable or better performance than human experts trained for a quarter. It first learns to meet hard-constraints (eg. gain, bandwidth), and then learns to optimize good-to-have targets (eg. area, power). Compared with grid search-aided human design, L2DC can achieve 250x higher sample efficiency with comparable performance. Under the same runtime constraint, the performance of L2DC is also better than Bayesian Optimization.

Abstract 12: **ReLeQ: A Reinforcement Learning Approach for Deep Quantization of Neural Networks in Machine Learning for Systems**,

02:40 PM

Despite numerous state-of-the-art applications of Deep Neural Networks (DNNs) in a wide range of real-world tasks, two major challenges hinder further advances in DNNs: hyperparameter optimization and constrained power resources, which is a significant concern in embedded devices. DNNs become increasingly difficult to train and deploy as they grow in size due to both computational intensity and the large memory footprint. Recent efforts show that quantizing weights of deep neural networks to lower bitwidths takes a significant step toward mitigating the mentioned issues, by reducing memory bandwidth and using limited computational resources which is important for deploying DNN models to devices with limited resources. This paper builds upon the algorithmic insight that the bitwidth of operations in DNNs can be reduced without compromising their classification accuracy. Deep quantization (quantizing bitwidths below eight) while maintaining accuracy, requires magnificent manual effort and hyper-parameter tuning as well as re-training. This paper tackles the aforementioned problems by designing an end to end framework, dubbed ReLeQ, to automate DNN quantization. We formulate DNN quantization as an optimization problem and use a state-of-the-art policy gradient based Reinforcement Learning (RL) algorithm, Proximal Policy Optimization (PPO) to efficiently explore the large design space of DNN quantization and solve the defined optimization problem. To show the effectiveness of ReLeQ, we evaluated it across several neural networks including MNIST, CIFAR10, SVHN. ReLeQ quantizes the weights of these networks to average bitwidths of 2.25, 5 and 4 respectively while maintaining the final accuracy loss below 0.3% .

Abstract 14: **Invited Speaker 6: Sanjay Krishnan in Machine Learning for Systems**, 04:00 PM

To integrate information from more than two tables, a SQL query optimizer must identify the most efficient nesting of two-way table join operations to answer the query. Recent advances in AI may provide an unexpected new perspective on this classical problem that has been studied for over 40 years. Join optimization can be posed as a Markov Decision Process where the state is a graph that represents the join conditions in a query and actions are edge contractions on this graph; thereby, allowing us to apply ideas from deep reinforcement learning and imitation learning to facilitate an improved query optimizer that learns from experience, handles uncertainty, and incorporates execution feedback. I describe how our group built a full-featured query optimizer based on this MDP architecture, and we present results across a variety of database designs and query workloads in Postgres SQL and Apache Spark. I conclude by highlighting some of the under-appreciated RL research challenges in exploration, parametrization, and policy evaluation unearthed by this application.

Machine Learning for the Developing World (ML4D): Achieving sustainable impact

Maria De-Arteaga, William Herlands, Amanda Coston

Room 510 BD, Sat Dec 08, 08:00 AM

Global development experts are beginning to employ ML for diverse problems such as aiding rescue workers allocate resources during natural disasters, providing intelligent educational and healthcare services in regions with few human experts, and detecting corruption in government contracts. While ML represents a tremendous hope for accelerated development and societal change, it is often difficult to ensure that machine learning projects provide their promised benefit. The challenging reality in developing regions is that pilot projects disappear after a few years or do not have the same effect when expanded beyond the initial test site, and prototypes of novel methodologies are often never deployed.

At the center of this year's program is how to achieve sustainable impact of Machine Learning for the Developing World (ML4D). This one-day workshop will bring together a diverse set of participants from across the globe to discuss major roadblocks and paths to action. Practitioners and development experts will discuss essential elements for ensuring successful deployment and maintenance of technology in developing regions. Additionally, the workshop will feature cutting edge research in areas such as transfer learning, unsupervised learning, and active learning that can help ensure long-term ML system viability. Attendees will learn about contextual components to ensure effective projects, development challenges that can benefit from machine learning solutions, and how these problems can inspire novel machine learning research.

The workshop will include invited and contributed talks, a poster session of accepted papers, panel discussions, and breakout sessions tailored to the workshop theme. We welcome paper submissions focussing on core ML methodology addressing ML4D roadblocks, application papers that showcase successful examples of ML4D, and research that evaluates the societal impact of ML.

Schedule

08:45 AM	**Introductory remarks**	*Dubrawski*
09:00 AM	**Early lessons in ML4d from the field**	*Panicker, Anandan*
09:30 AM	**Taking "Big Data" evidence to policy: Experiences from the Global South**	*Lokanathan*
10:00 AM	**Exploring data science for public good in South Africa: evaluating factors that lead to success**	*Moorosi*
11:00 AM	**Forecasting Internally Displaced Population Migration Patterns in Syria and Yemen**	*Huynh*
11:15 AM	**Regression by clustering using Metropolis Hastings**	*Ramírez Amaya*
11:30 AM	**Poster session: Contributed papers**	*Cvitkovic, Patra, Li, SANYA, Chi, Huynh, Alemohammad, Ramírez Amaya, Saquib, Abbott, de Campos, Prabhu, Riascos, Abera, dubey, Chattopadhyay, Hsu, Jain, Bhardwaj, Cadamuro, GRAM-HANSEN, Dorffner*
02:00 PM	**BCCNet: Bayesian classifier combination neural network**	*Isupova*

02:15 PM	**Point-of-care ultrasound in the global south: A case of fetal heart anomaly assessment with mobile devices**	
02:30 PM	**Machine Learning for Development: Challenges, Opportunities, and a Roadmap**	*Neill*
03:30 PM	**Using ML to locate hidden graves in Mexico**	*Meltis Vejar*
04:00 PM	**Real-Time Measures of Poverty and Vulnerability**	*Blumenstock*
04:30 PM	**Challenges and Opportunities in ML4D**	

Abstracts (9):

Abstract 2: **Early lessons in ML4d from the field in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Panicker, Anandan* 09:00 AM

In this talk, I will share lessons from our efforts on the ground in creating AI-for-social-good solutions, spanning cellphone-image-based anthropometry for babies, AI-enabled active case-finding in tuberculosis, and early pest detection in cotton farming. The promise of AI as a powerful aid for achieving global-development goals is bolstered by five current forces: large frontline workforces enabling service delivery and data collection, growing smartphone penetration providing compute, connectivity, imaging, localization, and interfaces, large tech-enabled development programs having established data pipelines, infrastructure, and processes, rural populations increasingly adopting technology, and strong policy and institutional support for AI in development. We recommend that AI-for-social-good efforts utilize these forces by piggybacking on large tech-enabled development programs to achieve scaled impact. I will provide examples for such programs, point to opportunity areas, list criteria for AI-for-social-good innovators to assess likelihood of scaled impact, discuss risks and mitigation strategies, and suggest frontier areas for AI-for-social-good research.

Abstract 3: **Taking "Big Data" evidence to policy: Experiences from the Global South in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Lokanathan* 09:30 AM

LIRNEasia has been working on leveraging big data for public purposes since 2012. As an
organization situated in a developing country, we have experienced challenges in developing
new insights, and informing policy and government processes. When leveraging big data and
machine learning for development purposes, developing countries face three main inter-related
challenges:
1. Skills: data scientists are in short supply and developing skills to make use of these new data
sources become paramount. How should we build these skills? What should be the
composition of research teams?
2. Data: accessing private sector data as well as government data can both be challenging. In
an imperfect, often inconsistent regulatory environment, how can we facilitate responsible
data access and use?
3. Policy impact and mainstreaming: Except in extreme cases most policy domains already
have pre-existing established processes for generating and incorporating evidence in policy
planning and implementation. How do we disrupt these ‘sticky’ processes with new forms of
data and techniques?
This talk will address these three sets of challenges and our experiences in tackling them.

Abstract 4: **Exploring data science for public good in South Africa: evaluating factors that lead to success in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Moorosi* 10:00 AM

In the pursuit of public service, governments have to oversee many complex systems. In recent years, data-driven methodologies have been adopted as tools to oversee and enhance service delivery. In this talk I will discuss the ways that the government of South Africa, and its agencies, use data tools as well as the policies and investments that they have been put into place; some of which have created a more enabling ecosystem while others have created difficulties and challenges. I will discuss the current data landscape from the lens of Open Data policies, data readiness policies, and human capital development initiatives. This talk will be a summary of the work we have done in the past four years. It will be a discussion of our observations, our successes, aspirations and challenges we encountered as we continue towards a data-driven governance.

Abstract 5: **Forecasting Internally Displaced Population Migration Patterns in Syria and Yemen in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Huynh* 11:00 AM

Armed conflict has contributed to an unprecedented number of internally displaced persons (IDPs) - individuals who are forced out of their homes but remain within their country. IDPs often urgently require shelter, food, and healthcare, yet prediction of when fluxes of IDPs will cross into an area remains a major challenge for aid delivery organizations. We sought to develop an approach to more accurately forecast IDP migration that could empower humanitarian aid groups to more effectively allocate resources during conflicts. We modeled monthly IDP flow between provinces within Syria and within Yemen using heterogeneous data on food prices, fuel prices, wages, location, time, and conflict reports. We show
that our machine learning approach outperforms baseline persistence methods of forecasting. Integrating diverse data sources into machine learning models thus appears to improve IDP migration prediction.

Abstract 6: **Regression by clustering using Metropolis Hastings in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Ramírez Amaya* 11:15 AM

High quality risk adjustment in health insurance markets weakens insurer incentives
to engage in inefficient behavior to attract lower-cost enrollees. We propose a

novel methodology based on Markov Chain Monte Carlo methods to improve
risk adjustment by clustering diagnostic codes into risk groups optimal for health
expenditure prediction. We test the performance of our methodology against
common alternatives using panel data from 3.5 million enrollees of the Colombian
Healthcare System. Results show that our methodology outperforms common
alternatives and suggest that it has potential to improve access to quality healthcare
for the chronically ill.

Abstract 8: **BCCNet: Bayesian classifier combination neural network in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Isupova* 02:00 PM

Machine learning research for developing countries can demonstrate clear sustainable impact by delivering actionable and timely information to in-country government organisations (GOs) and NGOs in response to their critical information requirements. We co-create products with UK and in-country commercial, GO and NGO partners to ensure the machine learning algorithms address appropriate user needs whether for tactical decision making or evidence-based policy decisions. In one particular case, we developed and deployed a novel algorithm, BCCNet, to quickly process large quantities of unstructured data to prevent and respond to natural disasters. Crowdsourcing provides an efficient mechanism to generate labels from unstructured data to prime machine learning algorithms for large scale data analysis. However, these labels are often imperfect with qualities varying among different citizen scientists, which prohibits their direct use with many state-of-the-art machine learning techniques. We describe BCCNet, a framework that simultaneously aggregates biased and contradictory labels from the crowd and trains an automatic classifier to process new data. Our case studies, mosquito sound detection for malaria prevention and damage detection for disaster response, show the efficacy of our method in the challenging context of developing world applications.

Abstract 9: **Point-of-care ultrasound in the global south: A case of fetal heart anomaly assessment with mobile devices in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, 02:15 PM

A major challenge in pre-natal healthcare delivery is the lack of devices and clinicians in several areas of the developing world. While the advent of portable ultrasound machines and more recently, handheld probes, have brought down the capital costs, the shortage of trained manpower is a serious impediment towards ensuring the mitigation of maternal and infant mortality. Diagnosis of pre-natal ultrasound towards several key pre-natal health indicators can be modelled as an image analysis problem amenable to present day state-of-the art deep learning based image and video understanding pipelines. However, deep learning based analysis typically involves memory intensive models and the requirement of significant computational resources, which is a challenging prospect in point-of-care healthcare applications in the developing world. With the advent of portable ultra-sound systems, it is increasingly possible to expand the reach of prenatal health diagnosis. To accomplish that, there is a need for lightweight architectures that can perform image analysis tasks without a large memory or computational footprint. We propose a lightweight convolutional architecture for assessment of ultrasound videos, suitable for those acquired using mobile probes or converted from a DI-COM standard from portable machines. As exemplar of approach, we validated our pipeline for fetal heart assessment (a first step towards identification of congenital heart defects) inclusive of viewing plane identification and visibility prediction in fetal echocardiography. This was attempted by models using optimised kernel windows and the construction of image representations using salient features from multiple scales with relative feature importance gauged at each of these scales using weighted attention maps for different stages of the convolutional operations. Such a representation is found to improve model performances at significant economization of model size, and has been validated on real-world clinical videos.

Abstract 10: **Machine Learning for Development: Challenges, Opportunities, and a Roadmap in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Neill* 02:30 PM

Researchers from across the social and computer sciences are increasingly using machine learning to study and address global development challenges, and an exciting new field of "Machine Learning for the Developing World", or "Machine Learning for Development" (ML4D) is beginning to emerge. In recent work (De Arteaga et al., ACM TMIS, 2018), we synthesize the prominent literature in the field and attempt to answer the key questions, "What is ML4D, and where is the field headed?". Based on the literature, we identify a set of best practices for ensuring that ML4D projects are relevant to the advancement of development objectives. Given the strong alignment between development needs and ML approaches, we lay out a roadmap detailing three technical stages where ML4D can play an essential role and meaningfully contribute to global development. Perhaps the most important aspect of ML4D is that development challenges are treated as research questions, not as roadblocks: we believe that the ML4D field can flourish in the coming years by using the unique challenges of the developing world as opportunities to inspire novel and impactful research across multiple machine learning disciplines. This talk is based on joint work with Maria de Arteaga, William Herlands, and Artur Dubrawski.

Abstract 12: **Real-Time Measures of Poverty and Vulnerability in Machine Learning for the Developing World (ML4D): Achieving sustainable impact**, *Blumenstock* 04:00 PM

In wealthy nations, novel sources of data from the internet and social media are enabling new approaches for social science research and public policy. In developing countries, by contrast, fewer sources of such data exist, and researchers and policymakers often rely on data that are unreliable or out of date. Here, we develop a new approach for measuring the dynamic welfare of individuals remotely by analyzing their logs of mobile phone use. We calibrate our approach with an original high-frequency panel survey of 1,200 Afghans, and an experimental protocol that randomized the timing and value of an unconditional cash transfer to each respondent. We show that mobile phone metadata, obtained with the respondent's consent from Afghanistan's largest mobile phone company, can be used to estimate the social and economic well-being of respondents, including the onset of positive and negative shocks. We discuss the potential for such methods to transform current practices of policy monitoring and impact evaluation.

CiML 2018 - Machine Learning competitions "in the wild": Playing in the real world or in real time

Isabelle Guyon, Evelyne Viegas, Sergio Escalera, Jacob D Abernethy

Room 511 ABDE, Sat Dec 08, 08:00 AM

Challenges in machine learning and data science are competitions running over several weeks or months to resolve problems using provided datasets or simulated environments. The playful nature of challenges naturally attracts students, making challenge a great teaching resource. For this fifth edition of the CiML workshop at NIPS we want to go beyond simple data science challenges using canned data. We will explore the possibilities offered by challenges in which code submitted by participants are evaluated "in the wild", directly interacting in real time with users or with real or simulated systems. Organizing challenges "in the wild" is not new. One of the most impactful such challenge organized relatively recently is the DARPA grant challenge 2005 on autonomous navigation, which accelerated research on autonomous vehicles, leading to self-driving cars. Other high profile challenge series with live competitions include RoboCup, which has been running from the past 22 years. Recently, the machine learning community has started being interested in such interactive challenges, with last year at NIPS the learning to run challenge, an reinforcement learning challenge in which a human avatar had to be controlled with simulated muscular contractions, and the ChatBot challenge in which humans and robots had to engage into an intelligent conversation. Applications are countless for machine learning and artificial intelligence programs to solve problems in real time in the real world, by interacting with the environment. But organizing such challenges is far from trivial

The workshop will give a large part to discussions around two principal axes: (1) Design principles and implementation issues; (2) Opportunities to organize new impactful challenges.

Our objectives include bringing together potential partner to organize new such challenges and stimulating "machine learning for good", i.e. the organization of challenges for the benefit of society.

CiML is a forum that brings together workshop organizers, platform providers, and participants to discuss best practices in challenge organization and new methods and application opportunities to design high impact challenges. Following the success of previous years' workshops, we propose to reconvene and discuss new opportunities for challenges "in the wild", one of the hottest topics in challenge organization. We have invited prominent speakers having experience in this domain.

The audience of this workshop is targeted to workshop organizers, participants, and anyone with scientific problem involving machine learning, which may be formulated as a challenge. The emphasis of the workshop is on challenge design. Hence it complements nicely the workshop on the NIPS 2018 competition track and will help paving the way toward next year's competition program.

Submit abstract (up to 2 pages) before October 10 by sending email to nips2018@chalearn.org. See http://ciml.chalearn.org/ciml2018#CALL.

Schedule

08:00 AM	**Morning Welcome - - Isabelle Guyon and Evelyne Viegas**	*Viegas*
08:10 AM	**Esteban Arcaute and Umut Ozertem - Facebook project on developing benchmarks of algorithms in realistic settings**	*Arcaute*
08:40 AM	**Laura Seaman - Project Alloy – Machine Learning Challenges for Researching Human-Machine Teaming**	*Seaman*
09:10 AM	**Live competiton 1: Pommerman**	
10:40 AM	**How to fail hosting data science contests with images, Evgeny Nizhibitsky, Artur Kuzin**	*Kuzin, Nizhibitskiy*
10:40 AM	**ML Benchmark Tools Package, Ryan Turner, Uber AI Labs**	*Turner*
10:40 AM	**Corpus for AutoML Pipelines, Richard Lippmann, Swoop Vattam, Pooya Khorrami, and Cagri Dagli**	*Lippmann, Khorrami*
10:40 AM	**Grand-Challenge.org, James A. Meakin Bram van Ginneken**	*Meakin*
10:40 AM	**Beyond the Leaderboard, Adrienne M. Mendrik, Stephen R. Aylward**	*Mendrik, Aylward*
10:40 AM	**Break, poster viewing 1**	
11:00 AM	**Julien Hay, Bich-Liên Doan, Fabrice Popineau - Renewal news recommendation platform**	*Hay, DOAN*
11:30 AM	**Panel discussion. Design principles and implementation issues.**	*Escalera*
12:00 PM	**Lunch Break, poster viewing 2**	
01:00 PM	**Afternoon Welcome - Isabelle Guyon and Evelyne Viegas**	*Guyon*
01:10 PM	**Mikhail Burtsev and Varvara Logacheva - Wild evaluation of chat-bots**	*Burtsev, Logacheva*
01:40 PM	**Larry Jackel - Measuring Progress in Robotics**	*Jackel*
02:10 PM	**Daniel Polani - Competitions to Challenge Artificial Intelligence: from the L-Game to RoboCup**	*Polani*

02:40 PM	**Antoine Marot - Learning to run a power network**	*Marot*
03:10 PM	**Break, poster viewing 3**	
03:10 PM	**AutoDL challenge design and beta tests, Zhengying Liu, Olivier Bousquet, Andre Elisseeff, Isabelle Guyon, Adrien Pavao, Lisheng Sun-Hosoya, and Sebastien Treguer**	*Liu, Treguer*
03:10 PM	**Multi-Agent RL in MalmO (MARLO) Competition, Diego Perez-Liebana Katja Hofmann Sharada Prasanna Mohanty Noburu Kuno Andre Kramer Sam Devlin Raluca D. Gaina Daniel Ionita**	*Mohanty*
03:10 PM	**L2RPN: Learning To Run a Power Network Competition, Antoine Marot, Balthazar Donon, Isabelle Guyon, Benjamin Donnot.**	*Donnot*
03:10 PM	**TrackML, a Particle Physics Tracking Machine Learning Challenge, Jean-Roch Vlimant (Caltech), Vincenzo Innocente, Andreas Salzburger (CERN), Isabelle Guyon (ChaLearn), Sabrina Amrouche, Tobias Golling, Moritz Kiehn (Geneva University),David Rousseau∗, Yet**	*Ustyuzhanin, vlimant*
03:10 PM	**NASA Frontier Development Lab 2018, A. Bella, A. Chopra, W. Fawcett, R. Talebi, D. Angerhausen, A. Berea, N.A. Cabrol, C. Kempes, M. Mascaro**	*Bell*
03:30 PM	**Live competition 2: The driving AI Olympics**	
05:30 PM	**Panel discussion: Opportunities to organize new impactful challenges.**	*Abernethy*

NeurIPS 2018 Competition Track Day 2

Ralf Herbrich, Sergio Escalera

Room 511 CF, Sat Dec 08, 08:00 AM

coming soon

Schedule

08:00 AM	**Morning welcome and introduction**	
08:10 AM	**Facebook project on developing benchmarks of algorithms in realistic settings**	
08:40 AM	**Machine Learning Challenges for Researching Human-Machine Teaming**	
09:10 AM	**Live competition Pommerman: Introduction to Pommerman, the community, and the sponsors**	*Resnick, Eldridge*
09:20 AM	**Live competition Pommerman: Competition battles projected in the room**	*Chauhan, Osogami*
10:25 AM	**Live competition Pommerman: Wrap-up, including a word from the winners**	
10:40 AM	**Break, poster viewing**	
11:00 AM	**Renewal news recommendation platform**	
11:30 AM	**Panel discussion. Design principles and implementation issues**	*Escalera*
12:00 PM	**Lunch break and poster viewing**	
01:00 PM	**Afternoon welcome and announcements (organizers)**	
01:10 PM	**Wild evaluation of chat-bots**	
01:40 PM	**Measuring Progress in Robotics**	
02:10 PM	**Competitions to Challenge Artificial Intelligence: from the L-Game to RoboCup**	
02:40 PM	**Learning to run a power network**	
03:10 PM	**Break and poster viewing**	
03:30 PM	**Live competition The AI Driving Olympics: Introduction to Duckietown and the AI Driving Olympics**	*Paull, Tani, Bowser, Jin, Peron*

03:30 PM	**Live competition The AI Driving Olympics: Live running of top competitor entries**	*Troeshestov, Plante, Bernasconi*
03:50 PM	**Live competition The AI Driving Olympics: Containerization**	*Considine*
03:55 PM	**Live competition The AI Driving Olympics: Reinforcement Learning approaches**	*Golemo*
04:00 PM	**Live competition The AI Driving Olympics: Supervised Learning approaches**	*Díaz, Zilly*
04:05 PM	**Live competition The AI Driving Olympics: Classical Robotics approaches**	*Mehta*
04:10 PM	**Live competition The AI Driving Olympics: AWS and Sagemaker as enablers in the competition**	*Mallya*
04:20 PM	**Live competition The AI Driving Olympics: nuTonomy sponsor placeholder**	*Censi*
04:45 PM	**Live competition The AI Driving Olympics: Final results and awards**	*Tani*
05:00 PM	**Panel discussion: Opportunities to organize new impactful challenges**	*Abernethy*

Wordplay: Reinforcement and Language Learning in Text-based Games

Adam Trischler, Angeliki Lazaridou, Yonatan Bisk, Wendy Tay, Nate Kushman, Marc-Alexandre Côté, Alessandro Sordoni, Daniel Ricks, Tom Zahavy, Hal Daumé III

Room 512 ABEF, Sat Dec 08, 08:00 AM

Video games, via interactive learning environments like ALE [Bellemare et al., 2013], have been fundamental to the development of reinforcement learning algorithms that work on raw video inputs rather than featurized representations. Recent work has shown that text-based games may present a similar opportunity to develop RL algorithms for natural language inputs [Narasimhan et al., 2015, Haroush et al., 2018]. Drawing on insights from both the RL and NLP communities, this workshop will explore this opportunity, considering synergies between text-based and video games as learning environments as well as important differences and pitfalls.

Video games provide infinite worlds of interaction and grounding defined by simple, physics-like dynamics. While it is difficult, if not impossible, to simulate the full and social dynamics of linguistic interaction (see, e.g., work on user simulation and dialogue [Georgila et al., 2006, El Asri et al., 2016]), text-based games nevertheless present complex, interactive simulations that ground language in world and action semantics. Games like Zork [Infocom, 1980] rose to prominence in the age before advanced computer graphics. They use simple language to describe the state of the environment and to report the effects of player actions. Players interact with the environment through text commands that respect a predefined grammar, which, though simplistic, must be discovered in each game. Through sequential decision making, language understanding, and language generation, players work toward goals that may or may not be specified explicitly, and earn rewards (points) at completion or along the way.

Text-based games present a broad spectrum of challenges for learning algorithms. In addition to language understanding, successful play generally requires long-term memory and planning, exploration/experimentation, affordance extraction [Fulda et al., 2017], and common sense. Text games also highlight major open challenges for RL: the action space (text) is combinatorial and compositional, while game states are partially observable, since text is often ambiguous or underspecific. Furthermore, in text games the set of actions that affect the state is not known in advance but must be learned through experimentation, typically informed by prior world/linguistic knowledge.

