

소셜 미디어 감성분석을 통한 주가 등락 예측에 관한 연구

A Study on the Stock Market Prediction Based on Sentiment Analysis of Social Media

김영민 · 정석재 · 이석준

Kim, Youngmin · Jeong, Sukjae · Lee, Sukjun

Abstract

본 연구는 소셜 미디어상의 데이터를 수집하여 주식시장 참여자들의 감성을 분석하고 이를 기반으로 주가 등락을 예측하는 절차를 제안하고자 한다. 이를 위해 국내 대표적인 소셜 미디어로 알려진 네이버, 다음 커뮤니케이션과 주가 정보 제공 사이트인 팍스넷이 제공하고 있는 특정 주식종목에 대해 참여자들이 자유롭게 글을 남길 수 있는 증권 종목 토론실로부터 사용자들의 메시지를 추출한 후 감성분석을 수행하였다. 사용자들의 감성을 분석하기 위해서는 감성사전이 필요하지만 공식적인 한국어 감성사전의 부재에 대한 대안으로 추출된 한국어 메시지들을 구글 응용 프로그래밍 인터페이스를 이용하여 영문으로 변환한 후 영문 긍정·부정 사전과 비교하여 긍정 및 부정 단어의 빈도를 산출하는 Lexicon 기반의 감성분석을 수행하였다. 사용자들의 긍정적/부정적 감성을 주가의 상승/하락으로 정의한 후 최근 1년간 메시지 개수가 풍부한 종목들을 대상으로 주가 등락 예측이 가능한지 실증 분석하였다. 분석 결과, 본 연구에서 제안한 연구 절차가 주가 등락 예측에 유용하다는 결과를 얻을 수 있었으며, 소셜 미디어 데이터의 양이 풍부할수록 주가 예측에 도움이 된다는 것을 확인할 수 있었다.

주제어 : 소셜 미디어, 감성분석, 감성사전, 주가 예측

This study proposes a process for the stock price prediction based on the sentiment analysis using messages from social media. To do this, we selected the well-known social medias, Naver, Daum, and Paxnet, in Korea, and collected the discussion messages of each social media by crawling. The messages were translated Korean into English using Google translation API, and then the sentiments (positive and negative) were extracted from translated messages through the sentiment analysis based on Lexicon. We defined positive and negative word to UP and DOWN of stock price and analyzed movement of stock price during the past one year. Through empirical studies, we realized that the proposed process is useful for stock prediction and a large amount of social media information increase the accuracy of stock forecasting.

Keywords : Social Media, Sentiment Analysis, Sentimental Dictionary, Stock Prediction

김영민 : 연세대학교, 정보산업공학과, 박사과정(제1저자)

정석재 : 광운대학교, 경영학부, 조교수(공동저자)

이석준 : 광운대학교, 경영학부, 조교수(교신저자)

1. 서론

웹 2.0의 출현은 기존의 웹 사용자들에게 정보를 일방적으로 전달하는 구조에서 사용자가 직접 정보를 생성하고 공유함으로써 사용자 스스로 정보를 생산 및 유통할 수 있는 시대를 열어 주었다. 이는 디지털 기술의 발달과 인터넷의 보급으로 일반 사용자들이 텍스트, 오디오, 영상 등 다양한 형태의 정보 및 콘텐츠를 손쉽게 생산하고 배포하는 것을 가속화시켰다. 더불어 스마트 기기의 대중화 및 다양한 온라인 정보 채널의 증가는 정보가 생성, 유통, 저장되는 과정에서 방대하고 새로운 정보의 탄생을 야기하고 있다[1]. 특히 사용자가 소셜 미디어(Social Media) 및 소셜 네트워크 서비스(Social Network Service, 이하 SNS)와 같은 정보 채널을 이용하여 실시간으로 대량의 데이터를 생산하기 때문에 이를 수집하여 사용자의 의견을 분석하고자 하는 오피니언 마이닝(Opinion Mining)이 각광받으면서 사용자의 의견뿐만 아니라 구체적인 사용자의 정서를 분석하는 감성분석(Sentiment Analysis, 이하センチメント 분석) 또한 주목을 받고 있다. 소셜 미디어를 통해 사용자들이 구두로 의사소통을 하는 것보다 온라인상에서 텍스트로 의사소통하며 자신의 정서를 타인에게 전달하는 것에 더 익숙하기 때문에 텍스트에 반영된 사용자의 감정을 분석하는 것이 매우 중요해졌기 때문이다[9]. 따라서 뉴스, 웹 포럼, 채팅룸, 블로그, SNS 등 다양한 소셜 미디어에서 생성되는 사용자의 텍스트 데이터로부터 감성을 분류하고 이를 활용하는センチメント 분석에 대한 연구가 활발히 진행되고 있다[19]. 하지만 소셜 미디어 및 소셜 네트워크 서비스상에서 주고받는 정보는 뉴스와 같이 공식적인 문장을 사용하지 않기 때문에 사용자들의センチメント를 분석하는 것은 쉽지 않다[12]. 따라서 뉴스를 통한 주식시장 예측 및 분석에 관한 연구는 활발하게 진행되어 왔으나 소셜 미디어를 통한 주식시장에 대한 예측 및 분석에 관한 연구는 상대적으로 부족한 실정이다. 특히 해외의 경우,センチメント 분석시 활용되는 감성사전(Sentiment Lexicon)에 관한 연구가 매우 높은 수준으로 진행되어 각 어휘가 가지고 있는 극성값이 미리 구축되어 제공되고 있으나, 국내의 경우에는 공신력 있는 한국어 감성사전이 아직까지 제공되지 않고 있어 각 연구별로 감성사전을 구축하여センチメント 분석을 진행하고 있는 실정이다. 하지만 이러한 연구 방법은 데이터의 양이 늘어나게 되면 각 어휘의 극성값이 변동한다는 한계점을 가지고 있다.

