

시장심리 7단계로 계량화... 주식투자 보조지표로



강세훈 경제부기자

【서울=뉴스시스】 = 250개 국내 주요 상장 기업에 대해 온라인과 소셜미디어 상에서의 평가를 총망라해 분석, 수치화한 빅데이터 시장심리지수(Market Sentiment Index · MSI)가 14일 첫선을 보인다.

주식시장에서 기존의 정량적 데이터 분석의 한계를 넘은 정성적 데이터 분석의 첫 시도로, 개별 기업에 대한 호감도와 비호감도, 투자 의향의 심리적 지표를 주가 예측에 적용한 만큼 강력한 투자 보조지표로 활용될 수 있을 것으로 기대를 모으고 있다.

14일 공감언론 '뉴스시스'와 자본시장의 IT파트너인 '코스콤'은 '빅데이터 정보 분석·활용'에 관한 상호업무협력 협약(MOU)을 맺고, 빅데이터를 통한 상장사 250곳에 대한 주가흐름을 분석, 예측하는 참고지표 '빅데이터 MSI(시장심리지수)' 제공을 시작한다고 밝혔다.

뉴스시스와 코스콤은 온라인상에서 언급되는 개별 종목에 대한 소셜 데이터를 준 실시간(1시간 단위)으로 분석해 주식투자에 활용 가능한 정보를 제공할 계획인데, 이달 중에는 시범 서비스를 실시하고 4월부터 본격 서비스에 들어갈 예정이다.

이 서비스는 투자 참고지표이며, 투자 결정은 당사자가 해당 종목에 대한 실적과 업황, 글로벌 경제 등 대내외적인 다양한 요소들을 고려해 자기 책임 아래 이뤄져야 한다는 점을 밝혀둔다.

빅데이터 MSI는 특정 종목에 대한 트위터 등 SNS(소셜네트워크서비스), 카페, 블로그, 증권사 사이트, 포털, 뉴스 등에서 관련 문장·단어 등을 포괄적으로 수집한 후 의미의 최소 단위가 되는 '형태소' 분석과 '긍정 부정어' 분석을 거쳐 해당 기업에 대한 호감·비호감도 및 투자의향을 종합적으로 나타내는 지표다.

이 지표는 독자와 네티즌, 투자자들이 해당기업에 대한 시장심리 상태를 직관적으로 파악할 수 있도록 산출된 지표를 다시 1레벨(매우 나쁨)에서 7레벨(매우 좋음)까지 7단계로 세분화하는 방식으로 제시된다.

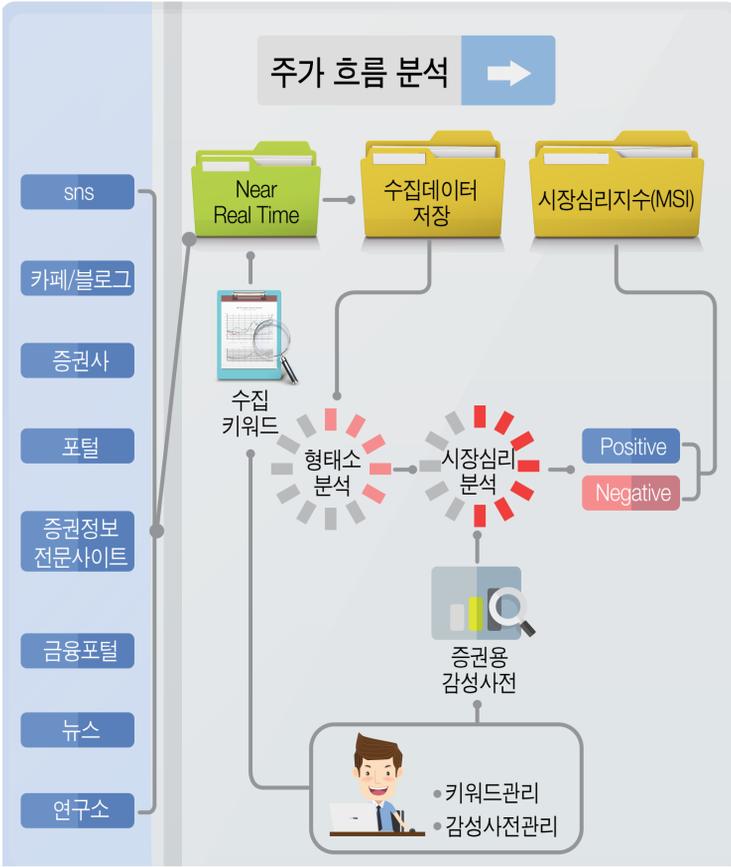
MSI 분석 대상 종목은 코스피200종목과 코스닥50종목이다.