There has been a host of recent work towards solving text games [Narasimhan et al., 2015, Fulda et al., 2017, Kostka et al., 2017, Zhilin, et al., 2017, Haroush et al., 2018]. Nevertheless, commercial games like Zork remain beyond the capabilities of existing approaches. We argue that addressing even a subset of the aforementioned challenges would represent important progress in machine learning. Agents that solve text-based games may further learn functional properties of language; however, it is unclear what limitations the constraints and simplifications of text games (e.g., on linguistic diversity) impose on agents trained to solve them.

This workshop will highlight research that investigates existing or novel RL techniques for text-based settings, what agents that solve text-based games (might) learn about language, and more generally whether text-based games provide a good testbed for research at the intersection of RL and NLP. The program will feature a collection of invited talks alongside contributed posters and spotlight talks, curated by a committee with broad coverage of the RL and NLP communities. Panel discussions will highlight perspectives of influential researchers from both fields and encourage open dialogue. We will also pose a text-based game challenge several months in advance of the workshop (a similar competition is held annually at the IEEE Conference on Computational Intelligence and Games). This optional component will enable participants to design, train, and test agents in a carefully constructed, interactive text environment. The best-performing agent(s) will be recognized and discussed at the workshop. In addition to the exchange of ideas and the initiation of collaboration, an expected outcome is that text-based games emerge more prominently as a benchmark task to bridge RL and NLP research.

Relevant topics to be addressed at the workshop include (but are not limited to):
- RL in compositional, combinatorial action spaces
- Open RL problems that are especially pernicious in text-based games, like (sub)goal identification and efficient experimentation

- Grounded language understanding
- Online language acquisition
- Affordance extraction (on the fly)
- Language generation and evaluation in goal-oriented settings
- Automatic or crowdsourcing methods for linguistic diversity in simulations
- Use of language to constrain or index RL policies [Andreas et al., 2017]

Schedule

08:30 AM	**Opening Remarks**	*Trischler*
08:40 AM	**Humans and models as embodied dialogue agents in text-based games**	*Weston*
09:20 AM	**Playing Text-Adventure Games with Graph-Based Deep Reinforcement Learning**	*Ammanabrolu*
09:40 AM	**Why Are Words so Dang Hard? Fundamental Challenges in Language-based Games (Nancy Fulda)**	
10:30 AM	**Coffee Break 1**	
11:00 AM	**Towards Solving Text-based Games by Producing Adaptive Action Spaces**	*Tao*
11:20 AM	**How Players Speak to an Intelligent Game Character Using Natural Language Messages**	*Hofmann*
12:00 PM	**Lunch**	
01:20 PM	**BabyAI: First Steps Towards Grounded Language Learning With a Human In the Loop**	*Chevalier-Boisvert*
02:00 PM	**Solving Interactive Fiction Games: A Colossal Adventure**	*Hausknecht, Chen*
02:20 PM	**Harnessing the synergy between natural language and interactive learning**	*Narasimhan*
03:00 PM	**Coffee Break 2**	
03:30 PM	**On the role of text-based games for language learning and RL**	
04:10 PM	**Hierarchical reinforcement learning for composite-task dialogues**	*Li*
04:50 PM	**Introducing "First TextWorld Problems": a text-based game competition**	*Côté*
05:10 PM	**Closing Remarks**	

Abstracts (9):

Abstract 2: **Humans and models as embodied dialogue agents in text-based games in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Weston* 08:40 AM

We describe new work that connects two separate threads of our previous research: (i) situated language learning in text adventure games such as (Bordes et al., AISTATS 2010) and (Weston et al, ICLR 2016); and (ii) non-situated dialogue agents such as in the recent PersonaChat dataset (Zhang et al, ACL 2018). The resulting approach aims to develop embodied agents with personas that can both act and speak, where the situated dialogue involves real language between models and humans that can be grounded within the game.

Abstract 3: **Playing Text-Adventure Games with Graph-Based Deep Reinforcement Learning in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Ammanabrolu* 09:20 AM

Text-based adventure games provide a platform on which to explore reinforcement learning in the context of a combinatorial action space, such as natural language. We present a deep reinforcement learning architecture that represents the game state as a knowledge graph which is learned during exploration. This graph is used to prune the action space, enabling more efficient exploration. The question of which action to take can be reduced to a question-answering task, a form of transfer learning that pre-trains certain parts of our architecture. In experiments using the TextWorld framework, we show that our proposed technique can learn a control policy faster than baseline alternatives.

Abstract 4: **Why Are Words so Dang Hard? Fundamental Challenges in Language-based Games (Nancy Fulda) in Wordplay: Reinforcement and Language Learning in Text-based Games**, 09:40 AM

As inherently linguistic creatures, we tend to think of text as a simple domain: After all, there are only twenty-six letters in the English language, and basic tasks like keyword recognition and part-of-speech tagging have been routinely applied in industry for more than a decade. But words are not a domain in and of themselves. Rather, they function as abstract representations for other types of input, resulting in daunting levels of complexity. This presentation discusses some of the challenges presented by language tasks in general and by text-based games in particular, including partially observable state spaces, compositional and combinatorial action spaces, word-sense disambiguation, consumable rewards, and goal-directed inference.

Abstract 6: **Towards Solving Text-based Games by Producing Adaptive Action Spaces in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Tao* 11:00 AM

To solve a text-based game, an agent needs to formulate valid text commands for a given context and find the one that leads to success. Recent attempts at solving text-based games with deep reinforcement learning have focused on the latter, i.e., learning to act optimally when valid actions are known in advance. In this work, we propose to tackle the first task and train a model that generates the set of all valid commands for a given context. We try three generative models on a dataset generated with Textworld (Côté et al., 2018). The best model can

generate valid commands which were unseen at training and achieve high F1 score on the test set.

Abstract 7: **How Players Speak to an Intelligent Game Character Using Natural Language Messages in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Hofmann* 11:20 AM

AI-driven characters that learn directly from human input are rare in digital games, but recent advances in several fields of machine learning suggests that they may soon be much more feasible to create. This study explores the design space for interacting with such a character through natural language text dialogue. We conducted an observational study with 18 high school students, who played Minecraft alongside a Wizard of Oz prototype of a companion AI character that learned from their actions and inputs. In this paper, we report on an analysis of the 186 natural language messages that players sent to the character, and review key variations in syntax, function and writing style. We find that players' behaviour and language was differentiated by the extent to which they expressed an anthropomorphic view of the AI character and the level of interest that they showed in interacting with it.

Abstract 9: **BabyAI: First Steps Towards Grounded Language Learning With a Human In the Loop in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Chevalier-Boisvert* 01:20 PM

Allowing humans to interactively train artificial agents to understand language instructions is desirable for both practical and scientific reasons, but given the poor data efficiency of the current learning methods, this goal may require substantial research efforts. Here, we introduce the BabyAI research platform to support investigations towards including humans in the loop for grounded language learning. The BabyAI platform comprises an extensible suite of 19 levels of increasing difficulty. The levels gradually lead the agent towards acquiring a combinatorially rich synthetic language which is a proper subset of English. The platform also provides a heuristic expert agent for the purpose of simulating a human teacher. We report baseline results and estimate the amount of human involvement that would be required to train a neural network-based agent on some of the BabyAI levels. We put forward strong evidence that current deep learning methods are not yet sufficiently sample efficient when it comes to learning a language with compositional properties.

Abstract 10: **Solving Interactive Fiction Games: A Colossal Adventure in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Hausknecht, Chen* 02:00 PM

Interactive fiction (IF) games present very different challenges than the vision and control-based games that learning agents have previously excelled at. Solving IF games requires human-like language understanding, commonsense reasoning, planning, and deduction skills. This paper provides a testbed for rapid development of new agents that exhibit these skills by introducing Jericho, a fast, fully-featured interface to fifty-six popular and challenging IF games. We also present initial work towards solving these games in the form of an agent that won the 2018 Text-Based Adventure AI Competition. Finally, we conduct a comprehensive evaluation between NAIL, our agent, and several other IF agents in a richer set of text game environments, and point to directions in which agents can improve. We are optimistic that tools such as Jericho and NAIL will help the community make progress towards language-understanding agents.

Abstract 11: **Harnessing the synergy between natural language and interactive learning in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Narasimhan* 02:20 PM

For most of the statistical ML era, the areas of computational linguistics and reinforcement learning (RL) have been studied separately. With the rise of deep learning, we now have tools that can leverage large amounts of data across multiple modalities. In this talk, I make the case for building holistic AI systems that learn by simultaneously utilizing signals from both language and environmental feedback. While RL has been used in recent work to help understand language, I will demonstrate that language can also help agents learn control policies that generalize over domains. Developing agents that can efficiently harness this synergy between language understanding and policy learning will be crucial for our progress towards stronger AI systems.

Abstract 14: **Hierarchical reinforcement learning for composite-task dialogues in Wordplay: Reinforcement and Language Learning in Text-based Games**, *Li* 04:10 PM

As in many complex text-based scenarios, a conversation can often be decomposed into multiple parts, each taking care of a subtopic or subtask that contributes to the success of the whole dialogue. An example is a travel assistant, which can converse with a user to deal with subtasks like hotel reservation, air ticket purchase, etc. In this talk, we will show how hierarchical deep reinforcement learning can be a useful framework for managing such "composite-task dialogues": (1) more efficient policy optimization with given subtasks; and (2) discovery of dialogue subtasks from corpus in an unsupervised way.

Privacy Preserving Machine Learning

Aurélien Bellet, Adria Gascon, Niki Kilbertus, Olga Ohrimenko, Mariana Raykova, Adrian Weller

Room 512 CDGH, Sat Dec 08, 08:00 AM

[Website](https://ppml-workshop.github.io/ppml/)

Description

This one day workshop focuses on privacy preserving techniques for training, inference, and disclosure in large scale data analysis, both in the distributed and centralized settings. We have observed increasing interest of the ML community in leveraging cryptographic techniques such as Multi-Party Computation (MPC) and Homomorphic Encryption (HE) for privacy preserving training and inference, as well as Differential Privacy (DP) for disclosure. Simultaneously, the systems security and cryptography community has proposed various secure frameworks for ML. We encourage both theory and application-oriented submissions exploring a range of approaches, including:

- secure multi-party computation techniques for ML
- homomorphic encryption techniques for ML
- hardware-based approaches to privacy preserving ML
- centralized and decentralized protocols for learning on encrypted data
- differential privacy: theory, applications, and implementations
- statistical notions of privacy including relaxations of differential privacy
- empirical and theoretical comparisons between different notions of

privacy
- trade-offs between privacy and utility

We think it will be very valuable to have a forum to unify different perspectives and start a discussion about the relative merits of each approach. The workshop will also serve as a venue for networking people from different communities interested in this problem, and hopefully foster fruitful long-term collaboration.

Schedule

08:30 AM	**Welcom and introduction**	
08:50 AM	**Invited talk 1: Scalable PATE and the Secret Sharer**	*Goodfellow*
09:40 AM	**Invited talk 2: Machine Learning and Cryptography: Challenges and Opportunities**	*Goldwasser*
10:30 AM	**Coffee Break 1**	
11:00 AM	**Contributed talk 1: Privacy Amplification by Iteration**	*Feldman*
11:15 AM	**Contributed talk 2: Subsampled Renyi Differential Privacy and Analytical Moments Accountant**	*Wang*
11:30 AM	**Contributed talk 3: The Power of The Hybrid Model for Mean Estimation**	*Dubey*
11:45 AM	**Contributed talk 4: Amplification by Shuffling: From Local to Central Differential Privacy via Anonymity**	*Erlingsson*
12:00 PM	**Lunch Break**	
01:30 PM	**Invited talk 3: Challenges in the Privacy-Preserving Analysis of Structured Data**	*Chaudhuri*
02:20 PM	**Invited talk 4: Models for private data analysis of distributed data**	*Smith*
03:10 PM	**Coffee Break 2**	
03:30 PM	**Contributed talk 5: DP-MAC: The Differentially Private Method of Auxiliary Coordinates for Deep Learning**	*Harder*
03:45 PM	**Contributed talk 6: Slalom: Fast, Verifiable and Private Execution of Neural Networks in Trusted Hardware**	*Tramer*
04:00 PM	**Contributed talk 7: Secure Two Party Distribution Testing**	*Shekel Nosatzki*
04:15 PM	**Contributed talk 8: Private Machine Learning in TensorFlow using Secure Computation**	*Dahl*
04:30 PM	**Spotlight talks**	
05:15 PM	**Poster Session**	*Schoppmann, Yu, Chen, Dick, Joye, Zhang, Harder, Saarikivi, Ryffel, Long, JOURDAN, Wang, Marcedone, Shekel Nosatzki, Dubey, Koskela, Bloem, Korolova, Bertran, Chen, Andrew, Martinez, Kulkarni, Passerat-Palmbach, Sapiro, Roy Chowdhury*
06:15 PM	**Wrap up**	

Abstracts (12):

Abstract 3: **Invited talk 2: Machine Learning and Cryptography: Challenges and Opportunities in Privacy Preserving Machine Learning**, *Goldwasser* 09:40 AM

At the mid eighties researchers in computational learning theory pointed the way to examples of hard learning tasks such as learning parity with noise (LPN) and learning with errors (LWE) which have been extremely useful for building sophisticated cryptographic primitives such as homomorphic encryption which are unbreakable if LPN and LWE are hard to learn.

Today, with the rise of machine learning algorithms that use large amounts of data to come up with procedures which have the potential to replace human decision processes, cryptography, in turn, stands to provide machine learning, tools for keeping data private during both training and inference phases of ML and to provide methods to verify adherence of models with data. These promise to be important steps in ensuring the safe transition of power from human to algorithmic decision making.

Abstract 5: **Contributed talk 1: Privacy Amplification by Iteration in Privacy Preserving Machine Learning**, *Feldman* 11:00 AM

Many commonly used learning algorithms work by iteratively updating an intermediate solution using one or a few data points in each iteration. Analysis of differential privacy for such algorithms often involves ensuring privacy of each step and then reasoning about the cumulative privacy cost of the algorithm. This is enabled by composition theorems for differential privacy that allow releasing of all the intermediate results. In this work, we demonstrate that for contractive iterations, not releasing the intermediate results strongly amplifies the privacy guarantees.

We describe several applications of this new analysis technique to solving convex optimization problems via noisy stochastic gradient descent. For example, we demonstrate that a relatively small number of non-private data points from the same distribution can be used to close

the gap between private and non-private convex optimization. In addition, we demonstrate that we can achieve guarantees similar to those obtainable using the privacy-amplification-by-sampling technique in several natural settings where that technique cannot be applied.

Abstract 6: **Contributed talk 2: Subsampled Renyi Differential Privacy and Analytical Moments Accountant in Privacy Preserving Machine Learning**, *Wang* 11:15 AM

We study the problem of subsampling in differential privacy (DP), a question that is the centerpiece behind many successful differentially private machine learning algorithms. Specifically, we provide a tight upper bound on the Renyi Differential Privacy (RDP) parameters for algorithms that: (1) subsample the dataset, and then (2) applies a randomized mechanism M to the subsample, in terms of the RDP parameters of M and the subsampling probability parameter. Our results generalize the moments accounting technique, developed by Abadi et al. [CCS'16] for the Gaussian mechanism, to any subsampled RDP mechanism.

Abstract 7: **Contributed talk 3: The Power of The Hybrid Model for Mean Estimation in Privacy Preserving Machine Learning**, *Dubey* 11:30 AM

In this work we explore the power of the hybrid model of differential privacy (DP) proposed in~\cite{blender}, where some users desire the guarantees of the local model of DP and others are content with receiving the trusted curator model guarantees. In particular, we study the accuracy of mean estimation algorithms for arbitrary distributions in bounded support. We show that a hybrid mechanism which combines the sample mean estimates obtained from the two groups in an optimally weighted convex combination performs a constant factor better for a wide range of sample sizes than natural benchmarks. We analyze how this improvement factor is parameterized by the problem setting and how it varies with sample size.

Abstract 8: **Contributed talk 4: Amplification by Shuffling: From Local to Central Differential Privacy via Anonymity in Privacy Preserving Machine Learning**, *Erlingsson* 11:45 AM

Sensitive statistics are often collected across sets of users, with repeated collection of reports done over time. For example, trends in users' private preferences or software usage may be monitored via such reports. We study the collection of such statistics in the local differential privacy (LDP) model, when each user's value may change only a small number of times. We describe an algorithm whose privacy cost is polylogarithmic in the number of times the statistic is collected.

More fundamentally---by building on anonymity of the users' reports---we also demonstrate how the privacy cost of our LDP algorithm can actually be much lower when viewed in the central model of differential privacy. We show, via a new and general technique, that any permutation-invariant algorithm satisfying ε-local differential privacy will satisfy $(O(\varepsilon \sqrt{\log \frac 1 \delta} / \sqrt{n}), \delta)$-central differential privacy. In the process, we clarify how the high noise and $\sqrt{n}$ overhead of LDP protocols is a consequence of them being significantly more private in the central model. As a final, practical corollary, our results also imply that industrial deployments of LDP mechanism may have much lower privacy cost than their advertised ε would indicate---at least if reports are anonymized.

Abstract 10: **Invited talk 3: Challenges in the Privacy-Preserving Analysis of Structured Data in Privacy Preserving Machine Learning**, *Chaudhuri* 01:30 PM

Much data analysis today is done on sensitive data, and particular privacy challenges arise when this data is sensitive and structured. In this talk I will describe two such challenges in the privacy-preserving analysis of complex, structured data that we have been working on in my group.

The first is continual release of graph statistics in an online manner from an expanding graph, which is motivated by a problem in HIV epidemiology. Even though node differentially private release of graph statistics is highly challenging, here we will describe how we can get a differentially private solution for this problem that performs better than the natural sequential composition baseline.

Next, I will talk about analysis of sensitive structured, correlated data, while still preserving the privacy of events in the data. Differential privacy does not adequately address privacy issues in this kind of data, and hence will look at a form of inferential privacy, called Pufferfish, that is more appropriate. We will provide mechanisms, establish their composition properties, and finally evaluate them on real data on physical activity measurements across time.

Abstract 11: **Invited talk 4: Models for private data analysis of distributed data in Privacy Preserving Machine Learning**, *Smith* 02:20 PM

This talk will present a partial survey of the proposed (and implemented) models for private analysis of distributed data, together with some new results.

Abstract 13: **Contributed talk 5: DP-MAC: The Differentially Private Method of Auxiliary Coordinates for Deep Learning in Privacy Preserving Machine Learning**, *Harder* 03:30 PM

Developing a differentially private deep learning algorithm is challenging, due to the difficulty in analyzing the sensitivity of objective functions that are typically used to train deep neural networks. Many existing methods resort to the stochastic gradient descent algorithm and apply a pre-defined sensitivity to the gradients for privatizing weights. However, their slow convergence typically yields a high cumulative privacy loss. Here, we take a different route by employing the method of auxiliary coordinates, which allows us to independently update the weights per layer by optimizing a per-layer objective function. This objective function can be well approximated by a low-order Taylor's expansion, in which sensitivity analysis becomes tractable. We perturb the coefficients of the expansion for privacy, which we optimize using more advanced optimization routines than SGD for faster convergence. We empirically show that our algorithm provides a decent trained model quality under a modest privacy budget.

Abstract 14: **Contributed talk 6: Slalom: Fast, Verifiable and Private Execution of Neural Networks in Trusted Hardware in Privacy Preserving Machine Learning**, *Tramer* 03:45 PM

As Machine Learning (ML) gets applied to security-critical or sensitive domains, there is a growing need for integrity and privacy for outsourced ML computations. A pragmatic solution comes from Trusted Execution Environments (TEEs), which use hardware and software protections to isolate sensitive computations from the untrusted software stack.

However, these isolation guarantees come at a price in performance, compared to untrusted alternatives. This paper initiates the study of high performance execution of Deep Neural Networks (DNNs) in TEEs by efficiently partitioning DNN computations between trusted and untrusted devices. Building upon an efficient outsourcing scheme for matrix multiplication, we propose Slalom, a framework that securely delegates execution of all linear layers in a DNN from a TEE (e.g., Intel SGX or Sanctum) to a faster, yet untrusted, co-located processor. We evaluate Slalom by executing DNNs in an Intel SGX enclave, which selectively delegates work to an untrusted GPU. For two canonical DNNs, VGG16 and MobileNet, we obtain 20× and 6× increases in throughput for verifiable inference, and 11× and 4× for verifiable and private inference.

Abstract 15: **Contributed talk 7: Secure Two Party Distribution Testing in Privacy Preserving Machine Learning**, *Shekel Nosatzki* 04:00 PM

We study the problem of discrete distribution testing in the two-party setting. For example, in the standard closeness testing problem, Alice and Bob each have t samples from, respectively, distributions a and b over [n], and they need to test whether a=b or a,b are ε-far (in the ■1 distance) for some fixed ε>0. This is in contrast to the well-studied one-party case, where the tester has unrestricted access to samples of both distributions, for which optimal bounds are known for a number of variations. Despite being a natural constraint in applications, the two-party setting has evaded attention so far.

We address two fundamental aspects of the two-party setting: 1) what is the communication complexity, and 2) can it be accomplished securely, without Alice and Bob learning extra information about each other's input. Besides closeness testing, we also study the independence testing problem, where Alice and Bob have t samples from distributions a and b respectively, which may be correlated; the question is whether a,b are independent of ε-far from being independent.

Our contribution is three-fold:

■ Communication: we show how to gain communication efficiency as we have more samples, beyond the information-theoretic bound on t. Furthermore, the gain is polynomially better than what one may obtain by adapting one-party algorithms.

For the closeness testing, our protocol has communication s=Θ_ε(n^2/t^2) as long as t is at least the information-theoretic minimum number of samples. For the independence testing over domain [n]×[m], where n≥m, we obtain s=O_ε(n^2m/t^2+nm/t+m^1/2).

■ Security: we define the concept of secure distribution testing and argue that it must leak at least some minimal information when the promise is not satisfied. We then provide secure versions of the above protocols with an overhead that is only polynomial in the security parameter.

■ Lower bounds: we prove tightness of our trade-off for the closeness testing, as well as that the independence testing requires tight Ω(m^1/2) communication for unbounded number of samples. These lower bounds are of independent interest as, to the best of our knowledge, these are the first 2-party communication lower bounds for testing problems, where the inputs represent a set of i.i.d. samples.

Abstract 16: **Contributed talk 8: Private Machine Learning in TensorFlow using Secure Computation in Privacy Preserving Machine Learning**, *Dahl* 04:15 PM

We present a framework for experimenting with secure multi-party computation directly in TensorFlow. By doing so we benefit from several properties valuable to both researchers and practitioners, including tight integration with ordinary machine learning processes, existing optimizations for distributed computation in TensorFlow, high-level abstractions for expressing complex algorithms and protocols, and an expanded set of familiar tooling. We give an open source implementation of a state-of-the-art protocol and report on concrete benchmarks using typical models from private machine learning.