본 연구에서는 주식시장과 관련된 소셜 미디어상의 데이터를 수집하여 개인 투자자들의 감성을 분석하고

이를 기반으로 주가의 등락을 예측하는 분석 절차를 제안한다. 이를 위해 국내 대표적인 포털 사이트인 네이버, 다음 커뮤니케이션과 주가 정보 제공 사이트인 팍스넷에서 제공하고 있는 특정 주식 종목에 대해 참여자들이 자유롭게 글을 남길 수 있는 증권 종목 토론실로부터 사용자들의 메시지를 추출한 후 감성분석을 수행하였다. 사용자들의 감성을 분석하기 위해서는 한국어 감성사전이 필요하지만 공식적인 한국어 감성사전이 존재하지 않는 점을 감안하여 추출된 한국어 메시지를 구글 번역 응용 프로그래밍 인터페이스(Google Translate API)를 이용하여 영문으로 변환한 후 Loughran and Macdonald[20]가 개발한 주식시장 관련 긍정·부정 사전과 비교하여 긍정 및 부정 단어의 빈도를 산출하는センチメント 분석 방법을 수행하였다. Kirchoff, et al.[18]은 기계 번역(Machine Translate) 방법을 이용하여 영어로 된 공공 건강 정보를 에스파냐어로 번역했을 때의 효율성에 대한 연구를 수행하였다. 기계 번역 방법으로 구글 번역기를 사용하였으며, 그 결과 기계 번역의 사용이 전문 번역가에 비해 시간과 비용적인 측면에서 효율적이었으며, 완벽한 번역은 아니지만 번역의 정확성과 신뢰성도 높은 것으로 밝혀졌다.

주가의 등락 예측률(Hit Rate)은 일별 사용자들의 긍정적/부정적 감성과 주가의 상승/하락 간의 일치 정도로 측정하였다. 실증분석 결과, 본 연구에서 제안한 국내 소셜 미디어 감성분석 절차로 주가 예측이 가능하다는 것을 판단할 수 있었으며, 소셜 미디어 데이터의 양이 많을수록 사용자들의 감정이 정확히 파악될 수 있다는 것을 확인할 수 있었다. 이러한 결과는 Antweiler and Frank[10]의 연구에서도 밝혀진 바 있다.

본 연구의 이후 구성은 다음과 같다. 제2장에서는センチメント 분석방법과 선행연구들을 고찰하였으며, 제3장에서는 본 연구가 제안하는 연구모형을 소개한다. 제4장에서는 실증분석을 통해 각 소셜 미디어별 및 주가 등락 예측률을 측정하였다. 마지막으로 제5장에서는 결론 및 한계점을 토의하고자 한다.

2.センチメント 분석 및 관련연구

センチメント 분석은 기본적으로 문서 또는 문장이 긍정, 중립, 부정 중 어떤 견해를 갖고 있는지 판별하는 방법으로, 소셜 미디어에서 발생하는 데이터를 분석할 수 있는 유용한 대안이 되었다. 특히センチメント 분석은 각 문서 최소 단위인 어휘의 감성 극성(Sentiment Polarity)을 기반으로 이루어지므로 감성 극성이 미리 정