이를 위해 코스콤은 250개 종목별로 해당 종목과 관련된 '기쁘다' '좋다' 등의 형용사뿐만 아니라 제품과 브랜드, 이미지, 인물 등의 문장·단어들이 주가에 미치는 상관 관계를 분석, 20만 단어가량의 '감성 사전(emotion Dictionary)'을 만들었다.

감성사전을 통해 해당 종목의 실시간 점수를 산출해내고 이를 최근 1년간의 흐름과 비교해 시장심리지수 레벨로 제시한 것이다.

예컨대 'OO전자의 부채율이 상승했다'는 문장은 해당 종목의 주가에 부정적인 영향을 줄 수 있는 내용으로 분류되고, 'OO중공업의 매출이 상승했다'는 문장은 주가에 긍정적인 영향을 줄 수 있는 내용으로 분류된다.

시장심리지수 레벨이 가장 낮은 수준인 1단계에서



증시 정량분석 한계 넘어선 정성적 데이터 분석은 최초 SNS, 카페, 블로그 등 실시간 정보 수치로 변환해 제공 레벨 1▶레벨 7로 바뀌면 투자자 반응 부정▶긍정 전환

가장 높은 수준인 7단계로 바뀌면 해당 종목에 대해 온라인 이용자나 투자자들의 반응이 부정에서 긍정으로 바뀌고 있다는 뜻으로 해석할 수 있는 것이다.

지금까지 투자자들이 매매를 결정하는 데 있어 투자 보조지표로 호가나 수급, 차트 등 정량적 데이터만을 사용해 왔다.

하지만 빅데이터 MSI를 통한 분석은 인간 행동을 실시간으로 반영한 정성적 데이터로 주식시장에서 한 단계 진화된 투자 참고지표로 활용될 수 있을 것으로 기대된다.

코스콤 강태홍 기술연구소장(상무)은 “빅데이터 주가 분석은 기업의 재무적 상황이나 기술적 분석 위주의 예측과 달리 개별 종목에 대해 실시간으로 변하는 투자자들의 심리를 직관적으로 평가할 수 있는 투자 참고지표”라고 설명했다.

주가 흐름에 영향을 주는 데는 헤아릴 수 없는 많은 지표가 작용한다는 게 정설이다. 단순히 기업의 실적

이나 업황만이 아니라 북한의 미사일 발사, 중국의 위안화 평가 절하 등 대내외적인 수많은 요소가 한데 엉켜 예측의 한계를 번번이 벗어난다.

하지만 온라인 상에서 네티즌이나 투자자들이 특정 종목이나 기업에 대해 어떤 심리상태에 있는지, 다시 말해 시장심리지수(MSI)가 어느 정도 인지를 파악하는 것과 그렇지 않은 것은 상당한 차이가 있으며, 이때문에 MSI는 투자의 결정적 지표는 아니더라도 유용한 참고지표로 의미가 있다.

그는 “주식시장에도 기존의 수량적 데이터 분석의 한계를 넘어 개별 기업에 대한 호감 및 투자심리까지 고려하는 빅데이터 분석의 시대가 이제 열린 것”이라며 “투자자들이 주식을 선택하는 데 있어 참고 지표가 등장하면서 주식 투자문화를 바꾸는 데 기여할 것으로 기대된다”고 밝혔다.

kangse@newsis.com

숫자만 보고 투자? 매번 실패 온라인 집단감성 투자 활용해야 증시 투자환경에 큰 변화 온다

【서울=뉴스시스】 강세훈 기자 = 기업의 가치, 즉 주가는 기본적으로 경제논리에 바탕을 둔다. 기업의 목표는 이익 추구이고, 실적은 곧 주가로 연결된다. 실적이 좋은 기업의 주가는 오르고 좋지 않은 기업 주가는 떨어진다.

이 때문에 기업은 성장을 위해 글로벌 경제, 한국 경제, 해당 산업 업황에 민감할 수밖에 없다. 투자자들 역시 기업의 실적을 무엇보다 중요하게 생각하는 이유다.