Abstract 17: **Spotlight talks in Privacy Preserving Machine Learning**, 04:30 PM

1. [Cynthia Dwork and Vitaly Feldman] Privacy-preserving Prediction (#05)
2. [Garrett Bernstein and Daniel Sheldon] Differentially Private Bayesian Inference for Exponential Families (#03)
3. [Di Wang, Adam Smith and Jinhui Xu] High Dimensional Sparse Linear Regression under Local Differential Privacy: Power and Limitations (#11)
4. [Ashwin Machanavajjhala and Kamalika Chaudhuri] Capacity Bounded Differential Privacy (#25)
5. [Aurélien Bellet, Rachid Guerraoui and Hadrien Hendrikx] Who started this gossip? Differentially private rumor spreading (#26)
6. [Antti Koskela and Antti Honkela] Learning rate adaptation for differentially private stochastic gradient descent (#38)
7. [Kareem Amin, Travis Dick, Alex Kulesza, Andres Medina and Sergei Vassilvitskii] Private Covariance Estimation via Iterative Eigenvector Sampling (#45)
8. [Kareem Amin, Alex Kulesza, Andres Munoz Medina and Sergei Vassilvitskii] Bias Variance Trade-off in Differential Privacy (#53)
9. [Nicolas Loizou, Peter Richtarik, Filip Hanzely, Jakub Konecny and Dmitry Grishchenko] A Privacy Preserving Randomized Gossip Algorithm via Controlled Noise Insertion (#57)
10. [Brendan McMahan and Galen Andrew] A General Approach to Adding Differential Privacy to Iterative Training Procedures (#62)
11. [Da Yu, Huishuai Zhang and Wei Chen] Improving the Gradient Perturbation Approach for Differentially Private Optimization (#70)
12. [Alexandra Schofield, Aaron Schein, Zhiwei Steven Wu and Hanna Wallach] A Variational Inference Approach for Locally PrivateInference of Poisson Factorization Models (#63)
13. [Judy Hoffman, Mehryar Mohri and Ningshan Zhang] Algorithms and Theory for Multiple-Source Adaptation (#52)
14. [Martin Bertran, Natalia Martinez, Afroditi Papadaki, Qiang Qiu, Miguel Rodrigues and Guillermo Sapiro] Learning Representations for Utility and Privacy: An Information-Theoretic Based Approach (#15)
15. [Fabrice Benhamouda and Marc Joye] How to Profile Privacy-Conscious Users in Recommender Systems (#32)
16. [Koen Lennart van der Veen, Ruben Seggers, Peter Bloem and Giorgio Patrini] Three Tools for Practical Differential Privacy (#29)
17. [Vasyl Pihur, Aleksandra Korolova, Frederick Liu, Subhash Sankuratripati, Moti Yung, Dachuan Huang and Ruogu Zeng] Differentially Private "Draw and Discard" Machine Learning (#60)
18. [Hsin-Pai Cheng, Patrick Yu, Haojing Hu, Feng Yan, Shiyu Li, Hai Li and Yiran Chen] LEASGD: an Efficient and Privacy-Preserving Decentralized Algorithm for Distributed Learning (#01)
19. [Joshua Allen, Bolin Ding, Janardhan Kulkarni, Harsha Nori, Olga

Ohrimenko and Sergey Yekhanin] An Algorithmic Framework For Differentially Private Data Analysis on Trusted Processors (#12)
20. [Antoine Boutet, Théo Jourdan and Carole Frindel] Toward privacy in IoT mobile devices for activity recognition (#13)
21. [Roshan Dathathri, Olli Saarikivi, Hao Chen, Kim Laine, Kristin Lauter, Saeed Maleki, Madanlal Musuvathi and Todd Mytkowicz] CHET: Compiler and Runtime for Homomorphic Evaluation of Tensor Programs (#14)
22. [Theo Ryffel, Andrew Trask, Morten Dahl, Bobby Wagner, Jason Mancuso, Daniel Rueckert and Jonathan Passerat-Palmbach] A generic framework for privacy preserving deep learning (#43)
23. [Valerie Chen, Valerio Pastro and Mariana Raykova] Secure Computation for Machine Learning With SPDZ (#44)
24. [Phillipp Schoppmann, Adria Gascon, Mariana Raykova and Benny Pinkas] Make Some ROOM for the Zeros: Data Sparsity in Secure Distributed Machine Learning (#51)
25. [Siddharth Garg, Zahra Ghodsi, Carmit Hazay, Yuval Ishai, Antonio Mercedone and Muthuramakrishnan Venkitasubramaniam] Oursourcing Private Machine Learning via Lightweight Secure Arithmetic Computation (#66)
26. [Hao Chen, Ilaria Chillotti, Oxana Poburinnaya, Ilya Razenshteyn and M. Sadegh Riazi] Scaling Up Secure Nearest Neighbor Search (#33)
27. [Yunhui Long, Vincent Bindschaedler and Carl Gunter] Towards Measuring Membership Privacy (#09) [video]

Medical Imaging meets NIPS

Ender Konukoglu, Ben Glocker, Hervé Lombaert, Marleen de Bruijne

Room 513 ABC, Sat Dec 08, 08:00 AM

Medical imaging and radiology are facing a major crisis with an ever-increasing complexity and volume of data and immense economic pressure. With the current advances in imaging technologies and their widespread use, interpretation of medical images pushes human abilities to the limit with the risk of missing critical patterns of disease. Machine learning has emerged as a key technology for developing novel tools in computer aided diagnosis, therapy and intervention. Still, progress is slow compared to other fields of visual recognition, which is mainly due to the domain complexity and constraints in clinical applications, i.e. robustness, high accuracy and reliability.

“Medical Imaging meets NIPS” aims to bring researchers together from the medical imaging and machine learning communities to discuss the major challenges in the field and opportunities for research and novel applications. The proposed event will be the continuation of a successful workshop organized in NIPS 2017 (https://sites.google.com/view/med-nips-2017). It will feature a series of invited speakers from academia, medical sciences and industry to give an overview of recent technological advances and remaining major challenges.

Different from last year and based on feedback from participants, we propose to implement two novelties.
1. The workshop will accept paper submissions and have oral presentations with a format that aims to foster in depth discussions of a few selected articles. We plan to implement a Program Committee who will be responsible for reviewing articles and initiating discussions. The abstract track organized last year has brought a significant number of submission and has clearly demonstrated an appetite for more.
2. Along the workshop, we will host a challenge on outlier detection in brain Magnetic Resonance Imaging (MRI), which is one of the main applications of advanced unsupervised learning algorithms and generative models in medical imaging. The challenge will highlight a problem where the machine learning community can have a huge impact. To facilitate the challenge and potential further research, we provide necessary pre-processed datasets to simplify the use of medical imaging data and lower data-related entry barrier. Data collection for this challenge is finalized and ethical approval for data sharing is in place. We plan to open the challenge as soon as acceptance of the workshop is confirmed.

Schedule

08:45 AM	**Welcome**	*Konukoglu, Glocker, Lombaert, de Bruijne*
09:00 AM	**Making the Case for using more Inductive Bias in Deep Learning**	*Welling*
09:45 AM	**The U-net does its job – so what next?**	*Ronneberger*
10:30 AM	**Coffee Break**	
11:00 AM	**Oral session I**	*Adler, Jalal, Cheng*
12:00 PM	**Lunch**	
01:00 PM	**Tackling the challenges of next generation healthcare**	*Roth*
01:45 PM	**Oral session II**	*van de Leemput, Dalca, Gopinath*
02:45 PM	**Poster session**	*Zeng, Tahaei, Chen, Meister, Shah, Gupta, Jalal, Arvaniti, Zimmerer, Kamnitsas, Ballester, Braman, Kumar, van de Leemput, Qadir, Kervadec, Akrout, Tousignant, Ng, Mehta, Monteiro, Basu, Adler, Dalca, Peng, Han, Li, Gopinath, Cheng, Georgescu, Quach, Sarma, Van Veen*
04:15 PM	**Is your machine learning method solving a real clinical problem?**	*Arbel*
05:00 PM	**Oral session III**	*Braman, Tousignant, Ng*
06:00 PM	**Closing remarks**	*Konukoglu, Glocker, Lombaert, de Bruijne*

Abstracts (2):

Abstract 3: **The U-net does its job – so what next? in Medical Imaging meets NIPS**, *Ronneberger* 09:45 AM

U-net based architectures have demonstrated very high performance in a wide range of medical image segmentation tasks, but a powerful segmentation architecture alone is only one part of building clinically

applicable tools. In my talk I'll present three projects from the DeepMind Health Research team that address these challenges.
The first project, a collaboration with University College London Hospital, deals with the challenging task of the precise segmentation of radiosensitive head and neck anatomy in CT scans, an essential input for radiotherapy planning [1]. With a 3D U-net we reach a performance similar to human experts on the majority of anatomical classes. Beside some minor architectural adaptations, e.g. to tackle the large imbalance of foreground to background voxels, a substantial focus of the project was in generating a high-quality test set [2] where each scan was manually segmented by two independent experts. Furthermore we introduced a new surface based performance metric, the surface DSC [3], designed to be a better proxy for the expected performance in a real-world radiotherapy setting than existing metrics.
The second project, together with Moorfields Eye Hospital, developed a system that analyses 3D OCT (optical coherence tomography) eye scans to provide referral decisions for patients [4]. The performance was on par with world experts with over 20 years experience. We use two network ensembles to decouple the variations induced by the imaging system from the patient-to-patient variations. The first ensemble of 3D U-nets creates clinically interpretable device-independent tissue map hypotheses; the second (3D dense-net based) ensemble maps the tissue map hypotheses to the diagnoses and referral recommendation. Adaptation to a new scanning device type only needed sparse manual segmentations on 152 scans, while the diagnosis model (trained with 14,884 OCT scans) could be reused without changes.
The third project deals with the segmentation of ambiguous images [5]. This is of particular relevance in medical imaging where ambiguities can often not be resolved from the image context alone. We propose a combination of a U-net with a conditional variational autoencoder that is capable of efficiently producing an unlimited number of plausible segmentation map hypotheses for a given ambiguous image. We show that each hypothesis provides an overall consistent segmentation, and that the probabilities of these hypotheses are well calibrated.

[1] Nikolov et al. (2018) "Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy" (soon available on ArXiv)
[2] Dataset will be soon available at https://github.com/deepmind/tcia-ct-scan-dataset
[3] Implementation available at https://github.com/deepmind/surface-distance
[4] De Fauw, et al. (2018) "Clinically applicable deep learning for diagnosis and referral in retinal disease" Nature Medicine (in press). https://doi.org/10.1038/s41591-018-0107-6 (fulltext available from https://deepmind.com/blog/moorfields-major-milestone/)
[5] Kohl, et al. (2018) "A Probabilistic U-Net for Segmentation of Ambiguous Images". NIPS 2018 (accepted). Preprint available at https://arxiv.org/abs/1806.05034

Abstract 7: **Tackling the challenges of next generation healthcare in Medical Imaging meets NIPS**, *Roth* 01:00 PM

TBD

Interpretability and Robustness in Audio, Speech, and Language

Mirco Ravanelli, Dmitriy Serdyuk, Ehsan Variani, Bhuvana Ramabhadran

Room 513DEF, Sat Dec 08, 08:00 AM

Domains of natural and spoken language processing have a rich history deeply rooted in information theory, statistics, digital signal processing and machine learning. With the rapid rise of deep learning (“deep learning revolution”), many of these systematic approaches have been replaced by variants of deep neural methods, that often achieve unprecedented performance levels in many fields. With more and more of the spoken language processing pipeline being replaced by sophisticated neural layers, feature extraction, adaptation, noise robustness are learnt inherently within the network. More recently, end-to-end frameworks that learn a mapping from speech (audio) to target labels (words, phones, graphemes, sub-word units, etc.) are becoming increasingly popular across the board in speech processing in tasks ranging from speech recognition, speaker identification, language/dialect identification, multilingual speech processing, code switching, natural language processing, speech synthesis and much much more.

A key aspect behind the success of deep learning lies in the discovered low and high-level representations, that can potentially capture relevant underlying structure in the training data. In the NLP domain, for instance, researchers have mapped word and sentence embeddings to semantic and syntactic similarity and argued that the models capture latent representations of meaning. Nevertheless, some recent works on adversarial examples have shown that it is possible to easily fool a neural network (such as a speech recognizer or a speaker verification system) by just adding a small amount of specially constructed noise. Such a remarkable sensibility towards adversarial attacks highlights how superficial the discovered representations could be, rising crucial concerns on the actual robustness, security, and interpretability of modern deep neural networks. This weakness naturally leads researchers to ask very crucial questions on what these models are really learning, how we can interpret what they have learned, and how the representations provided by current neural networks can be revealed or explained in a fashion that modeling power can be enhanced further. These open questions have recently raised the interest towards interpretability of deep models, as witness by the numerous works recently published on this topic in all the major machine learning conferences. Moreover, some workshops at NIPS 2016, NIPS 2017 and Interspeech 2017 have promoted research and discussion around this important issue.
With our initiative, we wish to further foster some progresses on interpretability and robustness of modern deep learning techniques, with a particular focus on audio, speech and NLP technologies. The workshop will also analyze the connection between deep learning and models developed earlier for machine learning, linguistic analysis, signal processing, and speech recognition. This way we hope to encourage a discussion amongst experts and practitioners in these areas with the expectation of understanding these models better and allowing to build upon the existing collective expertise.

The workshop will feature invited talks, panel discussions, as well as oral and poster contributed presentations. We welcome papers that specifically address one or more of the leading questions listed below:
1. Is there a theoretical/linguistic motivation/analysis that can explain how nets encapsulate the structure of the training data it learns from?
2. Does the visualization of this information (MDS, t-SNE) offer any

insights to creating a better model?
3. How can we design more powerful networks with simpler architectures?
4. How can we can exploit adversarial examples to improve the system robustness?
5. Do alternative methods offer any complimentary modeling power to what the networks can memorize?
6. Can we explain the path of inference?
7. How do we analyze data requirements for a given model? How does multilingual data improves learning power?

Schedule

08:45 AM	**Workshop Opening**	*Ravanelli, Serdyuk, Variani, Ramabhadran*
09:00 AM	**Rich Caruana, "Friends Don't Let Friends Deploy Black-Box Models: The Importance of Intelligibility in Machine Learning"**	*Caruana*
09:30 AM	**Jason Yosinski, "Good and bad assumptions in model design and interpretability"**	*Yosinski*
10:00 AM	**Brandon Carter, "Local and global model interpretability via backward selection and clustering"**	*Carter*
10:15 AM	**Andreas Krug, "Neuron Activation Profiles for Interpreting Convolutional Speech Recognition Models"**	*Krug*
10:30 AM	**Coffee break + posters 1**	*Myer, Hsu, Li, Dinculescu, Schönherr, Hosseini-Asl, Seto, Parker Jones, Sheikh, Manzini, Belinkov, Durrani, Amini, Hansen, Shalev, Shin, Smolensky, Fan, Zhu, Eghbalzadeh, Baer, Jimenez, Santos, Kremer, McDermott, Krug, Fuchs, Tang, Carter, Gifford, Zeyer, Merboldt, Pillutla, Lee, Parcollet, Firat, Bhattacharya, ALAM, Ravanelli*
11:00 AM	**Hynek Hermansky, "Learning - not just for machines anymore"**	*Hermansky*
11:30 AM	**Michiel Bacchiani, "Observations in Joint Learning of Features and Classifiers for Speech and Language"**	*Bacchiani*
12:00 PM	**Mirco Ravanelli, "Interpretable convolutional filters with SincNet"**	*Ravanelli*
12:15 PM	**Hamid Eghbal-zadeh, "Deep Within-Class Covariance Analysis for Robust Deep Audio Representation Learning"**	*Eghbalzadeh*
12:30 PM	**Lunch Break**	
01:30 PM	**Ralf Schlüter, "Automatic Speech Recognition Architectures: from HMM to End-to-End Modeling"**	*Schlüter*
02:00 PM	**Erik McDermott, "A Deep Generative Acoustic Model for Compositional Automatic Speech Recognition"**	*McDermott*
02:15 PM	**Jamin Shin, "Interpreting Word Embeddings with Eigenvector Analysis"**	*Shin*
02:30 PM	**Jan Kremer, "On the Inductive Bias of Word-Character-Level Multi-Task Learning for Speech Recognition"**	*Kremer*
02:45 PM	**Coffee break + posters 2**	*Kremer, McDermott, Carter, Zeyer, Krug, Liang, Lee, Basaj, Jimenez, Fan, Bhattacharya, Fuchs, Gifford, Lugosch, Firat, Baer, ALAM, Shin, Ravanelli, Smolensky, Zhu, Eghbalzadeh, Seto, Sheikh, Santos, Belinkov, Durrani, Parker Jones, Tang, Merboldt, Parcollet, Hsu, Pillutla, Hosseini-Asl, Dinculescu, Amini, Zhang, Cheng, Tapp*
03:30 PM	**Mike Schuster, "Learning from the move to neural machine translation at Google"**	*Schuster*
04:00 PM	**Alexander Rush, "Interprebility in Text Generation"**	*Rush*
04:30 PM	**Shuai Tang, "Learning Distributed Representations of Symbolic Structure Using Binding and Unbinding Operations"**	*Tang*
04:45 PM	**Paul Pu Liang, "Learning Robust Joint Representations for Multimodal Sentiment Analysis"**	*Shin*

05:00 PM	**Jason Eisner, "BiLSTM-FSTs and Neural FSTs"**	*Eisner*
05:30 PM	**Panel Discussion**	*Caruana, Schuster, Schlüter, Hermansky, De Mori, Bengio, Bacchiani, Eisner*

Abstracts (20):

Abstract 2: **Rich Caruana, "Friends Don't Let Friends Deploy Black-Box Models: The Importance of Intelligibility in Machine Learning" in Interpretability and Robustness in Audio, Speech, and Language**, *Caruana* 09:00 AM

In machine learning often a tradeoff must be made between accuracy and intelligibility: the most accurate models (deep nets, boosted trees and random forests) usually are not very intelligible, and the most intelligible models (logistic regression, small trees and decision lists) usually are less accurate. This tradeoff limits the accuracy of models that can be safely deployed in mission-critical applications such as healthcare where being able to understand, validate, edit, and ultimately trust a learned model is important. In this talk, I'll present a case study where intelligibility is critical to uncover surprising patterns in the data that would have made deploying a black-box model risky. I'll also show how distillation with intelligible models can be used to understand what is learned inside a black-box model such as a deep nets, and show a movie of what a deep net learns as it trains and then begins to overfit.

Abstract 3: **Jason Yosinski, "Good and bad assumptions in model design and interpretability" in Interpretability and Robustness in Audio, Speech, and Language**, *Yosinski* 09:30 AM

The seduction of large neural nets is that one simply has to throw input data into a big network and magic comes out the other end. If the output is not magic enough, just add more layers. This simple approach works just well enough that it can lure us into a few bad assumptions, which we'll discuss in this talk. One is that learning everything end-to-end is always best. We'll look at an example where it isn't. Another is that careful manual architecture design is useless because either one big stack of layers will work just fine, or
if it doesn't, we should just give up and use random architecture search and a bunch of computers. But perhaps we just need better tools and mental models to analyze the architectures we're building; in this talk we'll talk about one simple such tool. A final assumption is that as our models become large, they become inscrutable. This may turn out to be true for large models, but attempts at understanding persist,
and in this talk, we'll look at how the assumptions we put into our methods of interpretability color the results.

Abstract 4: **Brandon Carter, "Local and global model interpretability via backward selection and clustering" in Interpretability and Robustness in Audio, Speech, and Language**, *Carter* 10:00 AM

Local explanation frameworks aim to rationalize particular decisions made by a black-box prediction model. Existing techniques are often restricted to a specific type of predictor or based on input saliency, which may be undesirably sensitive to factors unrelated to the model's decision-making process. We instead propose sufficient input subsets that identify minimal subsets of features whose observed values alone suffice for the same decision to be reached, even if all other input feature values are missing. General principles that globally govern a model's decision-making can also be revealed by searching for clusters of such input patterns across many data points. Our approach is conceptually straightforward, entirely model-agnostic, simply implemented using instance-wise backward selection, and able to produce more concise rationales than existing techniques. We demonstrate the utility of our interpretation method on neural network models trained on text and image data.

Abstract 5: **Andreas Krug, "Neuron Activation Profiles for Interpreting Convolutional Speech Recognition Models" in Interpretability and Robustness in Audio, Speech, and Language**, *Krug* 10:15 AM

The increasing complexity of deep Artificial Neural Networks (ANNs) allows to solve complex tasks in various applications. This comes with less understanding of decision processes in ANNs. Therefore, introspection techniques have been proposed to interpret how the network accomplishes its task. Those methods mostly visualize their results in the input domain and often only process single samples. For images, highlighting important features or creating inputs which activate certain neurons is intuitively interpretable. The same introspection for speech is much harder to interpret. In this paper, we propose an alternative method which analyzes neuron activations for whole data sets. Its generality allows application to complex data like speech. We introduce time-independent Neuron Activation Profiles (NAPs) as characteristic network responses to certain groups of inputs. By clustering those time-independent NAPs, we reveal that layers are specific to certain groups. We demonstrate our method for a fully-convolutional speech recognizer. There, we investigate whether phonemes are implicitly learned as an intermediate representation for predicting graphemes. We show that our method reveals, which layers encode phonemes and graphemes and that similarities between phonetic categories are reflected in the clustering of time-independent NAPs.
Keywords: introspection, speech recognition, phoneme representation, grapheme representation, convolutional neural networks

Abstract 6: **Coffee break + posters 1 in Interpretability and Robustness in Audio, Speech, and Language**, *Myer, Hsu, Li, Dinculescu, Schönherr, Hosseini-Asl, Seto, Parker Jones, Sheikh, Manzini, Belinkov, Durrani, Amini, Hansen, Shalev, Shin, Smolensky, Fan, Zhu, Eghbalzadeh, Baer, Jimenez, Santos, Kremer, McDermott, Krug, Fuchs, Tang, Carter, Gifford, Zeyer, Merboldt, Pillutla, Lee, Parcollet, Firat, Bhattacharya, ALAM, Ravanelli* 10:30 AM

Jamin Shin, Andrea Madotto, Pascale Fung, "Interpreting Word Embeddings with Eigenvector Analysis"
Mirco Ravanelli, Yoshua Bengio, "Interpretable Convolutional Filters with SincNet"
Shuai Tang, Paul Smolensky, Virginia R. de Sa, "Learning Distributed Representations of Symbolic Structure Using Binding and Unbinding Operations"
Lisa Fan, Dong Yu, Lu Wang, "Robust Neural Abstractive Summarization Systems and Evaluation against Adversarial Information"
Zining Zhu, Jekaterina Novikova, Frank Rudzicz, "Semi-supervised classification by reaching consensus among modalities"
Hamid Eghbal-zadeh, Matthias Dorfer, Gerhard Widmer, "Deep Within-Class Covariance Analysis for Robust Deep Audio Representation Learning"
Benjamin Baer, Skyler Seto, Martin T. Wells, "Interpreting Word

Embeddings with Generalized Low Rank Models"
Abelino Jimenez, Benjamin Elizalde, Bhiksha Raj, "Sound event classification using ontology-based neural netowrks"
Hai Pham, Paul Pu Liang, Thomas Manzini, Louis-Philippe Morency, Barnabas Poczos, "Found in Translation: Learning Robust Joint Representations by Cyclic Translations Between Modalities"
Sri Harsha Dumpala, Imran Sheikh, Rupayan Chakraborty, Sunil Kumar Kopparapu, "Cycle-Consistent GAN Front-end to Improve ASR Robustness to Perturbed Speech"
Joao Felipe Santos, Tiago H. Falk, "Investigating the effect of residual and highway connections in speech enhancement models"
Jan Kremer, Lasse Borgholt, Lars Maaløe, "On the Inductive Bias of Word-Character-Level Multi-Task Learning for Speech Recognition"
Erik McDermott, "A Deep Generative Acoustic Model for Compositional Automatic Speech Recognition"
Andreas Krug, René Knaebel, Sebastian Stober, "Neuron Activation Profiles for Interpreting Convolutional Speech Recognition Models"
Anthony Bau, Yonatan Belinkov, Hassan Sajjad, Nadir Durrani, Fahim Dalvi, James Glass, "Identifying and Controlling Important Neurons in Neural Machine Translation"
Oiwi Parker Jones, Brendan Shillingford, "Composing RNNs and FSTs for Small Data: Recovering Missing Characters in Old Hawaiian Text"
Tzeviya Fuchs, Joseph Keshet, "Robust Spoken Term Detection Automatically Adjusted for a Given Threshold"
Shuai Tang, Virginia R. de Sa, "Improving Sentence Representations with Multi-view Frameworks"
Brandon Carter, Jonas Mueller, Siddhartha Jain, David Gifford, "Local and global model interpretability via backward selection and clustering"
Albert Zeyer, André Merboldt, Ralf Schlüter, Hermann Ney, "A comprehensive analysis on attention models"
Barbara Rychalska, Dominika Basaj, Przemys■aw Biecek, "Are you tough enough? Framework for Robustness Validation of Machine Comprehension Systems"
Jialu Li, Mark Hasegawa-Johnson, "A Comparable Phone Set for the TIMIT Dataset Discovered in Clustering of Listen, Attend and Spell"
Loren Lugosch, Samuel Myer, Vikrant Singh Tomar, "DONUT: CTC-based Query-by-Example Keyword Spotting"
Titouan Parcollet, Mirco Ravanelli, Mohamed Morchid, Georges Linarès, Renato De Mori, "Speech Recognition with Quaternion Neural Networks"
Wei-Ning Hsu, Yu Zhang, Ron J. Weiss, Yu-An Chung, Yuxuan Wang, Yonghui Wu, James Glass, "Disentangling Correlated Speaker and Noise for Speech Synthesis via Data Augmentation and Adversarial Factorization"
Jan Buys, Yonatan Bisk, Yejin Choi, "Bridging HMMs and RNNs through Architectural Transformations"
Katherine Lee, Orhan Firat, Ashish Agarwal, Clara Fannjiang, David Sussillo, "Hallucinations in neural machine translation"
Ehsan Hosseini-Asl, Yingbo Zhou, Caiming Xiong, Richard Socher, "Robust Domain Adaptation By Augmented Cyclic Adversarial Learning"
Cheng-Zhi Anna Huang, Monica Dinculescu, Ashish Vaswani, Douglas Eck, "Visualizing Music Transformer"
Lea Schönherr, Katharina Kohls, Steffen Zeiler, Dorothea Kolossa, Thorsten Holz, "Adversarial Attacks Against Automatic Speech Recognition Systems via Psychoacoustic Hiding"
Gautam Bhattacharya, Joao Monteiro, Jahangir Alam, Patrick Kenny, "SpeakerGAN: Recognizing Speakers in New Languages with Generative Adversarial Networks"
Jessica Thompson, Marc Schönwiesner, Yoshua Bengio, Daniel Willett, "How transferable are features in convolutional neural network acoustic models across languages?"
Ramin M. Hasani, Alexander Amini, Mathias Lechner, Felix Naser, Radu Grosu, Daniela Rus, "Response Characterization for Auditing Cell Dynamics in Long Short-term Memory Networks"

Abstract 7: **Hynek Hermansky, "Learning - not just for machines anymore" in Interpretability and Robustness in Audio, Speech, and Language**, *Hermansky* 11:00 AM

It is often argued that in processing of sensory signals such as speech, engineering should apply knowledge of properties of human perception - both have the same goal of getting information from the signal. We show on examples from speech technology that perceptual research can also learn from advances in technology. After all, speech evolved to be heard and properties of hearing are imprinted on speech. Subsequently, engineering optimizations of speech technology often yield human-like processing strategies. Further, fundamental difficulties that speech engineering still faces could indicate gaps in our current understanding of the human speech communication process, suggesting directions of further inquiries.