의된 감성사전(Sentiment Lexicon)을 구축한 후 새로운 문서에 출현한 어휘의 감성 극성에 따라 문서 전체의 감성 극성을 분류하게 된다. 이러한 센티멘트 분석 방법은 긍정 및 부정 감성 분류가 기록된 학습 데이터 집합에 기계학습을 적용하며, 감성 분류기로 서포트 벡터 머신(SVM, Support Vector Machine) 등을 활용하기도 한다. 또한 사전에 정의된 긍정/부정 Bag of Words(1-grams or Uni-grams)를 이용하여 문서에 포함된 긍정, 부정 단어의 출현 빈도로 긍정과 부정을 판별하는 기법으로 LIWC(Linguistic Inquiry and Word Count)나 POMS(Profile of Mood States) 같은 사전을 이용하는 Lexicon-based 방법이 있으며, 이때 긍정과 부정 코퍼스(Corpus)를 잘 만드는 것이 중요하다. 마지막으로 언어학적(Linguistic) 방법은 문서의 문법적인 구조를 파악하여 극성을 판별하는 방법이다. 이와 같은 방법은 Lexicon-based 방법과 주로 사용되며 문맥(Context)을 파악하여 극성을 판별한다(〈표 1〉 참조).

표 1. 센티멘트 분석 주요 기법

분석기법	설명
Machine Learning Approach	- 사전에 긍정/부정으로 분류된 학습데이터로 텍스트의 긍정/부정 의견을 분류하는 방식 (SVMs(Support Vector Machines)이 주로 쓰임).
Lexicon-based Approach	- 사전에 정의된 긍정/부정 단어를 이용하여 텍스트에 포함된 긍정/부정 단어의 출현 빈도로 긍정과 부정을 판별하는 방식.
Linguistic Approach	- 텍스트의 문법적인 구조를 파악하여 극성을 판별하는 방식.

본 연구에서는 소셜 미디어에서 수집된 한국어 메시지의 영문 번역을 실시하여 센티멘트 분석을 진행하기 때문에 오역 문제를 발생시킬 가능성이 크다. 따라서 본 연구에서는 문맥에 따라 극성을 판단하는 Linguistic 방법이 아닌 긍정과 부정 단어의 빈도를 판별하는 Lexicon-based 방법을 사용하였다. 센티멘트 분석에서는 긍·부정 표현에 해당하는 어휘 정보를 추출하는 것이 매우 중요하며, 기존에 구축된 감성사전을 이용하거나 수작업을 통해 해당 도메인의 긍정/부정을 표현하는 단어들을 기반으로 감성사전을 구축할 수 있다[3]. 따라서 국내에서도 이러한 감성사전 구축의 필요성에 따라 한국어 감성사전에 대한 연구가 현재 활발하게 진행되고 있지만 아직까지 공신력 있는 한국어 감성사전은 개발되지 않았다. 반면 영문인 경우에는 전 세계적으로 기존의 WordNet을 기반으로 SentiWordNet과 WordNetAffect 등을 활용하여 센티멘트

분석이 진행되고 있다. 즉, 각 단어 및 어구에 가중치 또는 극성이 부여되어 누구나 쉽게 활용할 수 있도록 사전이 구축되어 있으며, 다른 나라의 언어까지 확대되어 개발되고 있다.

센티멘트 분석 및 오피니언 마이닝의 초기 연구는 상품평이나 영화평을 분석하는 데 주로 사용되었다 [17][22]. 이후 뉴스 텍스트 마이닝을 통해 주가를 예측하는 연구가 활발히 진행되었으며, 뉴스 텍스트와 주가 간의 상관관계는 비교적 많은 연구결과에서 증명된 바 있다[4][6][15][21]. 그러나 최근에는 뉴스보다 개인의 감정이 직접적으로 반영되는 소셜 미디어 및 SNS가 활성화되면서 주가지수 및 특정 개별 기업의 주가 등락에 투자자들의 감성이 반영되었는지에 대한 연구가 활발히 진행되고 있다[10][13][14][23]. Antweiler and Fank[10]는 다우지수에 포함된 45개의 기업들을 대상으로 금융 관련 온라인 게시판인 야후 금융(Yahoo! Finance)과 Raging Bull 사이트에서 발생한 메시지를 수집하여 수익률, 거래량, 변동성에 영향을 미치는지에 대한 연구를 진행하였으며, 그 결과 게시판에 작성된 정보는 주가의 움직임에 영향을 미치며, 주가 등락을 예측함에 있어 메시지에 나타난 감정에 대한 정보가 통계적으로 유의함을 보였다. 특히 Bollen, et al.[11]은 2008년부터 약 1,000만 개의 트위터(Twitter) 콘텐츠에 표현된 메시지를 Calm, Alert, Sure, Vital, Kind, Happy 등 6가지의 감성으로 분류하였으며, 분류된 감성을 통해 트위터의 전반적인 감성이 어떤 특정한 감성으로 변화는 경우에 따라 익일 다우존수의 등락을 예측한 결과 86.7%의 높은 적중률을 얻었다. Hristidis, et al.[16]은 약 3억4천만 개의 트위터를 분석한 결과, 특정 기업에 대한 언급이 많을수록 주가상승 확률이 증가한다는 연구결과를 보였다. 하지만 국내에서는 소셜 미디어를 활용한 주가 등락 예측에 관한 연구는 미비한 상황이며, 이에 비해 뉴스를 통한 주가 등락 예측 및 영향에 관한 연구는 활발하게 진행되고 있는 실정이다 [5][6][8]. 소셜 미디어를 통한 주가 등락 예측에 관한 연구가 국내에서 부족한 이유는 언급했듯이 신뢰성 높은, 한국어에 대한 감성사전이 존재하지 않기 때문이라 판단된다. 그러나 이러한 연구 중 뉴스를 통해 감성분석을 진행한 대표적인 연구를 살펴보면, 김유신 외[2]는 뉴스 콘텐츠를 감성분석하여 주식시장의 영향도를 평가하고 이를 활용한 지능형 투자사결정 모형을 제안하였다. 그 결과 뉴스 콘텐츠의 감성분석 결과값과 주가지수 등락과는 유의한 관계를 가지고 있음을 보였다. 제안한 지능형 투자사결정 모형은 주가하락시 70%, 주가상승시 78.8%의 분류정확도 보였다. 하지만 실증분석에 수집된 데이터가 3개월, 즉 63