하지만 경제논리만으로 주식시장을 설명하기 어려운 부분도 많다. 세계 각국들이 수출을 주도하기 위해서 환율 전쟁을 일으키고, 글로벌 헤지펀드들이 막대한 자본으로 주식시장을 뒤흔들기도 한다.

개별 기업도 마찬가지다. 규모가 큰 기업은 기초체력(펀더멘털)에 의해 움직이는 경향이 강하지만 덩치가 작은 기업일수록 머니 게임, 즉 수급에 의해 주가가 결정되는 경우가 많다.

이뿐만 아니다. 오너나 CEO의 이미지 때문이든, 사회공헌을 많이 한 덕분이든 특정 기업을 좋게 보고 주식을 사려는 사람이 많으면 주가가 오르고, 기업을 안 좋게 보고 팔려고 하는 사람이 많으면 주가는 떨어진다. 다시 말해 사람들의 심리에 의해 좌우되는 측면도 상당하다는 얘기다.

이 때문에 호재든, 악재든 다른 사람보다 먼저 이런 상황을 인지한다는 건 주식시장에서는 승패와 직결된다. 또한 개별종목에 대한 다른 사람들의 전반적인 평가나 키워드가 무엇인지를 안다는 것은 주식시장에서 안정적 수익을 올리는 데 중요한 정보가 될 것이다.

빅데이터 환경의 도래로 실제로 이런 일이 가능해졌다. 요즘 사람들은 일상적으로 컴퓨터와 스마트폰, 소셜네트워크서비스(SNS)를 통해 엄청난 양의 디지털 흔적을 남기고 있다. 이를 순식간에 처리할 수 있는 기술이 생겨나면서 빅데이터 시대가 열렸다. 과거에는 불가능했던 일이다.

빅데이터는 스포츠, 미디어, 광고, 패션, 의료 등 각 분야에서 이미 활용되고 있지만 아직 결وما 수준이다. 특히 개별 기업에 대한 호감도와 비호감도, 투자심리까지 파악해 이를 주식시장에 본격 응용한 사례는 없었다.

뉴스시스와 코스콤이 공동으로 내놓은 '빅데이터 MSI(시장심리지수·Market Sentiment Index·MSI)'는 온라인상의 집단감성에 기반한 투자 매매 원리를 활용한다. 전체 시장의 감성뿐 아니라 개별 종목의 감성까지 읽어낸다. 빅데이터를 주식시장에 적용해 주가에 예측에 활용하는 첫 시도인 셈이다.

특히 단순히 종목에 대한 '좋음' '나쁨' '슬픔' 등의 감성표현뿐 아니라 해당 종목과 관련된 상품, 브랜드, 인물 등의 정보까지 수집해 주가에 긍정적인 영향을 미치는지, 부정적인 영향을 미치는지를 감성사전으로 분류하고 이를 준 실시간(1시간 단위 집계)으로 제공한다.

이 때문에 빅데이터 MSI는 주식시장에 새로운 투자 문화를 만드는 계기가 될 것으로 기대를 모으고 있다.

투자자들이 주식을 선택하는 방식으로 '답다운 방식'이 주로 쓰인다. 거시경제, 산업분석을 통해 유망산업을 찾아낸 뒤 종목의 기본적인 정보를 살펴보는 분석 방식이다.

kangse@newsis.com

뉴스시스-코스콤 빅데이터 MSI 소셜 빅데이터 첫 상용화 시도

【서울=뉴스시스】 심동준 기자 = 숫자로 구성된 자본 시장은 모든 현상이 수치로 드러나는 일종의 계량화된 세계이다. 하지만 시장은 숫자만으로는 움직이지 않는다.

유효 수요와 재정 정책을 강조한 영국의 경제학자 존 메이너드 케인스는 자본시장을 두고 “인기투표로 승자가 결정되는 미인대회”라고 평한 바 있다.

자본시장이 정교한 계산으로 움직이는 것이 아니라 투자자들의 대중심리에 따라 크게 흔들린다는 점을 지적한 것이다.

뉴스시스와 코스콤이 이번엔 국내 자본시장에 적용하는 '시장심리지수(Market Sentiment Index·MSI)'는 소셜 빅데이터 분석의 첫 상용화 시도이다.

그간 국내의 시장에서는 숫자 계산만으로는 예측이 불가능한 변수인 '심리'의 방향성을 확인하기 위한 노력이 이어져 왔다.

하지만 시장 심리를 계량화까지는 했더라도 이를 기존의 계량화된 분석과 성공적으로 접목해 시장의 정량·정성적 요인을 아울러 파악하고, 상용화까지 이른 사례는 흔치 않은 편이다.