Abstract 8: **Michiel Bacchiani, "Observations in Joint Learning of Features and Classifiers for Speech and Language" in Interpretability and Robustness in Audio, Speech, and Language**, *Bacchiani* 11:30 AM

In relation to launching the Google @home product, we were faced with the problem of far-field speech recognition. That setting gives rise to problems related to reverberant and noisy speech which degrades speech recognition performance. A common approach to address some of these detrimental effects is to use multi-channel processing. This processing is generally seen as an "enhancement" step prior to ASR and is developed and optimized as a separate component of the overall system. In our work, we integrated this component into the neural network that is tasked with the speech recognition classification task. This allows for a joint optimization of the enhancement and recognition components. And given that the structure of the input layer of the network is based on the "classical" structure of the enhancement component, it allows us to interpret what type of representation the network learned. We will show that in some cases this learned representation appears to mimic what was discovered by previous research and in some cases, the learned representation seems "esoteric".

The second part of this talk will focus on an end-to-end letter to sound model for Japanese. Japanese uses a complex orthography where the pronunciation of the Chinese characters, which are a part of the script, varies depending on the context. The fact that Japanese (like Chinese and Korean) does not explicitly mark word boundaries in the orthography further complicates this mapping. We show results of an end-to-end, encoder/decoder model structure to learn the letter-to-sound relationship. These systems are trained from speech data coming through our systems. This shows that such models are capable of learning the mapping (with accuracies exceeding 90% for a number of model topologies). Observing the learned representation and attention distributions for various architectures provides some insight as to what cues the model uses to learn the relationship. But it also shows that interpretation remains limited since the joint optimization of encoder and decoder components allows the model the freedom to learn implicit representations that are not directly amenable to interpretation.

Abstract 9: **Mirco Ravanelli, "Interpretable convolutional filters with SincNet" in Interpretability and Robustness in Audio, Speech, and**

Language, *Ravanelli* 12:00 PM

Deep learning is currently playing a crucial role toward higher levels of artificial intelligence. This paradigm allows neural networks to learn complex and abstract representations, that are progressively obtained by combining simpler ones. Nevertheless, the internal "black-box" representations automatically discovered by current neural architectures often suffer from a lack of interpretability, making of primary interest the study of explainable machine learning techniques.
This paper summarizes our recent efforts to develop a more interpretable neural model for directly processing speech from the raw waveform. In particular, we propose SincNet, a novel Convolutional Neural Network (CNN) that encourages the first layer to discover more meaningful filters by exploiting parametrized sinc functions. In contrast to standard CNNs, which learn all the elements of each filter, only low and high cutoff frequencies of band-pass filters are directly learned from data. This inductive bias offers a very compact way to derive a customized filter-bank front-end, that only depends on some parameters with a clear physical meaning. Our experiments, conducted on both speaker and speech recognition, show that the proposed architecture converges faster, performs better, and is more interpretable than standard CNNs.

Abstract 10: **Hamid Eghbal-zadeh, "Deep Within-Class Covariance Analysis for Robust Deep Audio Representation Learning" in Interpretability and Robustness in Audio, Speech, and Language**, *Eghbalzadeh* 12:15 PM

Deep Neural Networks (DNNs) are known for excellent performance in supervised tasks such as classification. Convolutional Neural Networks (CNNs), in particular, can learn effective features and build high-level representations that can be used for classification, but also for querying and nearest neighbor search. However, CNNs have also been shown to suffer from a performance drop when the distribution of the data changes from training to test data. In this paper, we analyze the internal representations of CNNs and observe that the representations of unseen data in each class, spread more (with higher variance) in the embedding space of the CNN compared to representations of the training data. More importantly, this difference is more extreme if the unseen data comes from a shifted distribution. Based on this observation, we objectively evaluate the degree of representation’s variance in each class by applying eigenvalue decomposition on the within-class covariance of the internal representations of CNNs and observe the same behavior. This can be problematic as larger variances might lead to misclassification if the sample crosses the decision boundary of its class. We apply nearest neighbor classification on the representations and empirically show that the embeddings with the high variance actually have significantly worse KNN classification performances, although this could not be foreseen from their end-to-end classification results. To tackle this problem, we propose Deep Within-Class Covariance Analysis (DWCCA), a deep neural network layer that significantly reduces the within-class covariance of a DNN’s representation, improving performance on unseen test data from a shifted distribution. We empirically evaluate DWCCA on two datasets for Acoustic Scene Classification (DCASE2016 and DCASE2017). We demonstrate that not only does DWCCA significantly improve the network’s internal representation, it also increases the end-to-end classification accuracy, especially when the test set exhibits a slight distribution shift. By adding DWCCA to a VGG neural network, we achieve around 6 percentage points improvement in the case of a distribution
mismatch.

Abstract 12: **Ralf Schlüter, "Automatic Speech Recognition Architectures: from HMM to End-to-End Modeling" in Interpretability and Robustness in Audio, Speech, and Language**, *Schlüter* 01:30 PM

For decades, the general architecture of the classical state-of-the-art statistical approach to automatic speech recognition (ASR) has not been significantly challenged. The classical statistical approach to ASR is based on Bayes decision rule, a separation of acoustic and language modeling, hidden Markov modeling (HMM), and a search organization based on dynamic programming and hypothesis pruning methods. Even when deep neural networks started to considerably boost ASR performance, the general architecture of
state-of-the-art ASR systems was not altered considerably.
The hybrid DNN/HMM approach, together with recurrent LSTM neural network language modeling currently marks the state-of-the-art on many tasks covering a large range of training set sizes.
However, currently more and more alternative approaches occur, moving gradually towards so-called end-to-end approaches. By and by, these novel end-to-end approaches replace explicit time alignment modeling and dedicated search space organization by more implicit, integrated neural-network based representations, also dropping the separation between acoustic and language modeling, showing promising results, especially for large training sets.

In this presentation, an overview of current approaches to ASR will be given, including variations of both HMM-based and end-to-end modeling. Approaches will be discussed w.r.t. their modeling, their performance against available training data, their search space complexity and control, as well as potential modes of comparative analysis.

Abstract 13: **Erik McDermott, "A Deep Generative Acoustic Model for Compositional Automatic Speech Recognition" in Interpretability and Robustness in Audio, Speech, and Language**, *McDermott* 02:00 PM

Inspired by the recent successes of deep generative models for Text-To-Speech (TTS) such as WaveNet (van den Oord et al., 2016) and Tacotron (Wang et al., 2017), this article proposes the use of a deep generative model tailored for Automatic Speech Recognition (ASR) as the primary acoustic model (AM) for an overall recognition system with a separate language model (LM). Two dimensions of depth are considered: (1) the use of mixture density networks, both autoregressive and non-autoregressive, to generate density functions capable of modeling acoustic input sequences with much more powerful conditioning than the first-generation generative models for ASR, Gaussian Mixture Models / Hidden Markov Models (GMM/HMMs), and (2) the use of standard LSTMs, in the spirit of the original tandem approach, to produce discriminative feature vectors for generative modeling. Combining mixture density networks and deep discriminative features leads to a novel dual-stack LSTM architecture directly related to the RNN Transducer (Graves, 2012), but with the explicit functional form of a density, and combining naturally with a separate language model, using Bayes rule. The generative models discussed here are compared experimentally in terms of log-likelihoods and frame accuracies.
Keywords: Automatic Speech Recognition, Deep generative models, Acoustic modeling, End-to-end speech recognition

Abstract 14: **Jamin Shin, "Interpreting Word Embeddings with Eigenvector Analysis" in Interpretability and Robustness in Audio, Speech, and Language**, *Shin* 02:15 PM

Dense word vectors have proven their values in many downstream NLP tasks over the past few years. However, the dimensions of such embeddings are not easily interpretable. Out of the d-dimensions in a word vector, we would not be able to understand what high or low values mean. Previous approaches addressing this issue have mainly focused on either training sparse/non-negative constrained word embeddings, or post-processing standard pre-trained word embeddings. On the other hand, we analyze conventional word embeddings trained with Singular Value Decomposition, and reveal similar interpretability. We use a novel eigenvector analysis method inspired from Random Matrix Theory and show that semantically coherent groups not only form in the row space, but also the column space. This allows us to view individual word vector dimensions as human-interpretable semantic features.

Abstract 15: **Jan Kremer, "On the Inductive Bias of Word-Character-Level Multi-Task Learning for Speech Recognition" in Interpretability and Robustness in Audio, Speech, and Language**, *Kremer* 02:30 PM

End-to-end automatic speech recognition (ASR) commonly transcribes audio signals into sequences of characters while its performance is evaluated by measuring the word-error rate (WER). This suggests that predicting sequences of words directly may be helpful instead. However, training with word-level supervision can be more difficult due to the sparsity of examples per label class. In this paper, we analyze an end-to-end ASR model that combines a word-and-character representation in a multi-task learning (MTL) framework. We show that it improves on the WER and study how the word-level model can benefit from character-level supervision by analyzing the learned inductive preference bias of each model component empirically. We find that by adding character-level supervision, the MTL model interpolates between recognizing more frequent words (preferred by the word-level model) and shorter words (preferred by the character-level model).
Keywords: speech recognition, multi-task learning, interpretability.

Abstract 16: **Coffee break + posters 2 in Interpretability and Robustness in Audio, Speech, and Language**, *Kremer, McDermott, Carter, Zeyer, Krug, Liang, Lee, Basaj, Jimenez, Fan, Bhattacharya, Fuchs, Gifford, Lugosch, Firat, Baer, ALAM, Shin, Ravanelli, Smolensky, Zhu, Eghbalzadeh, Seto, Sheikh, Santos, Belinkov, Durrani, Parker Jones, Tang, Merboldt, Parcollet, Hsu, Pillutla, Hosseini-Asl, Dinculescu, Amini, Zhang, Cheng, Tapp* 02:45 PM

Jamin Shin, Andrea Madotto, Pascale Fung, "Interpreting Word Embeddings with Eigenvector Analysis"
Mirco Ravanelli, Yoshua Bengio, "Interpretable Convolutional Filters with SincNet"
Shuai Tang, Paul Smolensky, Virginia R. de Sa, "Learning Distributed Representations of Symbolic Structure Using Binding and Unbinding Operations"
Lisa Fan, Dong Yu, Lu Wang, "Robust Neural Abstractive Summarization Systems and Evaluation against Adversarial Information"
Zining Zhu, Jekaterina Novikova, Frank Rudzicz, "Semi-supervised classification by reaching consensus among modalities"
Hamid Eghbal-zadeh, Matthias Dorfer, Gerhard Widmer, "Deep Within-Class Covariance Analysis for Robust Deep Audio Representation Learning"
Benjamin Baer, Skyler Seto, Martin T. Wells, "Interpreting Word Embeddings with Generalized Low Rank Models"
Abelino Jimenez, Benjamin Elizalde, Bhiksha Raj, "Sound event classification using ontology-based neural netowrks"
Hai Pham, Paul Pu Liang, Thomas Manzini, Louis-Philippe Morency, Barnabas Poczos, "Found in Translation: Learning Robust Joint Representations by Cyclic Translations Between Modalities"
Sri Harsha Dumpala, Imran Sheikh, Rupayan Chakraborty, Sunil Kumar Kopparapu, "Cycle-Consistent GAN Front-end to Improve ASR Robustness to Perturbed Speech"
Joao Felipe Santos, Tiago H. Falk, "Investigating the effect of residual and highway connections in speech enhancement models"
Jan Kremer, Lasse Borgholt, Lars Maaløe, "On the Inductive Bias of Word-Character-Level Multi-Task Learning for Speech Recognition"
Erik McDermott, "A Deep Generative Acoustic Model for Compositional Automatic Speech Recognition"
Andreas Krug, René Knaebel, Sebastian Stober, "Neuron Activation Profiles for Interpreting Convolutional Speech Recognition Models"
Anthony Bau, Yonatan Belinkov, Hassan Sajjad, Nadir Durrani, Fahim Dalvi, James Glass, "Identifying and Controlling Important Neurons in Neural Machine Translation"
Oiwi Parker Jones, Brendan Shillingford, "Composing RNNs and FSTs for Small Data: Recovering Missing Characters in Old Hawaiian Text"
Tzeviya Fuchs, Joseph Keshet, "Robust Spoken Term Detection Automatically Adjusted for a Given Threshold"
Shuai Tang, Virginia R. de Sa, "Improving Sentence Representations with Multi-view Frameworks"
Brandon Carter, Jonas Mueller, Siddhartha Jain, David Gifford, "Local and global model interpretability via backward selection and clustering"
Albert Zeyer, André Merboldt, Ralf Schlüter, Hermann Ney, "A comprehensive analysis on attention models"
Barbara Rychalska, Dominika Basaj, Przemys■aw Biecek, "Are you tough enough? Framework for Robustness Validation of Machine Comprehension Systems"
Jialu Li, Mark Hasegawa-Johnson, "A Comparable Phone Set for the TIMIT Dataset Discovered in Clustering of Listen, Attend and Spell"
Loren Lugosch, Samuel Myer, Vikrant Singh Tomar, "DONUT: CTC-based Query-by-Example Keyword Spotting"
Titouan Parcollet, Mirco Ravanelli, Mohamed Morchid, Georges Linarès, Renato De Mori, "Speech Recognition with Quaternion Neural Networks"
Wei-Ning Hsu, Yu Zhang, Ron J. Weiss, Yu-An Chung, Yuxuan Wang, Yonghui Wu, James Glass, "Disentangling Correlated Speaker and Noise for Speech Synthesis via Data Augmentation and Adversarial Factorization"
Jan Buys, Yonatan Bisk, Yejin Choi, "Bridging HMMs and RNNs through Architectural Transformations"
Katherine Lee, Orhan Firat, Ashish Agarwal, Clara Fannjiang, David Sussillo, "Hallucinations in neural machine translation"
Ehsan Hosseini-Asl, Yingbo Zhou, Caiming Xiong, Richard Socher, "Robust Domain Adaptation By Augmented Cyclic Adversarial Learning"
Cheng-Zhi Anna Huang, Monica Dinculescu, Ashish Vaswani, Douglas Eck, "Visualizing Music Transformer"
Lea Schönherr, Katharina Kohls, Steffen Zeiler, Dorothea Kolossa, Thorsten Holz, "Adversarial Attacks Against Automatic Speech Recognition Systems via Psychoacoustic Hiding"
Gautam Bhattacharya, Joao Monteiro, Jahangir Alam, Patrick Kenny, "SpeakerGAN: Recognizing Speakers in New Languages with Generative Adversarial Networks"
Jessica Thompson, Marc Schönwiesner, Yoshua Bengio, Daniel Willett, "How transferable are features in convolutional neural network acoustic models across languages?"
Ramin M. Hasani, Alexander Amini, Mathias Lechner, Felix Naser, Radu Grosu, Daniela Rus, "Response Characterization for Auditing Cell Dynamics in Long Short-term Memory Networks"

Abstract 17: **Mike Schuster, "Learning from the move to neural machine translation at Google" in Interpretability and Robustness in Audio, Speech, and Language**, *Schuster* 03:30 PM

At Google we replaced over the last few years the phrase-based machine translation system by GNMT, the Google Neural Machine Translation system. This talk will describe some of the history of this transition and explain the challenges we faced. As part of the new system we developed and used many features that hadn't been used before in production-scale translation systems: A large-scale sequence-to-sequence model with attention, sub-word units instead of a full dictionary to address out-of-vocabulary handling and improve translation accuracy, special hardware to improve inference speed, handling of many language pairs in a single model and other techniques that a) made it possible to launch the system at all and b) to significantly improve on previous production-level accuracy. Some of the techniques we used are now standard in many translation systems – we'd like to highlight some of the remaining challenges in interpretability, robustness and possible solutions to them.

Abstract 18: **Alexander Rush, "Interprebility in Text Generation" in Interpretability and Robustness in Audio, Speech, and Language**, *Rush* 04:00 PM

Neural encoder-decoder models have had significant empirical success in text generation, but there remain major unaddressed issues that make them difficult to apply to real problems. Encoder-decoders are largely (a) uninterpretable in their errors, and (b) difficult to control in areas as phrasing or content. In this talk, I will argue that combining probabilistic modeling with deep learning can help address some of these issues without giving up their advantages. In particular, I will present a method for learning discrete latent templates along with generation. This approach remains deep and end-to-end, achieves comparably good results, and exposes internal model decisions. I will end by discussing some related work on successes and challenges of visualization for interpreting encoder-decoder models.

Abstract 19: **Shuai Tang, "Learning Distributed Representations of Symbolic Structure Using Binding and Unbinding Operations" in Interpretability and Robustness in Audio, Speech, and Language**, *Tang* 04:30 PM

Widely used recurrent units, including Long-short Term Memory (LSTM) and Gated Recurrent Unit (GRU), perform well on natural language tasks, but their ability to learn structured representations is still questionable. Exploiting Tensor Product Representations (TPRs) --- distributed representations of symbolic structure in which vector-embedded symbols are bound to vector-embedded structural positions --- we propose the TPRU, a recurrent unit that, at each time step, explicitly executes structural-role binding and unbinding operations to incorporate structural information into learning. Experiments are conducted on both the Logical Entailment task and the Multi-genre Natural Language Inference (MNLI) task, and our TPR-derived recurrent unit provides strong performance with significantly fewer parameters than LSTM and GRU baselines. Furthermore, our learnt TPRU trained on MNLI demonstrates solid generalisation ability on downstream tasks.

Abstract 20: **Paul Pu Liang, "Learning Robust Joint Representations for Multimodal Sentiment Analysis" in Interpretability and Robustness in Audio, Speech, and Language**, *Shin* 04:45 PM

Multimodal sentiment analysis is a core research area that studies speaker sentiment expressed from the language, visual, and acoustic modalities. The central challenge in multimodal learning involves inferring joint representations that can process and relate information from these modalities. However, existing work learns joint representations using multiple modalities as input and may be sensitive to noisy or missing modalities at test time. With the recent success of sequence to sequence models in machine translation, there is an opportunity to explore new ways of learning joint representations that may not require all input modalities at test time. In this paper, we propose a method to learn robust joint representations by translating between modalities. Our method is based on the key insight that translation from a source to a target modality provides a method of learning joint representations using only the source modality as input. We augment modality translations with a cycle consistency loss to ensure that our joint representations retain maximal information from all modalities. Once our translation model is trained with paired multimodal data, we only need data from the source modality at test-time for prediction. This ensures that our model remains robust from perturbations or missing target modalities. We train our model with a coupled translation-prediction objective and it achieves new state-of-the-art results on multimodal sentiment analysis datasets: CMU-MOSI, ICT-MMMO, and YouTube. Additional experiments show that our model learns increasingly discriminative joint representations with more input modalities while maintaining robustness to perturbations of all other modalities.

Abstract 21: **Jason Eisner, "BiLSTM-FSTs and Neural FSTs" in Interpretability and Robustness in Audio, Speech, and Language**, *Eisner* 05:00 PM

How should one apply deep learning to tasks such as morphological reinflection, which stochastically edit one string to get another? Finite-state transducers (FSTs) are a well-understood formalism for scoring such edit sequences, which represent latent hard monotonic alignments. I will discuss options for combining this architecture with neural networks. The BiLSTM-FST scores each edit in its full input context, which preserves the ability to do exact inference over the aligned outputs using dynamic programming. The Neural FST scores each edit sequence using an LSTM, which requires approximate inference via methods such as beam search or particle smoothing. Finally, I will sketch how to use the language of regular expressionsto specify not only the legal edit sequences but also how to present them to the LSTMs.

Abstract 22: **Panel Discussion in Interpretability and Robustness in Audio, Speech, and Language**, *Caruana, Schuster, Schlüter, Hermansky, De Mori, Bengio, Bacchiani, Eisner* 05:30 PM

Panel Discussion on "Interpretability and Robustness in Audio, Speech, and Language" (moderated by Jason Eisner).

Panelists:
- Sami Bengio
- Rich Caruana
- Mike Schuster
- Ralf Schlueter
- Hynek Hermansky
- Renato DeMori
- Michiel Bacchiani
- Jason Eisner

NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018

Li Erran Li, Anca Dragan, Juan Carlos Niebles, Silvio Savarese

Room 514, Sat Dec 08, 08:00 AM

Our transportation systems are poised for a transformation as we make progress on autonomous vehicles, vehicle-to-vehicle (V2V) and vehicle-to-everything (V2X) communication infrastructures, and smart road infrastructures (like smart traffic lights). But many challenges stand in the way of this transformation. For example, how do we make perception accurate and robust enough to accomplish safe autonomous driving? How do we generate policies that equip autonomous cars with adaptive human negotiation skills when merging, overtaking, or yielding? How do we decide when a system is safe enough to deploy? And how do we optimize efficiency through intelligent traffic management and control of fleets?

To meet these requirements in safety, efficiency, control, and capacity, the systems must be automated with intelligent decision making. Machine learning will be an essential component of that. Machine learning has made rapid progress in the self-driving domain (e.g., in real-time perception and prediction of traffic scenes); has started to be applied to ride-sharing platforms such as Uber (e.g., demand forecasting); and by crowd-sourced video scene analysis companies such as Nexar (e.g., understanding and avoiding accidents). But to address the challenges arising in our future transportation system, we need to consider the transportation systems as a whole rather than solving problems in isolation, from prediction, to behavior, to infrastructure.

The goal of this workshop is to bring together researchers and practitioners from all areas of intelligent transportations systems to address core challenges with machine learning. These challenges include, but are not limited to
pedestrian detection, intent recognition, and negotiation,
coordination with human-driven vehicles,
machine learning for object tracking,
unsupervised representation learning for autonomous driving,
deep reinforcement learning for learning driving policies,
cross-modal and simulator to real-world transfer learning,
scene classification, real-time perception and prediction of traffic scenes,
uncertainty propagation in deep neural networks,
efficient inference with deep neural networks
predictive modeling of risk and accidents through telematics, modeling, simulation and forecast of demand and mobility patterns in large scale urban transportation systems,
machine learning approaches for control and coordination of traffic leveraging V2V and V2X infrastructures,

The workshop will include invited speakers, panels, presentations of accepted papers, and posters. We invite papers in the form of short, long, and position papers to address the core challenges mentioned above. We encourage researchers and practitioners on self-driving cars, transportation systems and ride-sharing platforms to participate. Since this is a topic of broad and current interest, we expect at least 150 participants from leading university researchers, auto-companies and ride-sharing companies.

This will be the 3rd NIPS workshop in this series. Previous workshops have been very successful and have attracted large numbers of participants from both academia and industry.