일 동안의 주가지수 등락에 대한 결과라는 점과, 주식 시장에 특화된 도메인 감성사전을 구축하지 못하고 범용사전을 활용한 점이 연구의 한계임을 인정하였다. 또한 유은지 외[7]는 주제지향 감성사전 구축 방안을 제시하였으며, 주식 관련 뉴스에 포함된 어휘를 추출하여 주식 도메인에서의 감성사전을 구축한 후, 주가지수의 등락 예측 모형을 구축하였다. 그 결과 주제지향 감성사전 구축이 주가지수 등락을 예측함에 있어 효과적임을 주장하였지만, 주식 관련 뉴스의 수집 시기에 따라 각 어휘의 극성이 변화할 가능성이 있으며, 실증분석에 사용된 단기간 데이터가 연구의 한계라고 지적하였다.

본 연구에서 제안하는 절차 및 방법과 유사한 연구로는 Das and Chen[13]의 연구가 있다. 그들은 야후(www.yahoo.com) 종목 게시판에 작성된 메시지들을 Web-scrapers라고 불리는 Crawler 프로그램을 활용하여 개별 주식종목(총 9개)의 데이터를 수집하였다. 수집된 데이터는 작성된 메시지의 날짜, 시간, 내용 등으로 이루어져 있으며, 수집된 데이터 전처리를 진행 후センチメント 분석을 진행하였다. 더 나아가 그들은 종목별로センチメント 지수(Sentiment Index)를 개발하고 분류 알고리즘들을 활용하여 매수(공정), 매도(부정), 중립이라는 투자전략 의사결정지원 시스템을 구축하였다. 또한 종목별로 산출된センチメント 지수와 주가의 흐름의 상관분석을 통해 통계적으로 유의함을 주장하였다.

본 연구에서 제안하는 연구절차는 Das and Chen[13]과 유사하지만, 한국어 감성사전의 부재로 인한 해결 방안으로 번역을 활용하였다. 따라서 본 연구에서는 국내의 소셜 미디어 감성분석을 통한 주가 등락 예측 절차를 제안하며, 그 가능성을 살펴보고자 한다.

3. 제안 연구절차

본 연구에서는 소셜 미디어로부터 데이터를 추출하여 주식시장 참여자들의 감성을 분석하고 이를 주가 등락 예측에 활용하는 방법을 제안하고자 한다. <그림 1>은 본 연구에서 제안하는 연구절차를 보여 주고 있으며, 각 단계별로 자세히 설명하면 다음과 같다.

제안 연구절차의 첫 단계에서는 실증분석에 사용될 주식 종목이 선정된다. 종목 선정은 각 소셜 미디어상의 메시지 개수를 파악하여 풍부한 메시지 개수를 보유한 종목들로 이루어졌다. 만일 메시지의 개수가 너무 적거나 종목별로 최초에 저장된 게시물의 날짜가 최근과 비교하여 1년 미만인 종목들은 제외시킨다.