설령 성공했다고 하더라도 그 산출물은 고정된 포트폴리오 형태로 나오거나, 투자 판단을 위한 하위 요소로 활용되는 경우가 많다.

영국계 투자회사 케이만 아틀란틱(Cayman Atlantic)은 SNS를 활용한 빅데이터 분석을 활용해 일정한 포트폴리오로 구성된 투자 상품인 MTA(Managed Trading Account)를 운영하고 있다.

케이만 아틀란틱은 유럽과 미국에 상장된 기업 정보, 여러 시장 지표와 함께 팔로워 1000명 이상의 트위터 계정을 대상으로 회사가 생산하는 상품 평가를 반영해 투자 대상의 주가 방향성을 가능한다.

이들은 가장 저위험·저수익 지표를 1단계로 두고 7단계로 구분된 수치를 산출해 매수, 매도 판단을 한다.

일본의 카부닷컴 증권은 매매 거래 과정에서 트위터와 SNS인 미시의 사용자 반응을 참고한다.

카부닷컴 증권은 SNS 분석은 문장 단위로 이뤄지며, 주가 변동과 연관성이 높다고 판단되면 투자 판단에 반영하는 방식으로 이뤄진다.

국내에서는 다음소프트가 소셜 매트릭스 기법을 응용한 포트폴리오 투자에 관한 연구를 진행하고 있지만 상용화까지는 이르지 못 했다.

이들은 주가수익비율(PER)과 주가순자산비율(PBR), 주식 차트의 추세 정보와 함께 트위터에서 바라보는 종목에 대한 의견을 토대로 투자를 진행, 머신러닝이 효과적으로 이뤄지고 있는지를 평가하고 있다.

반면 뉴스시스와 코스콤이 공동으로 진행하는 빅데이터 '형태소 분석'은 자본시장의 종목에 대한 시장의 심리 자체를 계량화해 지표를 산출하고, 이를 투자 판단의 참고 요소로 삼는 것에 초점을 맞추고 있다. s.won@newsis.com

SNS 다양한 문장을 최소단위로 '형태소 분석' 수십만개 단어 분석해 놓은 '감성사전'과 비교



심동준 경제부기자

레벨 분석 어떻게?

【서울=뉴스시스】 = 온라인과 소셜네트워크서비스(SNS)에 오르내리는 다양한 문장은 여러 개의 단어로 구성돼 있다.

뉴스시스와 코스콤이 빅데이터 MSI(시장심리지수)를 도출하는 과정에서 사용하는 '형태소 분석'은 문장을 최소 단위까지 분해해 온라인 이면에 숨어 있는 시장 심리를 파악하는 방법론이다.

우선 형태소 분석은 특정 사안에 대해 온라인에서 논의되는 문장을 형태소 수준으로 쪼개는 데서 시작한다.

예컨대 트위터에 “그 회사 물건을 써 보니 참 좋았어요”란 글이 올라왔다면 ‘좋았어요’라는 단어를 근거로 긍정적 의견이라고 판단할 수 있다.

반면 “거기 서비스 이용해 보니 참 별로였다”는 의견은 이용자가 해당 회사를 부정적인 눈으로 바라보고 있다는 점을 시사한다.

특히 복잡한 글보다는 형용사를 통해 감정을 직접적으로 표현하는 경우가 많은 SNS를 분석 대상으로 두면 온라인상의 분위기를 파악하기가 한결 쉬워진다.



기쁘다, 슬프다, 좋다, 싫다 등의 직접적 형용사 표현은 온라인에서의 심리를 조망하기 위한 좋은 지표가 된다.

실제 분석을 위한 데이터는 트위터 같은 SNS는 물론 블로그와 카페, 포털사이트의 게시 글과 뉴스, 증권사 홈페이지 등 온라인 전반에서 수집된다.

이렇게 형태소 분석을 통해 채집된 단어들은 '감성사전'을 통해 계량된다.

감성 사전은 수십만 개의 단어 데이터를 분석해 어떤 사안에 대해 긍정적, 부정적인 평가가 이뤄지고

있는지를 측정하는 일종의 자동화 시스템이다.

이를 통해 감정을 나타내는 형용사 이외의 연관 단어까지 포괄적으로 분석, 시장의 심리까지 가능할 수 있다는 게 코스콤의 설명이다.