Schedule

08:45 AM	**Opening Remark**	*Li, Dragan*
09:00 AM	**Invited talk: Alfredo Canziani, NYU**	*Canziani*
09:30 AM	**Invited Talk: Drew Bagnell, CMU and Aurora**	*Bagnell*
10:00 AM	**Invited Talk: Yimeng Zhang, Pony.ai**	*Zhang*
10:30 AM	**Coffee break: morning**	
11:00 AM	**Invited Talk: Nathaniel Fairfield, Waymo**	*Fairfield*
11:30 AM	**John J. Leonard, MIT and TRI**	*Leonard*
12:00 PM	**Lunch**	
01:30 PM	**Invited Talk: Dorsa Sadigh, Stanford**	*Sadigh*
02:00 PM	**Invited Talk: Marco Pavone, Stanford**	*Pavone*
02:30 PM	**Contributed Talks**	*Rhinehart, Shah*
03:00 PM	**Coffee break: afternoon**	
03:30 PM	**Invited Talk: Ingmar Posner, Oxford and Oxbotica**	*Posner*
04:00 PM	**Invited Talk: Ekaterina Taralova and Sarah Tariq, Zoox**	*Taralova, Tariq*
04:30 PM	**Panel**	*Zhang, Canziani, Pavone, Sadigh, Keutzer*
05:15 PM	**Poster Session**	*Ding, Mguni, Zhuang, Leurent, Oda, Tachibana, Gora, Davis, Djuric, Chou, amirloo*

Abstracts (11):

Abstract 2: **Invited talk: Alfredo Canziani, NYU in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Canziani* 09:00 AM

Title: Prediction and Planning Under Uncertainty: The Case of Autonomous Driving

Abstract: In order to achieve a well specified goal an agent may use two distinct approaches: trial & error, or careful planning. In the first case the agent has to fail multiple times before learning a task (e.g. playing a card game), in the second we leverage the knowledge of the environment to avoid any fatal failure (e.g. vehicle collision).

Autonomous driving relies on accurate planning, which requires a good model of the world that also considers other vehicles' future response to our own actions. Effectively learning to predict such response, stochastic by nature, is the key aspect to successfully obtain planning under uncertainty.

Bio: Alfredo Canziani is a Post-Doctoral Deep Learning Research Scientist and Lecturer at NYU Courant Institute of Mathematical Sciences, under the supervision of professors KyungHyun Cho and Yann LeCun. His research mainly focusses on Machine Learning for Autonomous Driving. He has been exploring deep policy networks actions uncertainty estimation and failure detection, and long term planning based on latent forward models, which nicely deal with the stochasticity and multimodality of the surrounding environment. Alfredo obtained both his Bachelor (2009) and Master (2011) degrees in EEng cum laude at Trieste University, his MSc (2012) at Cranfield University, and his PhD (2017) at Purdue University. In his spare time, Alfredo is a professional musician, dancer, and cook, and keeps expanding his free online video course on Deep Learning, Torch, and PyTorch.

Abstract 3: **Invited Talk: Drew Bagnell, CMU and Aurora in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Bagnell* 09:30 AM

Bio: James A. (Drew) Bagnell is Chief Technology Officer of Aurora (aurora.tech), where he works with an amazing team to develop and deliver self-driving safely, quickly and broadly. Dr. Bagnell has worked for 19 years at the intersection of machine learning and robotics with expertise in self-driving cars, imitation and reinforcement learning, planning, and computational perception. Aurora was founded in 2017 to enable autonomous driving solutions that will make roads safer, improve lives, revitalize cities, and expand transportation access.

Dr. Bagnell is also an adjunct professor at Carnegie Mellon University’s Robotics Institute and Machine Learning Department. His interests in artificial intelligence range from algorithmic and basic theoretical development to delivering fielded learning-based systems. Bagnell and his research group have received over a dozen research awards for publications in both the robotics and machine learning communities including best paper awards at ICML, RSS, and ICRA.

He received the 2016 Ryan Award, Carnegie Mellon University’s award for Meritorious Teaching, and served as the founding director of the Robotics Institute Summer Scholars program, a summer research experience that has enabled hundreds of undergraduates throughout the world to leap into robotics research.

Abstract 4: **Invited Talk: Yimeng Zhang, Pony.ai in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Zhang* 10:00 AM

Title: On the generalization of autonomous driving technologies

Abstract: Most L4 autonomous driving companies are working on the solution of 1-2 cities. The generalization from one city to 10 cities is a very challenging problem, in particular, when the 10 cities are very different and even cross different countries. This requires much more generalization ability of the algorithm, including the deep learning algorithms. In this talk, I will talk about the challenge of generalization in autonomous driving, and the interesting problems we encountered when testing the system in different countries.

Abstract 7: **John J. Leonard, MIT and TRI in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Leonard* 11:30 AM

Title: Guardian Research Challenges and Opportunities

Abstract: This talk will describe research underway at Toyota Research Institute
and its partner universities to create the Toyota Guardian system for increasing the safety of human driving by exploiting advanced navigation, perception, prediction, and planning capabilities that are becoming available. The objective of Guardian is to create a highly automated driving system that can act as a safety net for the human driver to help prevent an accident, with three primary aims: (1) stay on the road; (2) don't hit things; (3) don't get hit. We will discuss some of the research challenges and opportunities for realizing such a system, spanning a wide range of topics in computer vision, machine learning, and mobile robotics.

Bio: Dr. John J. Leonard is Samuel C. Collins Professor in the MIT Department of Mechanical Engineering and a member of the MIT Computer
Science and Artificial Intelligence Laboratory (CSAIL). His research addresses the problems of navigation and mapping for autonomous mobile
robots and underwater vehicles. He holds the degrees of B.S.E.E. in Electrical Engineering and Science from the University of Pennsylvania (1987) and D.Phil. in Engineering Science from the University of Oxford (1994). He was team leader for MIT's DARPA Urban Challenge team, which was one of eleven teams to qualify for the Urban Challenge final event and one of six teams to complete the race. He is the recipient of an NSF Career Award (1998) and the King-Sun Fu Memorial Best Transactions on Robotics Paper Award (2006). He is an IEEE Fellow (2014). Professor Leonard has recently been on partial leave from MIT serving as Vice President of Automated Driving Research at the Toyota Research Institute (TRI).

Abstract 9: **Invited Talk: Dorsa Sadigh, Stanford in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Sadigh* 01:30 PM

Title: Altruistic Autonomy: Beating Congestion on Shared Roads

Abstract: Today the emergence of autonomous cars on public roads has become a reality. Autonomous vehicles on roads shared with human-driven cars can create a variety of challenges including influencing traffic flow when the transportation network is under heterogeneous use: when cars of differing levels of autonomy co-exist on the same road. In this talk, we will address some of the challenges of mixed-autonomy traffic networks via leveraging the power of autonomous vehicles. Specifically we will focus on two main approaches that use autonomous cars to positively influence congestion. First we discuss how local interactions between the vehicles can affect the global behavior of a traffic network. We will examine a high-level queuing framework to study the capacity of a mixed-autonomy transportation network, and then outline a lower-level control framework that leverages local interactions between cars to achieve a more efficient traffic flow via intelligent reordering of the cars. We provide theoretical bounds on the capacity that can be achieved by the network for given autonomy level. Second, we formalize the notion of altruistic autonomy—autonomous vehicles that are incentivized to take longer routes in order to alleviate

congestion on mixed-autonomy roads. We then study the effects of altruistic autonomy on roads shared between human drivers and autonomous vehicles. We develop a formal model of road congestion on shared roads based on the fundamental diagram of traffic, and discuss algorithms that compute optimal equilibria robust to additional unforeseen demand, and plan for optimal routings when users have varying degrees of altruism. We find that even with arbitrarily small altruism, total latency can be unboundedly lower than without altruism.

Abstract 10: **Invited Talk: Marco Pavone, Stanford in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Pavone* 02:00 PM

Title: On safe and efficient human-robot vehicle interactions via CVAE-based intent modeling and reachability-based safety assurance

Abstract: In this talk I will present a decision-making and control stack for human-robot vehicle interactions. I will first discuss a data-driven approach for learning interaction dynamics between robot-driven and human-driven vehicles, based on recent advances in the theory of conditional variational autoencoders (CVAEs). I will then discuss how to incorporate such a learned interaction model into a real-time, intent-aware decision-making framework, with an emphasis on minimally-interventional strategies rooted in backward reachability analysis for ensuring safety even when other cars defy the robot's predictions. Experiments on a full-scale steer-by-wire platform entailing traffic weaving maneuvers demonstrate how the proposed autonomy stack enables more efficient and anticipative autonomous driving behaviors, while avoiding collisions even when the other cars defy the robot's predictions and take dangerous actions.

Abstract 11: **Contributed Talks in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Rhinehart, Shah* 02:30 PM

Deep Object Centric Policies for Autonomous Driving, Dequan Wang (presenter), Coline Devin, Qi-Zhi Cai, Fisher Yu, Trevor Darrell

Deep Imitative Models for Flexible Inference, Planning, and Control, Nicholas Rhinehart (presenter), Rowan McAllister, Sergey Levine

Learning to Drive in a Day, Alex Kendall, Jeffrey Hawke, David Janz, Przemyslaw Mazur, Daniele Reda, John-Mark Allen, Vinh-Dieu Lam, Alex Bewley, Amar Shah (presenter)

Abstract 13: **Invited Talk: Ingmar Posner, Oxford and Oxbotica in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Posner* 03:30 PM

Title: Driving Autonomy - Machine Learning for Intelligent Transportation Systems

Abstract:
Intelligent transportation systems, and autonomous driving in particular, have captured the public imagination. Most of us are excited about the art of the possible. Machine learning clearly has a role to play. In this talk I will argue that a systems view of autonomous driving affords the machine learning community particular opportunities - and poses some interesting challenges - beyond an out-of-the-box deployment of strategies and models developed in related fields.

Bio:
Prof. Ingmar Posner leads the Applied Artificial Intelligence Lab (A2I) at Oxford University. He also serves as Deputy Director of the Oxford Robotics Institute, which he co-founded in 2016. Ingmar has a significant track record in designing machine learning approaches (shallow and deep) which address core challenges in AI and machine learning. Ingmar's goal is to enable robots to robustly and effectively operate in complex, real-world environments. His research is guided by a vision to create machines which constantly improve through experience. In doing so Ingmar's work explores a number of intellectual challenges at the heart of robot learning, such as machine introspection in perception and decision making, data efficient learning from demonstration, transfer learning and the learning of complex tasks via a set of less complex ones. All the while Ingmar's intellectual curiosity remains grounded in real-world robotics applications such as autonomous driving, logistics, manipulation and space exploration. In 2014 Ingmar co-founded Oxbotica, a leading provider of mobile autonomy software solutions.

Abstract 14: **Invited Talk: Ekaterina Taralova and Sarah Tariq, Zoox in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Taralova, Tariq* 04:00 PM

Title: Sensing and simulating the real world for next generation autonomous mobility

Abstract: Zoox is developing the first ground-up, fully autonomous vehicle fleet and the supporting ecosystem required to bring this technology to market. Sitting at the intersection of robotics, machine learning, and design, Zoox aims to provide the next generation of mobility-as-a-service in urban environments. A key part of our design is safety: in addition to providing a great user experience, we aim to design robots that are significantly safer than human drivers. To ensure this, it is critical to maintain accurate perception of objects in the world that our robots need to react to. To that end, Zoox has taken a holistic approach to its sensor choice and placement, computational power, and algorithms. In the first half of our talk we will describe some of the sensors and algorithms we use to ensure that the robot is able to perceive all objects that it needs to react to, and that it is able to do so with sufficiently low latency.

In addition to developing the necessary technology, it is imperative to be able to validate it. Zoox is developing an advanced 3D simulation framework to help verify that our vehicle is safe while also being able to complete its missions successfully. This framework provides the foundation for generating highly realistic simulated data which is used as ground truth for testing algorithms as well as to train machine learning algorithms in cases when sufficient real-world data is not readily available. The second part of the talk will provide an overview of this framework and, in particular, discuss how we quantify the fidelity of the various simulated sensor types used by our robot in its perception stack.

Bios:
Sarah is the Director of Vision Detection and Tracking at Zoox, where her team focuses on perception for cameras, including detecting and tracking objects of interest reliably and in real time.
Sarah has been at Zoox for over three years and before that she has almost a decade of experience working at NVIDIA across multiple roles. Amongst her many achievements at NVIDIA, she contributed to the implementation of novel real-time simulation and rendering of algorithms for video games, managed a team working on profiling and optimizing code for high performance computing and super computers, and served as a technical lead for the computer vision team focusing on self-driving

technology.

Ekaterina is a senior research engineer at Zoox. Her goal is to quantify how realistic simulated sensors need to be to to enable end-to-end testing of the software stack and create synthetic training data to help improve perception models. Before Zoox, she was a postdoc with Tony Jebara and Rafael Yuste at Columbia University, where she developed large scale graphical models to quantify neural activity in the mouse visual cortex. Ekaterina obtained her PhD with Martial Hebert and Fernando De la Torre at Carnegie Mellon University, where her thesis work was on action classification and segmentation in videos.

Abstract 15: **Panel in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Zhang, Canziani, Pavone, Sadigh, Keutzer* 04:30 PM

Discussion on key challenges and approaches of AI for autonomous driving

Abstract 16: **Poster Session in NIPS Workshop on Machine Learning for Intelligent Transportation Systems 2018**, *Ding, Mguni, Zhuang, Leurent, Oda, Tachibana, Gora, Davis, Djuric, Chou, amirloo* 05:15 PM

Deep Reinforcement Learning for Intelligent Transportation Systems,
Xiao-Yang Liu, Zihan Ding (presenter), Sem Borst, Anwar Walid

GANtruth – an unpaired image-to-image translation method for driving scenarios
Sebastian Bujwid (presenter), Miquel Martí Rabadan, Hossein Azizpour, Alessandro Pieropan

Controlling the Crowd: Inducing Efficient Equilibria in Multi-Agent Systems
David Mguni (presenter), Joel Jennings, Sergio Valcarcel Macua, Sofia Ceppi, Enrique Munoz de Cote

Distributed Fleet Control with Maximum Entropy Deep Reinforcement Learning
Takuma Oda (presenter), Yulia Tachibana (presenter)

Robust Auto-parking: Reinforcement Learning based Real-time Planning Approach with Domain Template
Yuzheng Zhuang (presenter), Qiang Gu, Bin Wang, Jun Luo, Hongbo Zhang, Wulong Liu

Approximate Robust Control of Uncertain Dynamical Systems
Edouard Leurent (presenter), Yann Blanco, Denis Efimov, Odalric-Ambrym Maillard

Towards Comprehensive Maneuver Decisions for Lane Change Using Reinforcement Learning
Chen Chen, Jun Qian, Hengshuai Yao, Jun Luo, Hongbo Zhang, Wulong Liu (presenter)

Investigating performance of neural networks and gradient boosting models approximating microscopic traffic simulations in traffic optimization tasks
Pawe■ Gora (presenter), Maciej Brzeski, Marcin Mo■ejko, Arkadiusz Klemenko, Adrian Kocha■ski

Taxi Demand-Supply Forecasting: Impact of Spatial Partitioning on the Performance of Neural Networks
Neema Davis (presenter), Gaurav Raina, Krishna Jagannathan

Predicting Motion of Vulnerable Road Users using High-Definition Maps and Efficient ConvNets
Fang-Chieh Chou (presenter), Tsung-Han Lin, Henggang Cui, Vladan Radosavljevic, Thi Nguyen, Tzu-Kuo Huang, Matthew Niedoba, Jeff Schneider, Nemanja Djuric (presenter)

Towards Practical Hierarchical Reinforcement Learning for Multi-lane Autonomous Driving
Masoud S. Nosrati, Elmira Amirloo Abolfathi (presenter), Mohammed Elmahgiubi, Peyman Yadmellat, Jun Luo, Yunfei Zhang, Hengshuai Yao, Hongbo Zhang, Anas Jamil

Risk-averse Behavior Planning for Autonomous Driving under Uncertainty
Mohammad Naghshvar (presenter), Ahmed K. Sadek, Auke J. Wiggers

Machine Learning Open Source Software 2018: Sustainable communities

Heiko Strathmann, Viktor Gal, Ryan Curtin, Sergey Lisitsyn, Antti Honkela, Cheng Soon Ong

Room 515, Sat Dec 08, 08:00 AM

Machine learning open source software (MLOSS) is one of the cornerstones of open science and reproducible research. Once a niche area for ML research, MLOSS today has gathered significant momentum, fostered both by scientific community, and more recently by corporate organizations. Along with open access and open data, it enables free reuse and extension of current developments in ML. The past mloss.org workshops at NIPS06, NIPS08, ICML10, NIPS13, and ICML15 successfully brought together researchers and developers from both fields, to exchange experiences and lessons learnt, to encourage interoperability between people and projects, and to demonstrate software to users in the ML community.

Continuing the tradition in 2018, we plan to have a workshop that is a mix of invited speakers, contributed talks and discussion/activity sessions. This year’s headline aims to give an insight of the challenges faced by projects as they seek long-term sustainability, with a particular focus on community building and preservation, and diverse teams. In the talks, we will cover some of the latest technical innovations as done by established and new projects. The main focus, however, will be on insights on project sustainability, diversity, funding and attracting new developers, both from academia and industry. We will discuss various strategies that helps promoting gender diversity in projects (e.g. implementing quotas etc.) and how to promote developer growth within a project.

We aim to make this workshop as diverse as possible within the field. This includes a gender balanced speakers, focussing on programming languages from different scientific communities, and in particular most of our invited speakers represent umbrella projects with a hugely diverse set of applications and users (NumFOCUS, openML, tidyverse).

With a call for participation for software project demos, we aim to provide improved outreach and visibility, especially for smaller OSS projects as

typically present in academia. In addition, our workshop will serve as a gathering of OSS developers in academia, for peer to peer exchange of learnt lessons, experiences, and sustainability and diversity tactics.

The workshop will include an interactive session to produce general techniques for driving community engagement and sustainability, such as application templates (Google Summer of Code, etc), “getting started” guides for new developers, and a collection of potential funding sources. We plan to conclude the workshop with a discussion on the headline topic.

Schedule

08:25 AM	**Opening remarks**	
08:30 AM	**Gina Helfrich, NumFOCUS**	*Helfrich*
09:00 AM	**Christoph Hertzberg, Eigen3**	*Hertzberg*
09:30 AM	**Joaquin Vanschoren, OpenML**	
10:00 AM	**Sherpa: Hyperparameter Optimization for Machine Learning Models**	*Sadowski*
10:05 AM	**How to iNNvestigate neural network’s predictions!**	
10:10 AM	**mlpack open-source machine learning library and community**	*Edel*
10:15 AM	**Stochastic optimization library: SGDLibrary**	*Kasai*
10:20 AM	**Baseline: Strong, Extensible, Reproducible, Deep Learning Baselines for NLP**	*Lester*
10:25 AM	**McTorch, a manifold optimization library for deep learning**	*Kunchukuttan*
10:25 AM	**Tensorflex: Tensorflow bindings for the Elixir programming language**	*Chhabra*
10:25 AM	**Open Source Machine Learning Software Development in CERN(High-Energy Physics): lessons and exchange of experience**	*Gleyzer*
10:25 AM	**Accelerating Machine Learning Research with MI-Prometheus**	*Marois*
10:25 AM	**Gravity: A Mathematical Modeling Language for Optimization and Machine Learning**	*Hijazi*
10:25 AM	**skpro: A domain-agnostic modelling framework for probabilistic supervised learning**	*Kiraly*
10:25 AM	**xpandas - python data containers for structured types and structured machine learning tasks**	*Davydov*
10:25 AM	**Machine Learning at Microsoft with ML.NET**	*Weimer*
10:25 AM	**Open Fabric for Deep Learning Models**	
10:25 AM	**Towards Reproducible and Reusable Deep Learning Systems Research Artifacts**	*Moreau*
10:25 AM	**PyLissom: A tool for modeling computational maps of the visual cortex in PyTorch**	*Barijhoff*
10:25 AM	**Salad: A Toolbox for Semi-supervised Adaptive Learning Across Domains**	*Schneider*
10:25 AM	**Why every GBM speed benchmark is wrong**	*Ershov*
10:25 AM	**Discussion over morning coffee**	
11:20 AM	**Building, growing and sustaining ML communities**	*Andrews*
11:40 AM	**PyMC's Big Adventure: Lessons Learned from the Development of Open-source Software for Probabilistic Programming**	*Fonnesbeck*
12:00 PM	**Lunch (on your own)**	
02:00 PM	**James Hensman, GPFlow**	
02:30 PM	**Mara Averick, tidyverse**	
03:00 PM	**Afternoon coffee break**	
03:30 PM	**DeepPavlov: An Open Source Library for Conversational AI**	*Kuratov*
03:50 PM	**MXFusion: A Modular Deep Probabilistic Programming Library**	*Dai*
04:10 PM	**Flow: Open Source Reinforcement Learning for Traffic Control**	*Kheterpal*
04:30 PM	**Reproducing Machine Learning Research on Binder**	
04:50 PM	**Panel discussion**	

05:30 PM **Closing remarks**

Abstracts (24):

Abstract 5: **Sherpa: Hyperparameter Optimization for Machine Learning Models in Machine Learning Open Source Software 2018: Sustainable communities**, *Sadowski* 10:00 AM

Sherpa is a free open-source hyperparameter optimization library for machine learning models. It is designed for problems with computationally expensive iterative function evaluations, such as the hyperparameter tuning of deep neural networks. With Sherpa, scientists can quickly optimize hyperparameters using a variety of powerful and interchangeable algorithms. Additionally, the framework makes it easy to implement custom algorithms.
Sherpa can be run on either a single machine or a cluster via a grid scheduler with minimal configuration. Finally, an interactive dashboard enables users to view the progress of models as they are trained, cancel trials, and explore which hyperparameter combinations are working best. Sherpa empowers machine learning researchers by automating the tedious aspects of model tuning and providing an extensible framework for developing automated hyperparameter-tuning strategies. Its source code and documentation are available at https://github.com/LarsHH/sherpa and https://parameter-sherpa.readthedocs.io/, respectively. A demo can be found at https://youtu.be/L95sasMLgP4.

Abstract 6: **How to iNNvestigate neural network's predictions! in Machine Learning Open Source Software 2018: Sustainable communities**, 10:05 AM

In recent years, deep neural networks have revolutionized many application domains of machine learning and are key components of many critical decision or predictive processes such as autonomous driving or medical image analysis. In these and many other domains it is crucial that specialists can understand and analyze actions and predictions, even of the most complex neural network architectures. Despite these arguments neural networks are often treated as black boxes and their complex internal workings as well as the basis for their predictions are not fully understood.
In the attempt to alleviate this shortcoming many analysis methods were proposed, yet the lack of reference implementations often makes a systematic comparison between the methods a major effort. In this tutorial we present the library iNNvestigate which addresses the mentioned issue by providing a common interface and out-of-the-box implementation for many analysis methods. In the first part we will show how iNNvestigate enables users to easily compare such methods for neural networks. The second part will demonstrate how the underlying API abstracts a common operations in neural network analysis and show how users can use them for the development of (future) methods.

Abstract 7: **mlpack open-source machine learning library and community in Machine Learning Open Source Software 2018: Sustainable communities**, *Edel* 10:10 AM

mlpack is an open-source C++ machine learning library with an emphasis on speed and flexibility. Since its original inception in 2007, it has grown to be a large project, implementing a wide variety of machine learning algorithms. This short paper describes the general design principles and discusses how the open-source community around the library functions.

Abstract 8: **Stochastic optimization library: SGDLibrary in Machine Learning Open Source Software 2018: Sustainable communities**, *Kasai* 10:15 AM

SGDLibrary is a open source MATLAB library of stochastic optimization algorithms, which finds the minimizer of a function f: R^d -> R of the finite-sum form min f(w) = 1/n sum_i f_i(w). This problem has been studied intensively in recent years in the field of machine learning. One typical but promising approach for large-scale data is to use a stochastic optimization algorithm to solve the problem. SGDLibrary is a readable, flexible and extensible pure-MATLAB library of a collection of stochastic optimization algorithms. The purpose of the library is to provide researchers and implementers a comprehensive evaluation environment for the use of these algorithms on various machine learning problems.

Abstract 9: **Baseline: Strong, Extensible, Reproducible, Deep Learning Baselines for NLP in Machine Learning Open Source Software 2018: Sustainable communities**, *Lester* 10:20 AM

Natural Language Processing is now dominated by deep learning models. Baseline (https://github.com/dpressel/baseline) is a library to facilitate reproducible research and fast model development for NLP with deep learning. It provides easily extensible implementations and abstractions for data loading, model development, training, hyper-parameter tuning, deployment to production, and a leaderboard to track experimental results.

Abstract 10: **McTorch, a manifold optimization library for deep learning in Machine Learning Open Source Software 2018: Sustainable communities**, *Kunchukuttan* 10:25 AM

In this paper, we introduce McTorch, a manifold optimization library for deep learning that extends PyTorch. It aims to lower the barrier for users wishing to use manifold constraints in deep learning applications, i.e., when the parameters are constrained to lie on a manifold. Such constraints include the popular orthogonality and rank constraints, and have been recently used in a number of applications in deep learning. McTorch follows PyTorch's architecture and decouples manifold definitions and optimizers, i.e., once a new manifold is added it can be used with any existing optimizer and vice-versa. McTorch is available at https://github.com/mctorch.