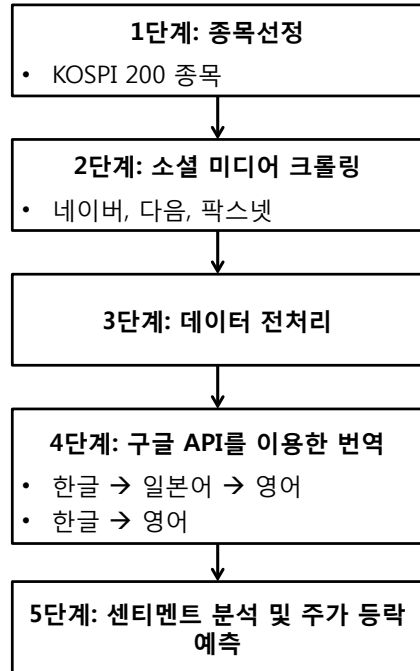


그림 1. 제안 연구절차

두 번째 단계에서는 첫 번째 단계에서 선정된 종목을 대상으로 네이버, 다음 커뮤니케이션, 팍스넷의 종목 토론실을 크롤링(Crawling)한다. 크롤링이란 무수히 많은 웹사이트상의 문서들을 수집하여 검색 대상의 색인으로 포함시키는 기술을 의미한다. 수집된 데이터는 날짜, 제목, 내용, 댓글로 구성된다.

세 번째 단계는 데이터 전처리 과정으로, 소셜 미디어에 작성된 메시지들은 한글 및 다양한 특수 기호(예: !, ?, @, #, \$, %, ^, &, * 등)로 작성되기 때문에 분석에 불필요한 데이터들이 이 과정에서 제거된다. 또한 게시판에 재인용된 신문기사 및 객관적 재무제표표, 인용 텍스트뿐만 아니라 유사한 내용을 반복적으로 사용하여 선동적 메시지를 담고 있는, 소위 “도배글”도 제거된다.

네 번째 단계에서는 전처리 데이터들을 구글 번역 API를 이용하여 영문으로 번역한다. 국내 소셜 미디어에 작성된 메시지들은 비속어, 욕설, 이모티콘 등이 혼재되어 있어 한글 긍정/부정 감성사전을 구축하여センチメント를 구별하는 것이 어렵다. 따라서 이러한 문제점의 대안으로 한글을 영문으로 번역한 후 마지막 단계에서 Loughran and McDonald[20]의 주식시장 영문 긍정/부정 감성사전을 이용하여センチメント 분석을 수행하였다. Loughran and McDonald[20]은 1994년부터 2008년

동안 주식시장 관련 단어들을 수집하고 기존 연구들의 방법들을 고려하여 감성사전을 구축하고 이를 검증하였다.

이 단계에서는 번역을 위해 두 가지 방법을 사용하였다. 첫 번째는 한글을 영문으로 번역하는 것이고, 두 번째는 한글과 어순이 유사한 언어로 번역한 후에 영문으로 재번역하는 것이다. 여기에서는 오역으로 인한 분석의 차이를 살펴보고자 의도적으로 두 가지 방법을 수행하였다.

마지막으로 번역된 메시지들은 긍정과 부정으로 나뉜 감성사전을 이용하여センチメント 분석을 실시한다. 해당 게시물별로 극성값을 계산하여 극성값이 0보다 크면 긍정, 0보다 작은 경우 부정으로 설정한다. 즉,センチメント 분석을 통해 해당 게시물의 긍정단어 개수가 부정단어 개수보다 많으면 긍정, 부정단어 개수가 긍정단어 개수보다 많으면 부정으로 정의한다. 해외 소셜 미디어 기반의センチメント 분석은 극성의 가중치가 결정된 감성사전, 또는 도메인에 따른 감성사전을 구축하여センチメント 분석이 실시되었다. 하지만 본 연구에서는 한글을 영문으로 번역하는 절차에서 오역문제를 보완하고자 단순히 긍정과 부정으로 나뉜 감성사전을 활용하여 어휘의 빈도를 통한センチメント 분석을 진행하였다.

4. 소셜 미디어センチメント 분석을 이용한 주가 등락예측

본 연구에서 수집된 소셜 미디어 데이터는 2013년 4월 1일부터 2014년 3월 31일까지 약 1년간 네이버, 다음, 팍스넷의 종목 토론실에 작성되어 있는 사용자들의 메시지이다. <표 2>은 선정된 10개 종목들의 소셜 미디어별 메시지 개수와 단어의 개수를 보여 주고 있다.

표 2. 네이버, 다음, 팍스넷에서 선정된 10개 종목들의 메시지 및 단어 개수

종목	네이버		다음		팍스넷	
	메시지 개수	단어 개수	메시지 개수	단어 개수	메시지 개수	단어 개수
LG유플러스	4,200	125,523	1,620	68,829	1,000	37,131
LG화학	4,600	130,065	1,590	76,461	1,150	62,730
POSCO	3,280	111,694	1,020	70,135	1,000	40,076
기아차	9,020	257,887	3,810	117,803	1,400	51,372
삼성SDI	3,980	142,714	1,230	66,659	1,100	44,603

삼성엔지니어링	14,400	425,503	1,170	41,696	3,000	79,227
한국전력	4,860	27,241	1,110	64,102	2,800	121,689
한국항공우주	3,620	201,649	1,020	67,216	1,350	69,588
한화케미칼	3,340	93,488	1,140	80,609	1,700	61,720
현대차	4,600	236,736	2,340	116,933	2,250	85,137
평균	5,590	175,250	1,605	77,044	1,675	65,327