이들레텐 분석 대상이 삼성전자라고 하면, 글자 그대로의 '삼성전자' 이외에도 '삼진', 'S전자'와 같은 사실상 같은 대상을 가리키는 단어가 들어간 문맥을 확인한다.

또 '갤럭시 노트', '갤럭시 기어' 등 서브 브랜드까지 포괄해 대상에 미칠 수 있는 심리적 영향을 분석하게 된다.

분석은 코스피200과 코스닥50 종목 등 모두 250곳을 대상으로 이뤄진다.

이들 종목에 대한 “좋다” “나쁘다”는 식의 감정 표현은 물론 “매출이 올랐다”, “부채율이 상승했다” 등과 같은 시장 심리에 영향을 미칠 수 있는 내용에 점수를 부여하는 방식으로 진행된다.

이렇게 수집된 정보는 한 시간 간격으로 형태소 분석과 감성 분석을 거쳐 시장 심리를 반영한 지표로 변환된다.

온라인에서 기업과 주식, 시장을 바라보는 비정형의 심리가, 계량 가능한 정형화된 수치로 치환되는 셈이다.

s.won@newsis.com

데이터 건수-MSI(시장심리지수)스코어-MSI지수 레벨-버즈워드

【서울=뉴스시스】 정필재 기자 = 뉴스시스와 코스콤이 투자자들에게 제공하는 빅데이터 주가 분석 서비스의 핵심 정보는 4가지로 요약된다.

뉴스시스와 코스콤은 온라인 상에 떠도는 수 십만개의 단어 데이터를 분석해 어떤 사안에 대해 긍정적, 부정적인 평가가 이뤄지고 있는지를 측정했다.

이를 통해 250개 상장사에 대한 빅데이터 MSI(Market Sentiment Index·시장심리지수)를 산출했다. 구체적인 데이터는 크게 수집건수와 MSI스코어, MSI레벨, 버즈워드(최다검색어)로 나뉜다.

데이터 수집건수는 해당 종목에 대해 나온 대화나 정보·데이터를 블로그나 카페, 뉴스, SNS, 주식 커뮤니티 등에서 모은 수를 뜻한다.

긍정과 부정의 분류는 감성사전을 통해 구분한다.

감성사전은 코스콤이 자체 제작한 시스템이다. 수집해 쌓인 데이터를 형태소로 분해해 이를 재검, MSI스코어가 분석된다.

'상승'이라는 키워드가 '매출' 뒤에 붙으면 긍정으로 분석되고 '부채' 뒤에 나타나면 부정적인 것을 걸러내는 방식이다.

이 과정에서 버즈워드도 함께 분석된다. 버즈워드란 수집된 데이터를 살펴본 뒤 평소와 잡치지 않던 키워

드가 발견되거나 특정 시점에 많이 오르내리는 단어를 찾아내는 장치다.

평소 1만건 수준의 데이터가 존재했던 한 종목이 어느 날 2만건의 분석건수를 기록했을 경우 사용된다. 버즈워드 분석으로 해당 종목의 이슈를 파악할 수 있다.

MSI스코어는 수집건수를 통해 확보된 단어를 감성 사전에 따라 긍정과 부정으로 분류하고 이를 점수화한 것이다.

수집건수가 갑자기 늘어난 경우 어떤 이슈가 발생했으며 어떤 단어가 가장 많이 모였는지를 분석해 해당 종목에 현재 어떤 일이 일어나고 있는지를 분석한다.

MSI스코어는 타 종목과 비교할 수 없는 만큼 해당 종목의 전달과 비교된다.

MSI레벨은 MSI스코어를 분석해 7단계로 나눈 최종 투자 참고지표다.

MSI스코어는 종목에 따라 상대적이지만 MSI레벨은 해당 종목에 대해 1단계(매우 나쁨)부터 7단계(매우 좋음)를 직관적으로 보여주는 참고지표다.

예컨대 MSI스코어가 최근 1년 동안 50에서 100사이를 움직여 온 종목이 당일 95나 105로 집계했다면 MSI레벨이 7단계가 부여되는 방식이다.



MSI레벨은 투자자들이 쉽게 이해할 수 있도록 이모티콘으로 표출된다.

코스콤 관계자는 “정보 소비자들이 MSI스코어만 제공할 경우 이해하고 받아들이기 복잡한 측면이 있다”며 “현재의 분석결과가 과거와 비교해 어떤지를 쉽게 제공하기 위해서 만들었다”고 설명했다.

rush@newsis.com