Abstract 11: **Tensorflex: Tensorflow bindings for the Elixir programming language in Machine Learning Open Source Software 2018: Sustainable communities**, *Chhabra* 10:25 AM

Recently, with the advent of programmatic and practical machine learning tools, programmers have been able to successfully integrate applications for the web and the mobile with artificial intelligence capabilities. This trend has largely been possible because of major organizations and software companies releasing their machine learning frameworks to the public-- such as Tensorflow (Google), MXnet (Amazon) and PyTorch (Facebook). Python has been the de facto choice as the programming language for these frameworks because of it's versatility and ease-of-use. In a similar vein, Elixir is the functional programming language equivalent of Python and Ruby, in that it combines the versatility and ease-of-use that Python and Ruby boast of, with functional programming paradigms and the Erlang VM's fault tolerance and robustness. However, despite these obvious advantages, Elixir, similar to other functional programming languages, does not

provide developers with a machine learning toolset which is essential for equipping applications with deep learning and statistical inference features. To bridge this gap, we present Tensorflex, an open source framework that allows users to leverage pre-trained Tensorflow models (written in Python, C or C++) for Inference (generating predictions) in Elixir. Moreover, Tensorflex was written as part of a Google Summer of Code (2018) project by Anshuman Chhabra, and José Valim was the mentor for the same.

Abstract 12: **Open Source Machine Learning Software Development in CERN(High-Energy Physics): lessons and exchange of experience in Machine Learning Open Source Software 2018: Sustainable communities**, *Gleyzer* 10:25 AM

In my talk, I would like to share the experience and lessons learned in creating an open source machine learning software at CERN with a significant user base in high-energy physics, expanding the open-source initiative to the wider physics community and interactions with the open source community at large.

Abstract 13: **Accelerating Machine Learning Research with MI-Prometheus in Machine Learning Open Source Software 2018: Sustainable communities**, *Marois* 10:25 AM

The paper introduces MI-Prometheus (Machine Intelligence - Prometheus), an open-source framework aiming at accelerating Machine Learning Research, by fostering the rapid development of diverse neural network-based models and facilitating their comparison. In its core, to accelerate the computations on their own, MI-Prometheus relies on PyTorch and extensively uses its mechanisms for the distribution of computations on CPUs/GPUs. The paper discusses the motivation of our work, the formulation of the requirements and presents the core concepts. We also briefly describe some of the problems and models currently available in the framework.

Abstract 14: **Gravity: A Mathematical Modeling Language for Optimization and Machine Learning in Machine Learning Open Source Software 2018: Sustainable communities**, *Hijazi* 10:25 AM

Gravity is an open source, scalable, memory efficient modeling language for solving mathematical models in Optimization and Machine Learning. Gravity exploits structure to reduce function evaluation time including Jacobian and Hessian computation. Gravity is implemented in C++ with a flexible interface allowing the user to specify the numerical accuracy of variables and parameters. It is also designed to offer efficient iterative model solving, convexity detection, multithreading of subproblems, and lazy constraint generation. When compared to state-of-the-art modeling languages such as JuMP, Gravity is 5 times faster in terms of function evaluation and up to 60 times more memory efficient. Gravity enables researchers and practitioners to access state-of-the-art optimization solvers with a user-friendly interface for writing general mixed-integer nonlinear models.

Abstract 15: **skpro: A domain-agnostic modelling framework for probabilistic supervised learning in Machine Learning Open Source Software 2018: Sustainable communities**, *Kiraly* 10:25 AM

We present skpro, a Python framework for domain-agnostic probabilistic supervised learning. It features a scikit-learn-like general API that supports the implementation and fair comparison of both Bayesian and frequentist prediction strategies that produce conditional predictive distributions for each individual test data point. The skpro interface also supports strategy optimization through hyper-paramter tuning, model composition, ensemble methods like bagging, and workflow automation. The package and documentation are released under the BSD-3 open source license and available at GitHub.com/alan-turing-institute/skpro.

Abstract 16: **xpandas - python data containers for structured types and structured machine learning tasks in Machine Learning Open Source Software 2018: Sustainable communities**, *Davydov* 10:25 AM

Data scientific tasks with structured data types, e.g., arrays, images, time series, text records, are one of the major challenge areas of contemporary machine learning and AI research beyond the ``tabular'' situation - that is, data that fits into a single classical data frame, and learning tasks on it such as the classical supervised learning task where one column is to be predicted from others.\\
With xpandas, we present a python package that extends the pandas data container functionality to cope with arbitrary structured types (such as time series, images) at its column/slice elements, and which provides a transformer interface to scikit-learn's pipeline and composition workflows.\\
We intend xpandas to be the first building block towards scikit-learn like toolbox interfaces for advanced learning tasks such as supervised learning with structured features, structured output prediction, image segmentation, time series forecasting and event risk modelling.

Abstract 17: **Machine Learning at Microsoft with ML.NET in Machine Learning Open Source Software 2018: Sustainable communities**, *Weimer* 10:25 AM

Machine Learning is transitioning from an art and science into a technology available to every developer.
In the near future, every application on every platform will incorporate trained models to encode data-based decisions that would be impossible for developers to author. This presents a significant engineering challenge, since currently data science and modeling are largely decoupled from standard software development processes.
This separation makes incorporating machine learning capabilities inside applications unnecessarily costly and difficult, and furthermore discourage developers from embracing ML in first place.
In this paper we introduce ML.NET, a framework developed at Microsoft over the last decade in response to the challenge of making it easy to ship machine learning models in large software applications.

Abstract 18: **Open Fabric for Deep Learning Models in Machine Learning Open Source Software 2018: Sustainable communities**, 10:25 AM

We will show the advantages of using a fabric of open source AI services and libraries, which have been launched by the AI labs in IBM Research, to train, harden and de-bias deep learning models. The motivation is that model building should not be monolithic. Algorithms, operations and pipelines to build and refine models should be modularized and reused as needed. The componentry presented meets these requirements and shares a philosophy of being framework- and vendor- agnostic, as well as modular and extensible. We focus on multiple aspects of machine learning that we describe in the following. To train models in the cloud in a distributed, framework-agnostic way, we use the Fabric for Deep Learning (FfDL). Adversarial attacks against models are mitigated using the Adversarial Robustness Toolbox (ART). We detect and remove bias using AI Fairness 360 (AIF360). Additionally, we publish to the open source developer community using the Model Asset Exchange (MAX). Overall, we demonstrate operations on deep learning models, and a set

of developer APIs, that will help open source developers create robust and fair models for their applications, and for open source sharing. We will also call for community collaboration on these projects of open services and libraries, to democratize the open AI ecosystem.

Abstract 19: **Towards Reproducible and Reusable Deep Learning Systems Research Artifacts in Machine Learning Open Source Software 2018: Sustainable communities**, *Moreau* 10:25 AM

This paper discusses results and insights from the 1st ReQuEST workshop, a collective effort to promote reusability, portability and reproducibility of deep learning research artifacts within the Architecture/PL/Systems communities.
ReQuEST (Reproducible Quality-Efficient Systems Tournament) exploits the open-source Collective Knowledge framework (CK) to unify benchmarking, optimization, and co-design of deep learning systems implementations and exchange results via a live multi-objective scoreboard.
Systems evaluated under ReQuEST are diverse and include an FPGA-based accelerator, optimized deep learning libraries for x86 and ARM systems, and distributed inference in Amazon Cloud and over a cluster of Raspberry Pis.
We finally discuss limitations to our approach, and how we plan improve upon those limitations for the upcoming SysML artifact evaluation effort.

Abstract 20: **PyLissom: A tool for modeling computational maps of the visual cortex in PyTorch in Machine Learning Open Source Software 2018: Sustainable communities**, *Barijhoff* 10:25 AM

Despite impressive advancements in the last years, human vision is still much more robust than machine vision. This article presents PyLissom, a novel software library for the modeling of the cortex maps in the visual system. The software was implemented in PyTorch, a modern deep learning framework, and it allows full integration with other PyTorch modules. We hypothesize that PyLissom could act as a bridge between the neuroscience and machine learning communities, driving to advancements in both fields.

Abstract 21: **Salad: A Toolbox for Semi-supervised Adaptive Learning Across Domains in Machine Learning Open Source Software 2018: Sustainable communities**, *Schneider* 10:25 AM

We introduce salad, an open source toolbox that provides a unified implementation of state-of-the-art methods for transfer learning, semi-supervised learning and domain adaptation. In the first release, we provide a framework for reproducing, extending and combining research results of the past years, including model architectures, loss functions and training algorithms. The toolbox along with first benchmark results and further resources is accessible at domainadaptation.org.

Abstract 22: **Why every GBM speed benchmark is wrong in Machine Learning Open Source Software 2018: Sustainable communities**, *Ershov* 10:25 AM

This pdf is a proposal for a small talk about speed benchmarks of gradient boosted decision libraries on MLoss 2018 NIPS workshop. There is no proper way to do speed benchmarks of boosting libraries today. Openly-available benchmarks are misleading, often contains mistakes. We want to show, what is wrong with the benchmarks and what everyone should keep in mind, if one want to compare different libraries in a meaningful way.

Abstract 24: **Building, growing and sustaining ML communities in Machine Learning Open Source Software 2018: Sustainable communities**, *Andrews* 11:20 AM

While there are multiple research-based groups for the ML community around the
world, the adoption of these skills by a broader base of developers will require
new communities that reach beyond researchers to flourish at a large scale.

The Singapore TensorFlow & Deep Learning community is a group of over 3,000
people from different backgrounds and levels that is pushing the adoption of ML in
South-East Asia, via monthly in-person meetings, guest talks, and special events.

In the proposed short talk, we will present some of the challenges, lessons learned
and solutions found to building machine learning communities at scale.

Abstract 25: **PyMC's Big Adventure: Lessons Learned from the Development of Open-source Software for Probabilistic Programming in Machine Learning Open Source Software 2018: Sustainable communities**, *Fonnesbeck* 11:40 AM

The PyMC project is a team of open source developers devoted to the development of software for
applied Bayesian statistics and probabilistic machine learning. Broadly, our objective is to produce
Python implementations of state-of-the-art methods that can be used by a wide range of non-expert
analysts, thereby democratizing probabilistic programming and putting powerful Bayesian methods
in the hands of those who need them most: economists, astronomers, epidemiologists, ecologists,
and more. Our current product, PyMC3, allows users to implement arbitrary probabilistic models
using a high-level API that is analogous to specifying a model on a whiteboard.

Abstract 30: **DeepPavlov: An Open Source Library for Conversational AI in Machine Learning Open Source Software 2018: Sustainable communities**, *Kuratov* 03:30 PM

An open-source DeepPavlov library is specifically tailored for development of dialogue systems. The library prioritizes efficiency, modularity, and extensibility with the goal to make it easier to create dialogue systems from scratch with limited data available. It supports modular as well as end-to-end approaches to implementation of conversational agents. In DeepPavlov framework an agent consists of skills and every skill can be decomposed into components. Components are usually trainable models which solve typical NLP tasks such as intent classification, named entity recognition, sentiment analysis or pre-trained encoders for word or sentence level embeddings.
Sequence-to-sequence chit-chat, question answering or task-oriented skills can be assembled from components provided in the library. ML models implemented in DeepPavlov have performance on par with current state of the art in the field.

Abstract 31: **MXFusion: A Modular Deep Probabilistic Programming Library in Machine Learning Open Source Software 2018: Sustainable communities**, *Dai* 03:50 PM

Modularity is a key feature of deep learning libraries but has not been fully exploited for probabilistic programming.
We propose to improve modularity of probabilistic programming language by offering not only plain probabilistic distributions but also sophisticated probabilistic model such as Bayesian non-parametric models as fundamental building blocks. We demonstrate this idea by presenting a modular probabilistic programming language MXFusion, which includes a new type of re-usable building blocks, called probabilistic modules. A probabilistic module consists of a set of random variables with associated probabilistic distributions and dedicated inference methods. Under the framework of variational inference, the pre-specified inference methods of individual probabilistic modules can be transparently used for inference of the whole probabilistic model.

Abstract 32: **Flow: Open Source Reinforcement Learning for Traffic Control in Machine Learning Open Source Software 2018: Sustainable communities**, *Kheterpal* 04:10 PM

This work presents Flow, an open-source Python library enabling the application of distributed reinforcement learning (RL) to mixed-autonomy traffic control tasks, in which autonomous vehicles, human-driven vehicles, and infrastructure interact. Flow integrates SUMO, a traffic microsimulator, with RLlib, a distributed reinforcement learning library. Using Flow, researchers can programatically design new traffic networks, specify experiment configurations, and apply control to autonomous vehicles and intelligent infrastructure. We have used Flow to train autonomous vehicles to improve traffic flow in a wide variety of representative traffic scenarios; the results and scripts to generate the networks and perform training have been integrated into Flow as benchmarks. Community use of Flow is central to its design; extensibility and clarity were considered throughout development. Flow is available for open-source use at flow-project.github.io and github.com/flow-project/flow.

Abstract 33: **Reproducing Machine Learning Research on Binder in Machine Learning Open Source Software 2018: Sustainable communities**, 04:30 PM

Binder is an open-source project that lets users share interactive, reproducible science. Binder's goal is to allow researchers to create interactive versions of their code utilizing pre-existing workflows and minimal additional effort. It uses standard configuration files in software engineering to let researchers create interactive versions of code they have hosted on commonly-used platforms like GitHub. Binder's underlying technology, BinderHub, is entirely open-source and utilizes entirely open-source tools. By leveraging tools such as Kubernetes and Docker, it manages the technical complexity around creating containers to capture a repository and its dependencies, generating user sessions, and providing public URLs to share the built images with others. BinderHub combines two open-source projects within the Jupyter ecosystem: repo2docker and JupyterHub. repo2docker builds the Docker image of the git repository specified by the user, installs dependencies, and provides various front-ends to explore the image. JupyterHub then spawns and serves instances of these built images using Kubernetes to scale as needed. Because each of these pieces is open-source and uses popular tools in cloud orchestration, BinderHub can be deployed on a variety of cloud platforms, or even on your own hardware.

Learning by Instruction

Shashank Srivastava, Igor Labutov, Bishan Yang, Amos Azaria, Tom Mitchell

Room 516 AB, Sat Dec 08, 08:00 AM

Today machine learning is largely about pattern discovery and function approximation. But as computing devices that interact with us in natural language become ubiquitous (e.g., Siri, Alexa, Google Now), and as computer perceptual abilities become more accurate, they open an exciting possibility of enabling end-users to teach machines similar to the way in which humans teach one another. Natural language conversation, gesturing, demonstrating, teleoperating and other modes of communication offer a new paradigm for machine learning through instruction from humans. This builds on several existing machine learning paradigms (e.g., active learning, supervised learning, reinforcement learning), but also brings a new set of advantages and research challenges that lie at the intersection of several fields including machine learning, natural language understanding, computer perception, and HCI.

The aim of this workshop is to engage researchers from these diverse fields to explore fundamental research questions in this new area, such as:
How do people interact with machines when teaching them new learning tasks and knowledge?
What novel machine learning models and algorithms are needed to learn from human instruction?
What are the practical considerations towards building practical systems that can learn from instruction?

Schedule

08:30 AM	**Introduction**	
08:35 AM	**Teaching Machines like we Teach People**	
09:00 AM	**Mapping Navigation Instructions to Continuous Control**	*Artzi*
09:30 AM	**An Cognitive Architecture Approach to Interactive Task Learning**	*Laird*
10:00 AM	**Compositional Imitation Learning: Explaining and executing one task at a time**	*Kipf*
10:15 AM	**Learning to Learn from Imperfect Demonstrations**	*Yang, Finn*
11:00 AM	**Natural Language Supervision**	*Liang*
11:30 AM	**Control Algorithms for Imitation Learning from Observation**	*Stone*

12:00 PM	**From Language to Goals: Inverse Reinforcement Learning for Vision-Based Instruction Following**	*Fu*
12:15 PM	**Teaching Multiple Tasks to an RL Agent using LTL**	*Toro Icarte, McIlraith*
01:30 PM	**Meta-Learning to Follow Instructions, Examples, and Demonstrations**	*Levine*
02:00 PM	**Learning to Understand Natural Language Instructions through Human-Robot Dialog**	*Mooney*
02:30 PM	**The Implicit Preference Information in an Initial State**	*Shah*
02:45 PM	**Modelling User's Theory of AI's Mind in Interactive Intelligent Systems**	*Peltola*
03:30 PM	**Poster Session**	*Trimbach, Siam, Toro Icarte, Dai, McIlraith, Rahtz, Sheline, MacLellan, Lawrence, Riezler, Hadfield-Menell, Hsiao*
04:15 PM	**Assisted Inverse Reinforcement Learning**	*Singla, Devidze*
04:30 PM	**Teaching through Dialogue and Games**	*Weston*
05:00 PM	**Panel Discussion**	

Abstracts (8):

Abstract 3: **Mapping Navigation Instructions to Continuous Control in Learning by Instruction**, *Artzi* 09:00 AM

Natural language understanding in grounded interactive scenarios is tightly coupled with the actions the system generates. The action space used determines much of the complexity of the problem and the type of reasoning required. In this talk, I will describe our approach to learning to map instructions and observations to continuous control of a realistic quadcopter drone. This scenario raises new challenging questions including how can we use demonstrations to learn to bridge the gap between the high-level concepts of language and low-level robot controls? And how do we design models that continuously observe, control, and react to a rapidly changing environment? This work uses a new publicly available evaluation benchmark.

Abstract 5: **Compositional Imitation Learning: Explaining and executing one task at a time in Learning by Instruction**, *Kipf* 10:00 AM

We introduce a framework for Compositional Imitation Learning and Execution (CompILE) of hierarchically-structured behavior. CompILE learns reusable, variable-length segments of behavior from demonstration data using a novel unsupervised, fully-differentiable sequence segmentation module. These learned behaviors can then be re-composed and executed to perform new tasks. At training time, CompILE auto-encodes observed behavior into a sequence of latent codes, each corresponding to a variable-length segment in the input sequence. Once trained, our model generalizes to sequences of longer length and from environment instances not seen during training. We evaluate our model in a challenging 2D multi-task environment and show that CompILE can find correct task boundaries and event encodings in an unsupervised manner without requiring annotated demonstration data. We demonstrate that latent codes and associated behavior policies discovered by CompILE can be used by a hierarchical agent, where the high-level policy selects actions in the latent code space, and the low-level, task-specific policies are simply the learned decoders. We found that our agent could learn given only sparse rewards, where agents without task-specific policies struggle.

Abstract 6: **Learning to Learn from Imperfect Demonstrations in Learning by Instruction**, *Yang, Finn* 10:15 AM

In the standard formulation of imitation learning, the agent starts from scratch without the means to take advantage of an informative prior. As a result, the expert's demonstrations have to either be optimal, or contain a known mode of sub-optimality that could be modeled. In this work, we consider instead the problem of imitation learning from imperfect demonstrations where a small number of demonstrations containing unstructured imperfections is available. In particular, these demonstrations contain large systematic biases, or fails to complete the task in unspecified ways. Our Learning to Learn From Imperfect Demonstrations (LID) framework casts such problem as a meta-learning problem, where the agent meta-learns a robust imitation algorithm that is able to infer the correct policy despite of these imperfections, by taking advantage of an informative prior. We demonstrate the robustness of this algorithm over 2D reaching tasks, multitask door opening and picking tasks with a simulated robot arm, where the demonstration merely gestures for the intended target. Despite not seeing a demonstration that completes the task, the
agent is able to draw lessons from its prior experience--correctly inferring a policy that accomplishes the task where the demonstration fails to.

Abstract 9: **From Language to Goals: Inverse Reinforcement Learning for Vision-Based Instruction Following in Learning by Instruction**, *Fu* 12:00 PM

Reinforcement learning is a promising framework for solving control problems, but its use in practical situations is hampered by the fact that reward functions are often difficult to engineer. Specifying goals and tasks for autonomous machines, such as robots, is a significant challenge: conventionally, reward functions and goal states have been used to communicate objectives. But people can communicate objectives to each other simply by describing or demonstrating them. How can we build learning algorithms that will allow us to tell machines what we want them to do? In this work, we investigate the problem of grounding language commands as reward functions using inverse reinforcement learning, and argue that language-conditioned rewards are more transferable than language-conditioned policies to new environments. We propose language-conditioned reward learning (LC-RL), which grounds language commands as a reward function represented by a deep neural network. We demonstrate that our model learns rewards that transfer to novel tasks and environments on realistic, high-dimensional visual environments with natural language commands, whereas directly learning a language-conditioned policy leads to poor performance.

Abstract 10: **Teaching Multiple Tasks to an RL Agent using LTL in Learning by Instruction**, *Toro Icarte, McIlraith* 12:15 PM

This paper examines the problem of how to teach multiple tasks to a Reinforcement Learning (RL) agent. To this end, we use Linear Temporal Logic (LTL) as a language for specifying multiple tasks in a manner that supports the composition of learned skills. We also propose a novel algorithm that exploits LTL progression and off-policy RL to speed up learning without compromising convergence guarantees, and show that our method outperforms the state-of-the-art.

Abstract 13: **The Implicit Preference Information in an Initial State in Learning by Instruction**, *Shah* 02:30 PM

Reinforcement learning (RL) agents optimize only the specified features and are indifferent to anything left out inadvertently. This means that we must not only tell a household robot what to do, but also the much larger space of what not to do. It is easy to forget these preferences, since we are so used to having them satisfied. Our key insight is that when a robot is deployed in an environment that humans act in, the state of the environment is already optimized for what humans want. We can therefore use this implicit information from the state to fill in the blanks. We develop an algorithm based on Maximum Causal Entropy IRL and use it to evaluate the idea in a suite of proof-of-concept environments designed to show its properties. We find that information from the initial state can be used to infer both side effects that should be avoided as well as preferences for how the environment should be organized.

Abstract 14: **Modelling User's Theory of AI's Mind in Interactive Intelligent Systems in Learning by Instruction**, *Peltola* 02:45 PM

Many interactive intelligent systems, such as recommendation and information retrieval systems, treat users as a passive data source. Yet, users form mental models of systems and instead of passively providing feedback to the queries of the system, they will strategically plan their actions within the constraints of the mental model to steer the system and achieve their goals faster. We propose to explicitly account for the user's theory of the AI's mind in the user model: the intelligent system has a model of the user having a model of the intelligent system. We study a case where the system is a contextual bandit and the user model is a Markov decision process that plans based on a simpler model of the bandit. Inference in the model can be reduced to probabilistic inverse reinforcement learning, with the nested bandit model defining the transition dynamics, and is implemented using probabilistic programming. Our results show that improved performance is achieved if users can form accurate mental models that the system can capture, implying predictability of the interactive intelligent system is important not only for the user experience but also for the design of the system's statistical models.

Abstract 16: **Assisted Inverse Reinforcement Learning in Learning by Instruction**, *Singla, Devidze* 04:15 PM

We study the problem of inverse reinforcement learning (IRL) with the added twist that the learner is assisted by a helpful teacher. More formally, we tackle the following algorithmic question: How could a teacher provide an informative sequence of demonstrations to an IRL agent to speed up the learning process? We prove rigorous convergence guarantees of a new iterative teaching algorithm that adaptively chooses demonstrations based on the learner’s current performance. Extensive experiments with a car driving simulator environment show that the learning progress can be speeded up drastically as compared to an uninformative teacher.

Infer to Control: Probabilistic Reinforcement Learning and Structured Control

Leslie Kaelbling, Martin Riedmiller, Marc Toussaint, Igor Mordatch, Roy Fox, Tuomas Haarnoja

Room 516 CDE, Sat Dec 08, 08:00 AM

Reinforcement learning and imitation learning are effective paradigms for learning controllers of dynamical systems from experience. These fields have been empowered by recent success in deep learning of differentiable parametric models, allowing end-to-end training of highly nonlinear controllers that encompass perception, memory, prediction, and decision making. The aptitude of these models to represent latent dynamics, high-level goals, and long-term outcomes is unfortunately curbed by the poor sample complexity of many current algorithms for learning these models from experience.

Probabilistic reinforcement learning and inference of control structure are emerging as promising approaches for avoiding prohibitive amounts of controller–system interactions. These methods leverage informative priors on useful behavior, as well as controller structure such as hierarchy and modularity, as useful inductive biases that reduce the effective size of policy search space and shape the optimization landscape. Intrinsic and self-supervised signals can further guide the training process of distinct internal components — such as perceptual embeddings, predictive models, exploration policies, and inter-agent communication — to break down the hard holistic problem of control into more efficiently learnable parts.

Effective inference methods are crucial for probabilistic approaches to reinforcement learning and structured control. Approximate control and model-free reinforcement learning exploit latent system structure and priors on policy structure, that are not directly evident in the controller–system interactions, and must be inferred by the learning algorithm. The growing interest of the reinforcement learning and optimal control community in the application of inference methods is synchronized well with the development by the probabilistic learning community of powerful inference techniques, such as probabilistic programming, variational inference, Gaussian processes, and nonparametric regression.