미디어별로 종목에 따른 메시지와 단어의 개수가 서로 상이한 것을 볼 수 있으며, 데이터 수집과정에서 각 소셜 미디어별로 정보를 제공하는 특성이 서로 다른 것을 알 수 있었다. 네이버의 경우에는 현재 시점으로부터 정확히 1년 동안의 게시물을 제공하고 있으며, 소셜 미디어 중 네이버가 해당 기간 동안 가장 많은 메시지 개수를 보유하고 있는 것을 알 수 있었다. 이것은 다른 소셜 미디어보다 많은 사용자가 정보를 공유하고 있다는 것을 의미한다. 팍스넷은 기간에 상관없이 한정된 메시지 개수를 열람할 수 있다. 특히 국내 포털 사이트 1위인 네이버는 일정 기간 내에 가장 많은 게시물들이 존재하는 반면에, 다음의 경우 최초 게시물부터 현재까지의 메시지를 제공하고 있어 장기간 분석에는 유용하지만 특정 종목을 제외하고 대부분 종목들의 단기간 내 데이터는 부족한 것으로 파악되었다. 따라서 본 연구에서 사용된 데이터의 제약은 3가지 소셜 미디어의 특성이 고려된 것이다. 종목을 선정 기준은 KOSPI 200 종목들 중 메시지의 개수가 소셜 미디어 3사 모두 1,000건 이상이다.

본 연구에서는 소셜 미디어로부터 데이터를 추출하기 위해 JAVA Programing을 이용하였으며, 수집된 메시지들은 3장에서 언급한 데이터 전처리 과정을 통해 정제되었다. 또한 추출된 메시지들은 구글 번역 API를 사용하여 전술한 바와 같이 두 가지 방법으로 영문번역이 수행되었다. <그림 2>는 소셜 미디어에서 데이터를 수집하고 분석을 위한 데이터 변형 과정을 간략히 보여 주고 있다.

본 연구에서는 국내 주식시장에서 소셜 미디어 데이터를 이용한 주가 등락 예측이 가능한지를 분석하기 위해 사용자들의 메시지 작성 시간을 기준으로 메시지들을 일별(Daily) 및 주별(Weekly)로 정리한 후에センチメント 분석을 수행하였다.

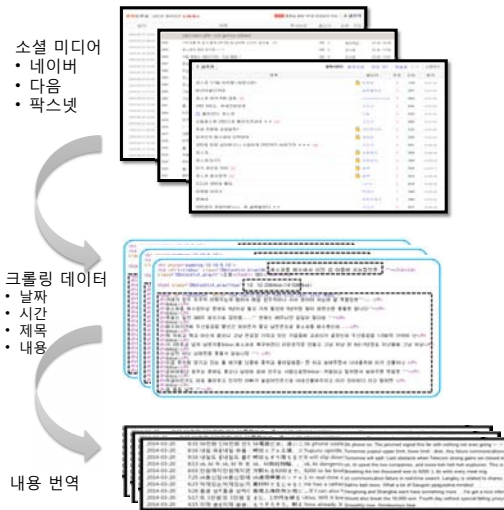


그림 2. 데이터 수집 및 변형 과정 예시

일별 데이터의 경우, 장중 작성된 게시물들도 주가에 영향을 미칠 것으로 판단하여 게시물이 작성된 시간을 기준으로 어제(작일) 15시부터 오늘(금일) 15시까지의 게시물의 긍정/부정과 금일 종가의 상승/하락을 비교하였다. 주식시장은 주중 오전 9시부터 오후 15시까지 운영된다. 단, 개장되지 않는 날(휴일과 국경일)은 t-1 시점의 데이터를 반영하여 분석을 수행하였다. 즉, 월요일 주가 예측을 위해서 금요일 오후 3시부터 토요일과 일요일을 포함하여 월요일 오후 3시까지의 메시지에 대해センチメント 분석을 수행하였다.

주가의 등락은 t 시점의 주가가 t-1보다 크면 '상승', t 시점의 주가가 t-1보다 작으면 '하락'으로 정의하고,センチメント 분석을 통한 극성값이 0보다 크면 상승, 0보다 작으면 하락으로 정의한 후 비교 분석하였다.