This workshop is a venue for the inference and reinforcement learning communities to come together in discussing recent advances, developing insights, and future potential in inference methods and their application to probabilistic reinforcement learning and structured control. The goal of this workshop is to catalyze tighter collaboration within and between the communities, that will be leveraged in upcoming years to rise to the challenges of real-world control problems.

=== Intel AI is proud to sponsor Infer2Control @ NeurIPS 2018 === Early detection of tumors. Predicting equipment failures before they happen. Having a natural conversation with your home or car. Making retail more personal than ever. This is Artificial Intelligence powered by Intel, and companies around the globe are using it to make money, save money, and advance the future of their industry. At Intel, we’re using

decades of expertise in silicon, software, communications, memory and storage to create the new technologies that AI demands. Technologies that break barriers between data center and edge, server and network, training and inference, model and reality – maximizing the economics of AI to take data from theory to real-world success. Learn more: ai.intel.com

Schedule

08:20 AM	**Opening Remarks**	*Fox*
08:30 AM	**Control as Inference and Soft Deep RL (Sergey Levine)**	*Levine*
09:00 AM	**Unsupervised Learning of Image Embedding for Continuous Control (Carlos Florensa)**	*Florensa*
09:10 AM	**Variational Inference Techniques for Sequential Decision Making in Generative Models (Igor Kiselev)**	*Kiselev*
09:20 AM	**Probabilistic Planning with Sequential Monte Carlo (Alexandre Piché)**	*Piche*
09:30 AM	**Inference and control of rules in human hierarchical reinforcement learning (Anne Collins)**	*Collins*
10:00 AM	**Hierarchical RL: From Prior Knowledge to Policies (Shie Mannor)**	*Mannor*
10:30 AM	**-- Coffee Break 1 --**	
11:00 AM	**Off-policy Policy Optimization (Dale Schuurmans)**	*Schuurmans*
11:30 AM	**Spotlights 1**	*Huang, Cui, Mehrjou, Duan, Vikram, Wang, Goel, Hunt, Wu, Zhang, Fellows*
11:45 AM	**Poster Session 1**	*Ambert, Araki, Cao, Choi, Cui, Degrave, Duan, Fellows, Florensa, Goel, Gopalan, Huang, Hunt, Ibrahim, Ichter, Igl, Ke, Kiselev, Mahajan, Mehrjou, Pertsch, Piche, Rhinehart, Ringstrom, Russel, Rybkin, Stoica, Vikram, Wang, Wei, Wen, Wu, Wu, Xie, Zhang*
12:15 PM	**-- Lunch Break --**	
01:45 PM	**Solving inference and control problems with the same machinery (Emo Todorov)**	*Todorov*
02:15 PM	**Spotlights 2**	*Gopalan, Choi, Ringstrom, Fox, Degrave, Cao, Pertsch, Igl, Ichter*
02:30 PM	**Inference and Control of Learning Behavior in Rodents (Ryan Adams)**	*Adams*
03:00 PM	**-- Coffee Break 2 --**	
03:30 PM	**On the Value of Knowing What You Don't Know: Learning to Sample and Sampling to Learn for Robot Planning (Leslie Kaelbling)**	*Kaelbling*
04:00 PM	**Learning to Plan with Logical Automata (Brandon Araki)**	*Araki*
04:10 PM	**Tight Bayesian Ambiguity Sets for Robust MDPs (Reazul Hasan Russel)**	*Russel*
04:20 PM	**Deep Imitative Models for Flexible Inference, Planning, and Control (Nicholas Rhinehart)**	*Rhinehart*
04:30 PM	**Probabilistic Reasoning for Reinforcement Learning (Nicolas Heess)**	*Heess*
05:00 PM	**Discussion Panel: Ryan Adams, Nicolas Heess, Leslie Kaelbling, Shie Mannor, Emo Todorov (moderator: Roy Fox)**	*Adams, Heess, Kaelbling, Mannor, Todorov, Fox*
06:00 PM	**Poster Session 2**	

Relational Representation Learning

Aditya Grover, Paroma Varma, Fred Sala, Steven Holtzen, Jennifer Neville, Stefano Ermon, Chris Ré

Room 517 A, Sat Dec 08, 08:00 AM

Relational reasoning, *i.e.*, learning and inference with relational data, is key to understanding how objects interact with each other and give rise to complex phenomena in the everyday world. Well-known applications include knowledge base completion and social network analysis. Although many relational datasets are available, integrating them directly into modern machine learning algorithms and systems that rely on continuous, gradient-based optimization and make strong i.i.d. assumptions is challenging. Relational representation learning has the potential to overcome these obstacles: it enables the fusion of recent advancements like deep learning and relational reasoning to learn from high-dimensional data. Success of such methods can facilitate novel applications of relational reasoning in areas like scene understanding, visual question-answering, reasoning over chemical and biological domains, program synthesis and analysis, and decision-making in multi-agent systems.

How should we rethink classical representation learning theory for relational representations? Classical approaches based on dimensionality reduction techniques such as isoMap and spectral decompositions still serve as strong baselines and are slowly paving the way for modern methods in relational representation learning based on random walks over graphs, message-passing in neural networks, group-invariant deep architectures etc. amongst many others. How can systems be designed and potentially deployed for large scale representation learning? What are promising avenues, beyond traditional applications like knowledge base and social network analysis, that can benefit from relational representation learning?

This workshop aims to bring together researchers from both academia and industry interested in addressing various aspects of representation learning for relational reasoning.Topics include, but are not limited to:

* Algorithmic approaches. E.g., probabilistic generative models, message-passing neural networks, embedding methods, dimensionality reduction techniques, group-invariant architectures etc. for relational data
* Theoretical aspects. E.g., when and why do learned representations aid relational reasoning? How does the non-i.i.d. nature of relational data conflict with our current understanding of representation learning?
* Optimization and scalability challenges due to the inherent discreteness and curse of dimensionality of relational datasets
* Evaluation of learned relational representations
* Security and privacy challenges
* Domain-specific applications
* Any other topic of interest

Schedule

08:45 AM	**Contributed Talk 1**	
09:00 AM	**Invited Talk 1**	*Meila*
09:30 AM	**Contributed Talk 2**	*Wu*
09:45 AM	**Invited Talk 2**	*Lillicrap*
10:15 AM	**Spotlights**	*Hu, Li, Kumar, Tran, Fadel, Kuznetsova, Kang, Haji Soleimani, An, de Lara, Kumar, Weyde, Weber, Altenburger, Amizadeh, Xu, Nandwani, Guo, Pacheco, Fedus, Jaume, Yoneda, Ma, Bai, Kapicioglu, Nickel, Malliaros, Zhu, Bojchevski, Joseph, Roig, Balkir, Steenbrugge*
11:00 AM	**Invited Talk 3**	*Bruna*
11:30 AM	**Contributed Talk 3**	*Bai*
11:45 AM	**Invited Talk 4**	*Nickel*
02:00 PM	**Invited Talk 5**	*Getoor*
02:30 PM	**Contributed Talk 4**	*Csordás*
02:45 PM	**Spotlights 2**	*, Anand, Singla, Koc, Klinger, Naderi, Lyu, Amizadeh, Dwivedi, Zu, Feng, Ravindran, Pineau, Celikkanat, Venugopal*
03:30 PM	**Invited Talk 6**	*Domingos*
04:00 PM	**Panel**	*Varma, Grover, Hamilton, Hamrick, Kipf, Zitnik*
04:45 PM	**Poster Session**	
05:45 PM	**Concluding Remarks**	

Abstracts (1):

Abstract 1: **Contributed Talk 1 in Relational Representation Learning**, 08:45 AM

Adversarial training has become the de facto standard for generative modeling. While adversarial approaches have shown remarkable success in learning a distribution that faithfully recovers a reference distribution in its entirety, they are not applicable when one wishes the generated distribution to recover some —but not all— aspects of it. For example, one might be interested in modeling purely relational or topological aspects (such as cluster or manifold structure) while ignoring or constraining absolute characteristics (e.g., global orientation in Euclidean spaces). Furthermore, such absolute aspects are not available if the data is provided in an intrinsically relational form, such as a weighted graph. In this work, we propose an approach to learn generative models across such incomparable spaces that relies on the Gromov-Wasserstein distance, a notion of discrepancy that compares distributions relationally rather than absolutely. We show how the resulting framework can be used to learn distributions across spaces of different dimensionality or even different data types.

AI for social good

Margaux Luck, Tristan Sylvain, Joseph Paul Cohen, Arsene Fansi Tchango, Valentine Goddard, Aurelie Helouis, Yoshua Bengio, Sam Greydanus, Cody Wild, Taras Kucherenko, Arya Farahi, Jonnie Penn, Sean McGregor, Mark Crowley, Abhishek Gupta, Kenny Chen, Myriam Côté, Rediet Abebe

Room 517 B, Sat Dec 08, 08:00 AM

AI for Social Good

Important information
[Workshop website](https://aiforsocialgood.github.io/2018/)

[Submission website](https://cmt3.research.microsoft.com/User/Login?ReturnUrl=%2FAISG2018)

Abstract
The “AI for Social Good” will focus on social problems for which artificial intelligence has the potential to offer meaningful solutions. The problems we chose to focus on are inspired by the United Nations Sustainable Development Goals (SDGs), a set of seventeen objectives that must be addressed in order to bring the world to a more equitable, prosperous,

and sustainable path. In particular, we will focus on the following areas: health, education, protecting democracy, urban planning, assistive technology for people with disabilities, agriculture, environmental sustainability, economic inequality, social welfare and justice. Each of these themes present opportunities for AI to meaningfully impact society by reducing human suffering and improving our democracies.

The AI for Social Good workshop divides the in-focus problem areas into thematic blocks of talks, panels, breakout planning sessions, and posters. Particular emphasis is given to celebrating recent achievements in AI solutions, and fostering collaborations for the next generation of solutions for social good.

First, the workshop will feature a series of invited talks and panels on agriculture and environmental protection, education, health and assistive technologies, urban planning and social services. Secondly, it will bring together ML researchers, leaders of social impact, people who see the needs in the field as well as philanthropists in a forum to present and discuss interesting research ideas and applications with the potential to address social issues. Indeed, the rapidly expanding field of AI has the potential to transform many aspects of our lives. However, two main problems arise when attempting to tackle social issues. There are few venues in which to share successes and failures in research at the intersection of AI and social problems, an absence this workshop is designed to address by showcasing these marginalized but impactful works of research. Also, it is difficult to find and evaluate problems to address for researchers with an interest on having a social impact. We hope this will inspire the creation of new tools by the community to tackle these important problems. Also, this workshop promotes the sharing of information about datasets and potential projects which could interest machine learning researchers who want to apply their skills for social good.

The workshop also explores how artificial intelligence can be used to enrich democracy, social welfare, and justice. A focus on these topics will connect researchers to civil society organizations, NGOs, local governments, and other organizations to enable applied AI research for beneficial outcomes. Various case-studies and discussions are introduced around these themes: summary of existing AI for good projects and key issues for the future, AI's impact on economic inequality, AI approaches to social sciences, and civil society organizations.
The definition of what constitutes social good being essential to this workshop, we will have panel discussions with leading social scholars to frame how contemporary AI/ML applications relate to public and philosophical notions of social good. We also aim to define new, quantifiable, and impactful research questions for the AI/ML community. Also, we would like as an outcome of this event the creation of a platform to share data, a pact with leading tech companies to support research staff sabbaticals with social progress organizations, and the connection of researchers to on-the-ground problem owners and funders for social impact.

We invite contributions relating to any of the workshop themes or more broadly any of the UN SDGs. The models or approaches presented do not necessarily need to be of outstanding theoretical novelty, but should demonstrate potential for a strong social impact. We invite two types of submissions. First, we invite research work as short papers (4 page limit) for oral and/or poster presentation. Second, we invite two page abstracts presenting a specific solution that would, if accepted, be discussed during round-table events. The short papers should focus on past and current work, showcasing actual results and ideally demonstrated beneficial effect on society, whereas the two page abstracts could highlight ideas that have not yet been applied in practice. These are designed to foster sharing different points of view ranging from the scientific assessment of feasibility, to discussion of practical constraints that may be encountered when they are deployed, also attracting interest from philanthropists invited to the event. The workshop provides a platform for developing these two page abstracts into real projects with a platform to connect with stakeholders, scientists, and funders.

Schedule

08:30 AM	**Opening remarks**	*Bengio*
08:45 AM	**AI for agriculture, environmental protection and sustainability**	*Mwebaze*
09:15 AM	**How AI can empower the blind community**	*Koul*
09:45 AM	**Rural Infrastructure Health Monitoring System: Using AI to Increase Rural Water Supply Reliability**	*Tadesse*
10:00 AM	**Exploiting data and human knowledge for predicting wildlife poaching**	*Fang*
10:15 AM	**Inferring Work Task Automatability from AI Expert Evidence**	*Graham*
10:30 AM	**Poster session**	*Mayet, Orenstein, Greeff, Rutkowski, Yu, Marin, He, Doshi, Boix, Janssoone, Kesari, Li, Vigodny, Gordon, Moshe, Nevo, Wu, Lee, Corriveau, Lomonaco, Pruksachatkun, Zurutuza, Mehta, Pelletier, Hitti, Latessa, Glowacki, Gkantiragas, Nina, Martínez de Rituerto de Troya, Sekara, Madaio, Jang, Moreno, Houri-Yafin, Babirye*
11:00 AM	**AI's impact on economic inequality, class**	*Havens, Ajunwa*
12:00 PM	**AI's Impact on Art, Music, and Culture, featuring Yo-Yo Ma, Element AI, and Special Guests**	
01:00 PM	**Bias and fairness in AI**	*Gebru, Mitchell, Saunders*
02:00 PM	**Anti-Malaria Operations Planning and Management**	*Vigodny*

02:15 PM	**A Wearable, Biomarker-Tracking Device Platform Using Machine Learning to Predict and Prevent Opioid Relapses and Overdoses in Real Time**	*Gordon*
02:30 PM	**Machine Learning-Based Screening for Fetal Alcohol Spectrum Disorder**	*Zhang*
02:45 PM	**Enabling better pregnancy monitoring: The case of point-of-care diagnosis in fetal echocardiography**	*Patra*
03:00 PM	**Poster Session**	
03:30 PM	**The role of civil society in the age of AI: Beyond buzzwords**	*Siminyu, Tambe, Skirpan, Kim*
04:30 PM	**Improving Traffic Safety Through Video Analysis in Jakarta, Indonesia**	
04:45 PM	**Inverse Optimal Power Flow: Assessing the Vulnerability of Power Grid Data**	*Donti*
05:00 PM	**Academia, Corporations, Society, Responsibility**	*Danks, Cornebise, Di Jorio*

Abstracts (6):

Abstract 1: **Opening remarks in AI for social good**, *Bengio* 08:30 AM

Opening remarks by Yoshua Bengio.

Abstract 2: **AI for agriculture, environmental protection and sustainability in AI for social good**, *Mwebaze* 08:45 AM

A key challenge in Africa is the lack of sufficient domain experts to effectively solve the problems in health, agriculture, education, transport, etc. Artificial Intelligence with all the current advances provides hope for a solution by enabling the automation of these varied expert tasks. Africa is largely an agro-economy where majority of the people depend on agriculture for their livelihood. In this talk I will give some examples of interventions in AI being employed by the AI and Data Science research lab (AIR) in Makerere University and the UN Pulse Lab to address some of these challenges. I will talk about automating disease diagnosis in the field on smartphones and crowdsourcing surveillance data from farmers with smartphones and how these impact the livelihoods of these farmers. I will also give some examples in other fields like health where AI is impacting the livelihoods of people in Africa.

Abstract 3: **How AI can empower the blind community in AI for social good**, *Koul* 09:15 AM

In this demo based talk, Anirudh will discuss the real-life impact that AI is already bringing in the daily lives of the blind and low vision community. Learning from failures and success while converting research to product, the talk showcases a range of real-world scenarios which can benefit from both classical computer vision as well as deep learning based techniques. Separating hype from reality, it also highlights open opportunities for innovation where many traditional datasets & benchmarks do not convert to in-the-wild usage beyond fancy demos. Deep learning techniques can also help improve human-computer interaction, which might be the key to making these advances usable. The key underlying theme to recognize is how developing for differently abled communities can lead to innovation for mainstream audiences.

Abstract 9: **AI's Impact on Art, Music, and Culture, featuring Yo-Yo Ma, Element AI, and Special Guests in AI for social good**, 12:00 PM

This special program features a demo, two short presentations, and panel discussion, hosted over lunch provided by Element AI. First, renowned cellist Yo-Yo Ma will perform in a live demonstration of music and neuroinformatics, using an Emotiv wireless EEG headset operated by Emotiv President, Olivier Oullier. Next, Valérie Bécaert, Director of Research and Scientific Programs at Element AI, will present on arts and AI initiatives in Montreal, as well as a new collaborative residency launched. Additionally, Karina Kesserwan of Kesserwan Arteau will discuss her firm's work on issues relating to Indigenous and Northern communities. To conclude, all presenters will engage in a panel discussion regarding the impact of artificial intelligence on such domains as art, music, and culture.

Abstract 11: **Anti-Malaria Operations Planning and Management in AI for social good**, *Vigodny* 02:00 PM

Despite malaria’s tremendous impact on public health and economic development in Africa, and a 3 billion USD annual investment in its control, the prospect of a malaria-free Africa seems distant. In this session, I will discuss how Zzapp is harnessing AI and mobile technologies to address the key challenges to malaria elimination today. Through the use of deep learning algorithms, we are developing a planning tool that will design ever-improving intervention strategies, customized to individual urban and rural communities. To simplify the implementation of these strategies, and to ensure effective execution, we have launched a mobile app that relays instructions directly to workers in the field, and monitors them using GPS. In collaboration with several research institutes and malaria control programs in Africa, we adapt these tools to the current and future needs of malaria elimination campaigns.

Abstract 12: **A Wearable, Biomarker-Tracking Device Platform Using Machine Learning to Predict and Prevent Opioid Relapses and Overdoses in Real Time in AI for social good**, *Gordon* 02:15 PM

For individuals with substance use disorder (SUD), the propensity for returning to drug use (relapsing) is high, topping 90 percent for heroin users. The current methods and tools to counter drug addiction have been inefficient, resulting in frequent relapses and increasing overdose rates. Historically, these tools have been retrospective, with the intervention happening far too late and with a lack of focus on eliminating trigger scenarios. The goal at Behaivior is to refocus the recovery process onto the status of the individual before relapse occurs, observing both internal and external factors that affect craving states in real time.

Reinforcement Learning under Partial Observability

Joni Pajarinen, Chris Amato, Pascal Poupart, David Hsu

Room 517 C, Sat Dec 08, 08:00 AM

Reinforcement learning (RL) has succeeded in many challenging tasks such as Atari, Go, and Chess and even in high dimensional continuous domains such as robotics. Most impressive successes are in tasks where the agent observes the task features fully. However, in real world problems, the agent usually can only rely on partial observations. In real time games the agent makes only local observations; in robotics the agent has to cope with noisy sensors, occlusions, and unknown dynamics. Even more fundamentally, any agent without a full a priori world model or without full access to the system state, has to make decisions based on partial knowledge about the environment and its dynamics.

Reinforcement learning under partial observability has been tackled in the operations research, control, planning, and machine learning communities. One of the goals of the workshop is to bring researchers from different backgrounds together. Moreover, the workshop aims to highlight future applications. In addition to robotics where partial observability is a well known challenge, many diverse applications such as wireless networking, human-robot interaction and autonomous driving require taking partial observability into account.

Partial observability introduces unique challenges: the agent has to remember the past but also connect the present with potential futures requiring memory, exploration, and value propagation techniques that can handle partial observability. Current model-based methods can handle discrete values and take long term information gathering into account while model-free methods can handle high-dimensional continuous problems but often assume that the state space has been created for the problem at hand such that there is sufficient information for optimal decision making or just add memory to the policy without taking partial observability explicitly into account.

In this workshop, we want to go further and ask among others the following questions.
* How can we extend deep RL methods to robustly solve partially observable problems?
* Can we learn concise abstractions of history that are sufficient for high-quality decision-making?
* There have been several successes in decision making under partial observability despite the inherent challenges. Can we characterize problems where computing good policies is feasible?
* Since decision making is hard under partial observability do we want to use more complex models and solve them approximately or use (inaccurate) simple models and solve them exactly? Or not use models at all?
* How can we use control theory together with reinforcement learning to advance decision making under partial observability?
* Can we combine the strengths of model-based and model-free methods under partial observability?
* Can recent method improvements in general RL already tackle some partially observable applications which were not previously possible?
* How do we scale up reinforcement learning in multi-agent systems with partial observability?
* Do hierarchical models / temporal abstraction improve RL efficiency under partial observability?

Schedule

08:30 AM	**Opening Remarks**	
08:40 AM	**Joelle Pineau**	*Pineau*
09:05 AM	**Leslie Kaelbling**	*Kaelbling*
09:30 AM	**Contributed Talk 1: High-Level Strategy Selection under Partial Observability in StarCraft: Brood War**	
09:45 AM	**David Silver**	*Silver*
10:10 AM	**Contributed Talk 2: Joint Belief Tracking and Reward Optimization through Approximate Inference**	
11:00 AM	**Contributed Talk 3: Learning Dexterous In-Hand Manipulation**	
11:15 AM	**Pieter Abbeel**	*Abbeel*
11:40 AM	**Spotlights & Poster Session**	*Preiss, Grishin, Kyrki, Moreno, Narayan, Leong, Tan, Weng, Sugawara, Young, Shu, Gehring, Beirami, Amato, katt, Baisero, Kuznetsov, Humplik, Petrík*
02:00 PM	**Peter Stone**	*Stone*
02:25 PM	**Contributed Talk 4: Differentiable Algorithm Networks: Learning Wrong Models for Wrong Algorithms**	
02:40 PM	**Jilles Dibangoye**	*Dibangoye*
03:35 PM	**Anca Dragan**	*Dragan*
04:00 PM	**Panel Discussion**	
05:00 PM	**Poster Session**	

Machine Learning for Health (ML4H): Moving beyond supervised learning in healthcare

Andrew Beam, Tristan Naumann, Marzyeh Ghassemi, Matthew McDermott, Madalina Fiterau, Irene Y Chen, Brett Beaulieu-Jones, Mike Hughes, Farah Shamout, Corey Chivers, Jaz Kandola, Alexandre Yahi, Sam G Finlayson, Bruno Jedynak, Peter Schulam, Natalia Antropova, Jason Fries, Adrian Dalca, Irene Y Chen

Room 517 D, Sat Dec 08, 08:00 AM

Machine learning has had many notable successes within healthcare and medicine. However, nearly all such successes to date have been driven by supervised learning techniques. As a result, many other important areas of machine learning have been neglected and under appreciated in healthcare applications. In this workshop, we will convene a diverse set of leading researchers who are pushing beyond the boundaries of traditional supervised approaches. Attendees at the

workshop will gain an appreciation for problems that are unique to healthcare and a better understanding of how machine learning techniques, including clustering, active learning, dimensionality reduction, reinforcement learning, causal inference, and others, may be leveraged to solve important clinical problems.

This year's program will also include spotlight presentations and two poster sessions highlighting novel research contributions at the intersection of machine learning and healthcare. We will invite submission of two page abstracts (not including references) for poster contributions. Topics of interest include but are not limited to models for diseases and clinical data, temporal models, Markov decision processes for clinical decision support, multiscale data-integration, modeling with missing or biased data, learning with non-stationary data, uncertainty and uncertainty propagation, non i.i.d. structure in the data, critique of models, interpretable models, causality, model biases, transfer learning, and incorporation of non-clinical (e.g., socioeconomic) factors.

The broader goal of the NIPS 2018 Machine Learning for Health Workshop (ML4H) is to foster collaborations that meaningfully impact medicine by bringing together clinicians, health data experts, and machine learning researchers. Attendees at this workshop can also expect to broaden their network of collaborators to include clinicians and machine learning researchers who are focused on solving some of the most import problems in medicine and healthcare.