<그림 3>은 제안 연구절차의 자세한 수행과정을 설명하기 위해 네이버 금융 종목토론실에서 LG화학 종목과 관련된 메시지를 추출하여 분석하는 모습을 보여 주고 있다. 우선 종목 선정을 위해서는 KOSPI 200 종목 중 게시판의 게시물 개수 확인과정이 선행되어야 하며, 선정된 종목의 게시물을 크롤링한다. 수집된 자료는 정제가 필요하기 때문에 제안 연구절차에서 언급한 데이터 전처리 과정을 수행한다. 정제된 텍스트는 구글 번역 API를 통해 영문으로 변환된 후 Loughran and McDonald[20]의 감성사전과 비교하여 극성을 산출한다. 극성의 긍정/부정을 주가의 상승/하락으로 정의한 후 주가의 실제 등락과 비교 분석하였다.

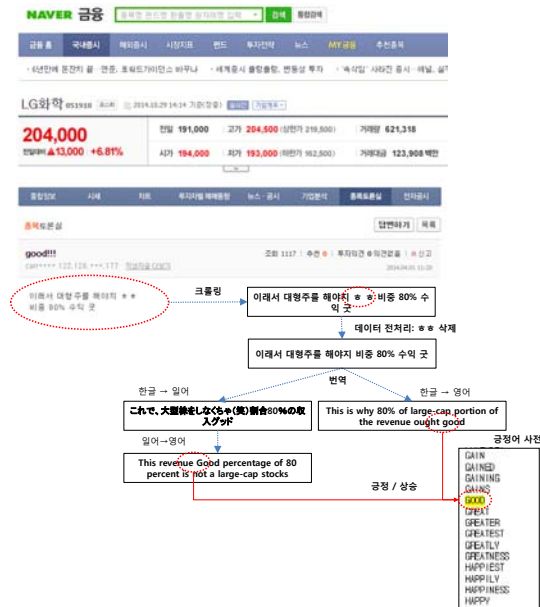


그림 3. 제안 연구절차의 실제 분석 과정

<표 3>은 일별センチメント를 이용한 주가 예측 결과를 보여 주고 있다. 미디어들의 조합이 주가 예측률에 영향을 미치는지 알아보기 위해 미디어들의 메시지 조합에 따른 주가 예측률도 분석하였으며(<표 3> (b) 참고), '한→일→영'과 '한→영'은 번역 순서를 의미한다. 또한 제안한 주가 예측 방법의 우수성을 알아보기 위해 단순 예측법(Naïve Approach)의 예측률 결과와 비교했다. 단순 예측법이란 다음 기간의 값이 현재 기간의 값과 동일할 것으로 추정하는 방법으로, t-1 시점에서 주식이 상승하였다면, t 시점에서 상승이라고 예측하고, t-1 시점에서 주식이 하락하였다면 t 시점에서 하락이라고 예측하는 일반적인 방법이다.

우선 두 가지 번역 방법에 따른 예측률 결과는 종목별로 조금씩 차이가 있지만 평균 예측률로 살펴보면 유사한 정도를 볼 수 있다. 미디어별로 살펴보면, 서로 유사한 예측률을 보이고 있지만 네이버가 다른 미디어들에 비해 약 1% 정도 높은 예측률을 보이고 있다. 3개의 소셜 미디어 모두 평균 예측률이 약 55% 이상이며, 이것은 단순 예측법과 비교했을 때 약 5% 이상의 높은 예측률을 보여 주었다. 또한 개별 미디어들의 예측률보다 미디어들의 조합 개수가 2개(네이버 + 파스넷, 네이버 + 다음, 파스넷 + 다음)에서 3개(네이버 + 다음 + 파스넷)가 됨에 따라 예측률이 조금 더 상승하여 단순 예측법과 비교시 약 7% 이상 차이가 발생하는 것을 알 수 있다. 이 결과는 Antweiler and Fank[10]가 분석한, 메시지 개수가 많으면 많을수록 주가 예측에 도움이 된다는 결과와 부합된다.

표 3. 일별センチメント 분석에 의한 주가 등락 예측률(%)

(a) 소셜 미디어별

종목	네이버		다음		팍스넷		단순 예측법
	한→일 →영	한→영	한→일 →영	한→영	한→일 →영	한→영	
LG유플러스	56.49	53.66	53.72	57.02	47.77	50.2	44.53
LG화학	59.32	59.32	53.47	57.96	55.47	55.06	58.70
POSCO	59.32	61.35	54.29	56.33	46.96	49.80	54.25
기아차	53.66	54.47	56.91	54.07	61.94	61.13	49.80
삼성SDI	55.28	56.89	59.11	52.44	53.44	56.28	46.15
삼성 엔지니어링	59.32	57.30	57.72	55.69	59.58	61.25	52.23
한국전력	53.25	55.68	57.96	51.84	56.68	53.44	54.66
한국항공우주	56.09	57.30	56.33	53.88	53.04	51.82	42.91
한화케미칼	54.87	51.23	52.65	58.37	58.70	53.85	47.77
현대차	54.47	58.11	57.20	56.38	52.23	52.23	57.09
평균	56.21	56.53	55.94	55.4	54.58	54.51	50.81