Schedule

08:30 AM	**Welcome and Opening Remarks**	
08:45 AM	**Miguel Hernan**	*Hernan*
09:15 AM	**Finale Doshi-Velez**	*Doshi-Velez*
09:45 AM	**Spotlight Presentations**	
10:30 AM	**Coffee Break I**	
11:00 AM	**Barbara Engelhardt**	*Engelhardt*
11:30 AM	**Poster Session I**	*Raghu, Jarrett, Lewis, Chaibub Neto, Mastronarde, Akbar, Chao, Zhu, Stafford, Zhang, Lu, Lee, Radhakrishnan, Falck, Shen, Neil, Roohani, Balagopalan, Marinelli, Rossman, Giesselbach, Gonzalez Ortiz, De Brouwer, Kim, Mahmood, Hsu, Ribeiro, Chunara, Orfanoudaki, Severson, Mai, Parbhoo, Haque, Prabhu, Jin, Harley, Dubourg-Felonneau, Hu, Raghu, Warrell, Johansen, Li, Järvenpää, Shukla, Tan, Fortuin, Norgeot, Hsu, Saltz, Tozzo, Miller, Ausset, Asgarian, Casale, Neuraz, Rawat, Ayer, Li, Motani, Braman, Shao, Dalca, Lee, Pierson, Ghimire, Kawai, Lahav, Goldenberg, Wu, Krishnaswamy, Pawlowski, Ukil, Zhang*
12:30 PM	**Lunch**	*Yu, Rawat, Ukil, Saib, Novikova, Hughes, Zhang, V, Kim, Taati, Ravishankar, Clifford, Kobayashi, Taati, Xu, Cheng, Cannings, Kalpathy-Cramer, Kalpathy-Cramer, Sobhani, Perros, Weng, Raykov, Lorch, Jin, Teng, Ferlaino, Rei, Beaulac, Verma, Keller, Cunningham, Evers, Rodriguez, Satone, Liu, Yasodhara, Tison, Solamen, He, Ladhania, Shi, Hamid, Mashouri, Hwang, Park, Chen, Kaur, Blalock, Wiberg, bhatia, Yu, LI, Sakuma, Ding, Babier, Cai, Pratap, O'Connor, Nie, Kang, Covert, Wang, Luo, Yeung, Boag, Tachikawa, Saltz, Lahav, Lee, Teasley, Kamp, Patel, Mhasawade, Samarin, Uchimido, Khalvati, Cruz, Symul, Nabulsi, Mihailescu, Picard*
01:30 PM	**Poster Session II**	
02:30 PM	**Katherine Heller**	*Heller*
03:00 PM	**Coffee Break II**	
03:30 PM	**Paul Varghese**	*varghese*
04:00 PM	**Suchi Saria**	*Saria*
04:30 PM	**Panel**	

Second Workshop on Machine Learning for Creativity and Design

Luba Elliott, Sander Dieleman, Rebecca Fiebrink, Jesse Engel, Adam Roberts, Tom White

Room 518, Sat Dec 08, 08:00 AM

Over the past few years, generative machine learning and machine creativity have continued grow and attract a wider audience to machine learning. Generative models enable new types of media creation across images, music, and text - including recent advances such as sketch-rnn and the Universal Music Translation Network. This one-day workshop broadly explores issues in the applications of machine learning to creativity and design. We will look at algorithms for generation and creation of new media and new designs, engaging researchers building the next generation of generative models (GANs, RL, etc). We investigate the social and cultural impact of these new models, engaging researchers from HCI/UX communities and those using machine learning to develop new creative tools. In addition to covering the technical advances, we also address the ethical concerns ranging from the use of biased datasets to building tools for better “DeepFakes”. Finally, we’ll hear from some of the artists and musicians who are adopting machine learning including deep learning and reinforcement learning as part of their own artistic process. We aim to balance the technical issues and challenges of applying the latest generative models to creativity and design with philosophical and cultural issues that surround this area of research.

Background
In 2016, DeepMind’s AlphaGo made two moves against Lee Sedol that were described by the Go community as “brilliant,” “surprising,” “beautiful,” and so forth. Moreover, there was little discussion surrounding the fact that these very creative moves were actually made by a machine; it was enough that they were great examples of go playing. At the same time, the general public showed more concern for other applications of generative models. Algorithms that allow for convincing voice style transfer (Lyrebird) or puppet-like video face control (Face2Face) have raised ethical concerns that generative ML will be used to make convincing forms of fake news

Balancing this, the arts and music worlds have positively embraced generative models. Starting with DeepDream and expanding with image and video generation advances (e.g. GANs) we’ve seen lots of new and interesting art and music technologies provided by the machine learning community. We’ve seen research projects like Google Brain’s Magenta, Sony CSL’s FlowMachines and IBM’s Watson undertake collaborations and attempt to build tools and ML models for use by these communities.

Research
Recent advances in generative models enable new possibilities in art and music production. Language models can be used to write science fiction film scripts (Sunspring), theatre plays (Beyond the Fence) and even replicate the style of individual authors (Deep Tingle). Generative models for image and video allow us to create visions of people, places and things that resemble the distribution of actual images (GANs etc). Sequence modelling techniques have opened up the possibility of generating realistic musical scores (MIDI generation etc) and even raw audio that resembles human speech and physical instruments (DeepMind’s WaveNet, MILA’s Char2Wav and Google’s NSynth). In addition, sequence modelling allows us to model vector images to construct stroke-based drawings of common objects according to human doodles (sketch-rnn). Lately, domain transfer techniques (FAIR’s Universal Music Translation Network) have enabled the translation of music across musical instruments, genres, and styles.

In addition to field-specific research, a number of papers have come out that are directly applicable to the challenges of generation and evaluation such as learning from human preferences (Christiano et al., 2017) and CycleGAN. The application of Novelty Search (Stanley), evolutionary complexification (Stanley - CPPN, NEAT, Nguyen et al - Plug&Play GANs, Innovation Engine) and intrinsic motivation (Oudeyer et al 2007, Schmidhuber on Fun and Creativity) techniques, where objective functions are constantly evolving, is still not common practice in art and music generation using machine learning.

Another focus of the workshop is how to better enable human influence over generative models. This could include learning from human preferences, exposing model parameters in ways that are understandable and relevant to users in a given application domain (e.g., similar to Morris et al. 2008), enabling users to manipulate models through changes to training data (Fiebrink et al. 2011), allowing users to dynamically mix between multiple generative models (Akten & Grierson 2016), or other techniques. Although questions of how to make learning algorithms controllable and understandable to users are relatively nascent in the modern context of deep learning and reinforcement learning, such questions have been a growing focus of work within the human-computer interaction community (e.g., examined in a CHI 2016 workshop on Human-Centred Machine Learning), and the AI Safety community (e.g. Christiano et al. 22017, using human preferences to train deep reinforcement learning systems). Such considerations also underpin the new Google “People + AI Research” (PAIR) initiative.

Artists and Musicians
All the above techniques improve our capabilities of producing text, sound and images and have helped popularise the themes of machine learning and artificial intelligence in the art world with a number of art exhibitions (ZKM’s Open Codes, Frankfurter Kunstverein’s I am here to learn, NRW Forum’s Pendoran Vinci) and media art festivals (Impakt Festival 2018 Algorithmic Superstructures, Retune 2016) dedicated to the topic.

Art and music that stands the test of time however requires more than generative capabilities. Recent research includes a focus on novelty in creative adversarial networks (Elgammal et al., 2017) and considers how generative algorithms can integrate into human creative processes, supporting exploration of new ideas as well as human influence over generated content (Atken & Grierson 2016a, 2016b). Artists including Mario Klingemann, Roman Lipski, Mike Tyka, and Memo Akten have further contributed to this space of work by creating artwork that compellingly demonstrates capabilities of generative algorithms, and by publicly reflecting on the artistic affordances of these new tools. Other artists such as Mimi Onuoha, Caroline Sinders, and Adam Harvey have explored the ethical dimensions of machine learning technologies, reflecting on the issues of biased datasets and facial recognition.

The goal of this workshop is to bring together researchers interested in advancing art and music generation to present new work, foster collaborations and build networks.

In this workshop, we are particularly interested in how the following can be used in art and music generation: reinforcement learning, generative adversarial networks, novelty search and evaluation as well as learning from user preferences. We welcome submissions of short papers, demos and extended abstracts related to the above.

Like last year, there will be an open call for a display of artworks incorporating machine learning techniques. The exhibited works serve as a separate and more personal forum for collecting and sharing some of the latest creative works incorporating machine learning techniques with the NIPS community.

Schedule

08:30 AM	**Introduction**	
08:45 AM	**Kenneth Stanley**	*Stanley*
09:15 AM	**Yaroslav Ganin**	*Ganin*
09:45 AM	**David Ha**	*Ha*
10:15 AM	**AI art gallery overview**	*Elliott*
11:00 AM	**Yaniv Taigman**	*Taigman*
11:30 AM	**Performing Structured Improvisations with Pre-existing Generative Musical Models**	
11:45 AM	**Legend of Wrong Mountain: Full Generation of Traditional Chinese Opera Using Multiple Machine Learning Algorithms**	*Huang, Sun, Jiang*
12:00 PM	**Lunch**	
01:30 PM	**Poster Session 1**	*Casey, Raffel, Simon, Li, Saunders, Gemeinboeck, Kang, Ge, Hawthorne, Huang, Su, Chu, Akten, Damani, Gupta, Singh, Hutchings*
02:30 PM	**Allison Parrish**	*Parrish*
03:00 PM	**Break**	
03:30 PM	**TimbreTron: A WaveNet(CycleGAN(CQT(Audio))) Pipeline for Musical Timbre Transfer**	*Huang, Anil, Bao*
03:45 PM	**Infilling Piano performances**	*Ippolito*
04:00 PM	**Improvised Robotic Design with Found Objects**	*Maekawa*
04:15 PM	**SpaceSheets: Interactive Latent Space Exploration through a Spreadsheet Interface**	*White*
04:30 PM	**Runway: Adding artificial intelligence capabilities to design and creative platforms**	*Valenzuela, Germanidis, Matamala*
04:45 PM	**Open Discussion**	
05:15 PM	**AI art show**	*Epstein, Chaney, Champandard, Kogan, Davis*
05:15 PM	**Poster Session 2**	*Gero, Zhou, Yu, Gao, Donahue, Li, KWON, Hutchings, Martin, Kang, Kitamoto, Jiang, Sun, Schmitt, Attarian, Lamb, CLANUWAT, Martino, Grimm, Jetchev*

Machine Learning for Molecules and Materials

Jose Miguel Hernández-Lobato, Klaus-Robert Müller, Brooks Paige, Matt Kusner, Stefan Chmiela, Kristof Schütt

Room 519, Sat Dec 08, 08:00 AM

Website http://www.quantum-machine.org/workshops/nips2018/

The success of machine learning has been demonstrated time and time again in classification, generative modelling, and reinforcement learning. This revolution in machine learning has largely been in domains with at least one of two key properties: (1) the input space is continuous, and thus classifiers and generative models are able to smoothly model unseen data that is ‘similar’ to the training distribution, or (2) it is trivial to generate data, such as in controlled reinforcement learning settings such as Atari or Go games, where agents can re-play the game millions of times.
Unfortunately there are many important learning problems in chemistry, physics, materials science, and biology that do not share these attractive properties, problems where the input is molecular or material data.

Accurate prediction of atomistic properties is a crucial ingredient toward rational compound design in chemical and pharmaceutical industries. Many discoveries in chemistry can be guided by screening large databases of computational molecular structures and properties, but high level quantum-chemical calculations can take up to several days per molecule or material at the required accuracy, placing the ultimate achievement of in silico design out of reach for the foreseeable future. In large part the current state of the art for such problems is the expertise of individual researchers or at best highly-specific rule-based heuristic systems. Efficient methods in machine learning, applied to the prediction of atomistic properties as well as compound design and crystal structure prediction, can therefore have pivotal impact in enabling chemical discovery and foster fundamental insights.

Because of this, in the past few years there has been a flurry of recent work towards designing machine learning techniques for molecule and material data [1-38]. These works have drawn inspiration from and made significant contributions to areas of machine learning as diverse as learning on graphs to models in natural language processing. Recent advances enabled the acceleration of molecular dynamics simulations, contributed to a better understanding of interactions within quantum many-body system and increased the efficiency of density based quantum mechanical modeling methods. This young field offers unique opportunities for machine learning researchers and practitioners, as it

presents a wide spectrum of challenges and open questions, including but not limited to representations of physical systems, physically constrained models, manifold learning, interpretability, model bias, and causality.

The goal of this workshop is to bring together researchers and industrial practitioners in the fields of computer science, chemistry, physics, materials science, and biology all working to innovate and apply machine learning to tackle the challenges involving molecules and materials. In a highly interactive format, we will outline the current frontiers and present emerging research directions. We aim to use this workshop as an opportunity to establish a common language between all communities, to actively discuss new research problems, and also to collect datasets by which novel machine learning models can be benchmarked. The program is a collection of invited talks, alongside contributed posters. A panel discussion will provide different perspectives and experiences of influential researchers from both fields and also engage open participant conversation. An expected outcome of this workshop is the interdisciplinary exchange of ideas and initiation of collaboration.

Call for papers:

The 1 day NIPS 2018 Workshop on Machine Learning for Molecules and Materials is calling for contributions on theoretical models, empirical studies, and applications of machine learning for molecules and materials. We also welcome challenge papers on possible applications or datasets. Topics of interest (though not exhaustive) include: chemoinformatics, applications of deep learning to predict molecular properties, drug-discovery and material design, retrosynthesis and synthetic route prediction, modeling and prediction of chemical reaction data, and the analysis of molecular dynamics simulations. We invite submissions that either address new problems and insights for chemistry and quantum physics or present progress on established problems. The workshop includes a poster session, giving the opportunity to present novel ideas and ongoing projects. Submissions should be no longer than 10 pages in any format. Please email all submissions to: nips2018moleculesworkshop@gmail.com

References

[1] Behler, J., Lorenz, S., Reuter, K. (2007). Representing molecule-surface interactions with symmetry-adapted neural networks. J. Chem. Phys., 127(1), 07B603.
[2] Behler, J., Parrinello, M. (2007). Generalized neural-network representation of high-dimensional potential-energy surfaces. Phys. Rev. Lett., 98(14), 146401.
[3] Kang, B., Ceder, G. (2009). Battery materials for ultrafast charging and discharging. Nature, 458(7235), 190.
[4] Bartók, A. P., Payne, M. C., Kondor, R., Csányi, G. (2010). Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. Phys. Rev. Lett., 104(13), 136403.
[5] Behler, J. (2011). Atom-centered symmetry functions for constructing high-dimensional neural network potentials. J. Chem. Phys, 134(7), 074106.
[6] Behler, J. (2011). Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. Phys. Chem. Chem. Phys., 13(40), 17930-17955.
[7] Rupp, M., Tkatchenko, A., Müller, K.-R., von Lilienfeld, O. A. (2012). Fast and accurate modeling of molecular atomization energies with machine learning. Phys. Rev. Lett., 108(5), 058301.
[8] Snyder, J. C., Rupp, M., Hansen, K., Müller, K.-R., Burke, K. (2012). Finding density functionals with machine learning. Phys. Rev. Lett., 108(25), 253002.
[9] Montavon, G., Rupp, M., Gobre, V., Vazquez-Mayagoitia, A., Hansen, K., Tkatchenko, A., Müller, K.-R., von Lilienfeld, O. A. (2013). Machine learning of molecular electronic properties in chemical compound space. New J. Phys., 15(9), 095003.
[10] Hansen, K., Montavon, G., Biegler, F., Fazli, S., Rupp, M., Scheffler, M., Tkatchenko, A., Mu■ller, K.-R. (2013). Assessment and validation of machine learning methods for predicting molecular atomization energies. J. Chem. Theory Comput., 9(8), 3404-3419.
[11] Bartók, A. P., Kondor, R., Csányi, G. (2013). On representing chemical environments. Phys. Rev. B, 87(18), 184115.
[12] Schütt K. T., Glawe, H., Brockherde F., Sanna A., Müller K.-R., Gross E. K. U. (2014). How to represent crystal structures for machine learning: towards fast prediction of electronic properties. Phys. Rev. B., 89(20), 205118.
[13] Ramsundar, B., Kearnes, S., Riley, P., Webster, D., Konerding, D., Pande, V. (2015). Massively multitask networks for drug discovery. arXiv preprint arXiv:1502.02072.
[14] Rupp, M., Ramakrishnan, R., & von Lilienfeld, O. A. (2015). Machine learning for quantum mechanical properties of atoms in molecules. J. Phys. Chem. Lett., 6(16), 3309-3313.
[15] V. Botu, R. Ramprasad (2015). Learning scheme to predict atomic forces and accelerate materials simulations., Phys. Rev. B, 92(9), 094306.
[16] Hansen, K., Biegler, F., Ramakrishnan, R., Pronobis, W., von Lilienfeld, O. A., Müller, K.-R., Tkatchenko, A. (2015). Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space. J. Phys. Chem. Lett, 6(12), 2326-2331.
[17] Alipanahi, B., Delong, A., Weirauch, M. T., Frey, B. J. (2015). Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. ■Nat. Biotechnol., 33(8), 831-838.
[18] Duvenaud, D. K., Maclaurin, D., Aguilera-Iparraguirre, J., Gomez-Bombarelli, R., Hirzel, T., Aspuru-Guzik, A., Adams, R. P. (2015). Convolutional networks on graphs for learning molecular fingerprints. NIPS, 2224-2232.
[19] Faber F. A., Lindmaa A., von Lilienfeld, O. A., Armiento, R. (2016). Machine learning energies of 2 million elpasolite (A B C 2 D 6) crystals. Phys. Rev. Lett., 117(13), 135502.
[20] Gomez-Bombarelli, R., Duvenaud, D., Hernandez-Lobato, J. M., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P., Aspuru-Guzik, A. (2016). Automatic chemical design using a data-driven continuous representation of molecules. arXiv preprint arXiv:1610.02415.
[21] Wei, J. N., Duvenaud, D, Aspuru-Guzik, A. (2016). Neural networks for the prediction of organic chemistry reactions. ACS Cent. Sci., 2(10), 725-732.
[22] Sadowski, P., Fooshee, D., Subrahmanya, N., Baldi, P. (2016). Synergies between quantum mechanics and machine learning in reaction prediction. J. Chem. Inf. Model., 56(11), 2125-2128.
[23] Lee, A. A., Brenner, M. P., Colwell L. J. (2016). Predicting protein-ligand affinity with a random matrix framework. Proc. Natl. Acad. Sci., 113(48), 13564-13569.
[24] Behler, J. (2016). Perspective: Machine learning potentials for atomistic simulations. J. Chem. Phys., 145(17), 170901.
[25] De, S., Bartók, A. P., Csányi, G., Ceriotti, M. (2016). Comparing molecules and solids across structural and alchemical space. Phys. Chem. Chem. Phys., 18(20), 13754-13769.
[26] Schütt, K. T., Arbabzadah, F., Chmiela, S., Müller, K.-R., Tkatchenko, A. (2017). Quantum-chemical insights from deep tensor neural networks. Nat. Commun., 8, 13890.
[27] Segler, M. H., Waller, M. P. (2017). Neural■symbolic machine learning for retrosynthesis and reaction prediction. ■Chem. Eur. J.,

23(25), 5966-5971.
[28] Kusner, M. J., Paige, B., Hernández-Lobato, J. M. (2017). Grammar variational autoencoder. arXiv preprint arXiv:1703.01925.
[29] Coley, C. W., Barzilay, R., Jaakkola, T. S., Green, W. H., Jensen K. F. (2017). Prediction of organic reaction outcomes using machine learning. ACS Cent. Sci., 3(5), 434-443.
[30] Altae-Tran, H., Ramsundar, B., Pappu, A. S., Pande, V. (2017). Low data drug discovery with one-shot learning. ACS Cent. Sci., 3(4), 283-293.
[31] Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., Dahl, G. E. (2017). Neural message passing for quantum chemistry. arXiv preprint arXiv:1704.01212.
[32] Chmiela, S., Tkatchenko, A., Sauceda, H. E., Poltavsky, Igor, Schütt, K. T., Müller, K.-R. (2017). Machine learning of accurate energy-conserving molecular force fields. Sci. Adv., 3(5), e1603015.
[33] Ju, S., Shiga T., Feng L., Hou Z., Tsuda, K., Shiomi J. (2017). Designing nanostructures for phonon transport via bayesian optimization. Phys. Rev. X, 7(2), 021024.
[34] Ramakrishnan, R, von Lilienfeld, A. (2017). Machine learning, quantum chemistry, and chemical space. Reviews in Computational Chemistry, 225-256.
[35] Hernandez-Lobato, J. M., Requeima, J., Pyzer-Knapp, E. O., Aspuru-Guzik, A. (2017). Parallel and distributed Thompson sampling for large-scale accelerated exploration of chemical space. arXiv preprint arXiv:1706.01825.
[36] Smith, J., Isayev, O., Roitberg, A. E. (2017). ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. Chem. Sci., 8(4), 3192-3203.
[37] Brockherde, F., Li, L., Burke, K., Müller, K.-R. By-passing the Kohn-Sham equations with machine learning. Nat. Commun., 8, 872.
[38] Schütt, K. T., Kindermans, P. J., Sauceda, H. E., Chmiela, S., Tkatchenko, A., Müller, K. R. (2017). SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. NIPS 30.

Schedule

08:00 AM	**Contributed Work**	*Moustafa Dieb, Balu, Khasahmadi, Shah, Knyazev, Das, Goh, Derevyanko, De Fabritiis, Hagawa, Ingraham, Belanger, Song, Nicoli, Skalic, Wu, Gebauer, Jørgensen, Griffiths, Liu, Mysore, Chieu, Schwaller, Olsthoorn, Cristescu, Tseng, Ryu, Drori, Yang, Sanyal, Boukouvalas, Bedi, Paul, Ghosal, Bash, Fare, Ren, Oskooei, Wong, Sinz, Gaudin, Jin, Leu*
08:40 AM	**Invited Talk Session 1**	*Noe*
11:00 AM	**Invited Talk Session 2**	*Marks, Isayev, Smidt, Thomas*
02:00 PM	**Invited Talk Session 3**	*Tkatchenko, Jaakkola, Wei*
03:30 PM	**Invited Talk Session 4**	*Clementi*

Emergent Communication Workshop

Jakob Foerster, Angeliki Lazaridou, Ryan Lowe, Igor Mordatch, Douwe Kiela, Kyunghyun Cho

Room 524, Sat Dec 08, 08:00 AM

Abstract
Communication is one of the most impressive human abilities. The question of how communication arises has been studied for many decades, if not centuries. However, due to computational and representational limitations, past work was restricted to low dimensional, simple observation spaces. With the rise of deep reinforcement learning methods, this question can now be studied in complex multi-agent settings, which has led to flourishing activity in the area over the last two years. In these settings agents can learn to communicate in grounded multi-modal environments and rich communication protocols emerge.

Last year at NIPS 2017 we successfully organized the inaugural workshop on emergent communication (https://sites.google.com/site/emecom2017/). We had a number of interesting submissions looking into the question of how language can emerge using evolution (see this Nature paper that was also presented at the workshop last year, https://www.nature.com/articles/srep34615) and under what conditions emerged language exhibits compositional properties, while others explored specific applications of agents that can communicate (e.g., answering questions about textual inputs, a paper presented by Google that was subsequently accepted as an oral presentation at ICLR this year, etc.).

While last year's workshop was a great success, there are a lot of open questions. In particular, the more challenging and realistic use cases come from situations where agents do not have fully aligned interests and goals, i.e., how can we have credible communication amongst self-interested agents where each agent maximizes its own individual rewards rather than a joint team reward? This is a new computational modeling challenge for the community and recent preliminary results (e.g. "Emergent Communication through Negotiation", Cao et al., ICLR 2018.) reinforce the fact that it is no easy feat.

Since machine learning has exploded in popularity recently, there is a tendency for researchers to only engage with recent machine learning literature, therefore at best reinventing the wheel and at worst recycling the same ideas over and over, increasing the probability of being stuck in local optima. For these reasons, just like last year, we want to take an interdisciplinary approach on the topic of emergent communication, inviting researchers from different fields (machine learning, game theory, evolutionary biology, linguistics, cognitive science, and programming languages) interested in the question of communication and emergent language to exchange ideas.

This is particularly important for this year's focus, since the question of communication in general-sum settings has been an active topic of research in game theory and evolutionary biology for a number of years, while it's a nascent topic in the area of machine learning.

Schedule

08:45 AM	**Opening Remarks**	
09:00 AM	**Invited Talk 1: Luc Steels**	*Steels*
09:30 AM	**Contributed Talks I and II**	

10:00 AM	**Invited Talk 2: Simon Huttegger**	*Huttegger*
11:00 AM	**Invited Talk 3: He He**	*He*
11:30 AM	**Spotlight Talks I**	*Leni, Spranger, Bogin, Steinert-Threlkeld, Tomlin, Li, Noukhovitch, Jain, Lee, Kuo, Correa, Hausman*
12:40 PM	**Poster Session and Lunch**	
02:45 PM	**Contributed Talk 3**	
03:00 PM	**Contributed Talk 4**	
03:15 PM	**Invited Talk 5: Nando de Freitas**	*de Freitas*
03:45 PM	**Coffee Break**	
04:15 PM	**Invited Talk 4: Jeff Clune**	
04:45 PM	**Panel Discussion**	
05:50 PM	**Closing Remarks**	