(b) 소셜 미디어 조합

종목	네이버 + 팍스넷		네이버 + 다음		팍스넷 + 다음		네이버 + 다음 + 팍스넷		단순 예측법
	한→일 →영	한→영	한→일 →영	한→영	한→일 →영	한→영	한→일 →영	한→영	
LG유플러스	53.72	57.44	54.13	60.33	54.55	54.55	52.48	59.09	44.53
LG화학	54.69	56.73	57.96	62.45	57.14	57.14	60.41	62.86	58.70
POSCO	51.43	53.88	57.55	63.27	60.00	61.22	55.92	62.86	54.25
기아차	60.16	58.13	58.94	54.07	56.50	56.50	59.35	56.50	49.80
삼성SDI	57.78	53.33	63.56	58.67	55.11	56.89	59.11	56.89	46.15
삼성 엔지니어링	60.17	59.34	57.09	58.30	63.07	59.75	60.17	57.68	52.23
한국전력	58.37	55.92	54.25	53.44	55.51	56.73	56.33	54.29	54.66
한국항공우주	56.73	52.24	57.09	55.87	58.78	58.37	58.78	58.37	42.91
한화케미칼	56.33	56.73	53.44	55.06	57.55	58.37	56.73	58.78	47.77
현대차	60.08	58.85	53.85	57.09	55.14	55.14	58.44	61.32	57.09
평균	56.95	56.26	56.78	57.85	57.34	57.47	57.77	58.86	50.81

<표 4>는 주별センチメント를 이용한 주가 예측 결과를 보여 주고 있다. <표 3>의 일별 예측에 비해 각각의 미디어 그리고 미디어 조합에 따른 예측률이 약 2% 정도 증가한 것을 볼 수 있으며, ‘네이버 + 다음’의 평균 예측률은 62.26%로, 단순 예측법보다 약 16% 정도 높은 것을 볼 수 있다. ‘네이버 + 다음 + 팍스넷’ 조합의 예측률 또한 약 60% 정도 일별 주가 예측보다 다소 상승한 것을 파악할 수 있었다.

표 4. 주별センチメント 분석에 의한 소셜 미디어 예측률(%)

(a) 소셜 미디어별

종목	네이버		다음		팍스넷		단순 예측법
	한→일 →영	한→영	한→일 →영	한→영	한→일 →영	한→영	
LG유플러스	61.75	64.71	59.82	50.94	56.9	55.98	43.76
LG화학	60.54	61.54	56.77	53.77	55.85	60.77	41.63
POSCO	49.00	56.00	49.00	54.00	38.00	35.00	41.67
기아차	60.54	57.69	60.62	65.31	52.00	62.69	35.43
삼성SDI	56.14	53.06	54.06	59.22	59.14	58.06	35.39
삼성 엔지니어링	50.02	53.06	62.22	69.43	69.35	72.35	54.12
한국전력	63.00	58.00	63.00	54.00	68.00	57.00	62.49
한국항공우주	60.54	55.77	60.62	61.46	57.77	56.92	54.16
한화케미칼	66.31	59.62	60.62	51.85	73.15	62.69	33.35
현대차	47.08	51.92	51.00	61.46	46.23	49.23	58.29
평균	57.49	57.14	57.77	58.14	57.64	57.07	46.03

(b) 소셜 미디어 조합

종목	네이버 + 팍스넷		네이버 + 다음		팍스넷 + 다음		네이버 + 다음 + 팍스넷		단순 예측법
	한→일 →영	한→영	한→일 →영	한→영	한→일 →영	한→영	한→일 →영	한→영	
LG유플러스	64.71	68.67	66.67	62.75	58.82	52.94	64.71	63.75	43.76
LG화학	63.46	63.54	63.46	65.38	55.77	59.62	63.46	64.46	41.63
POSCO	46.00	46.00	56.00	64.00	54.00	48.00	54.00	51.00	41.67
기아차	55.77	59.69	61.54	69.23	61.54	63.46	61.54	68.31	35.43
삼성SDI	51.02	57.10	59.18	63.27	55.1	57.14	55.10	62.22	35.39
삼성 엔지니어링	63.27	59.14	59.18	55.10	69.39	69.39	69.39	60.18	54.12
한국전력	64.00	52.00	58.00	64.00	58.00	54.00	62.00	47.00	62.49
한국항공우주	61.54	57.77	65.38	63.46	59.62	61.54	61.54	58.69	54.16
한화케미칼	65.38	65.46	57.69	55.77	63.46	55.77	57.69	56.77	33.35
현대차	44.23	55.85	50.00	59.62	50.00	59.62	44.23	58.69	58.29
평균	57.94	58.52	59.71	62.26	58.57	58.15	59.37	59.11	46.03

본 연구에서는 두 가지 번역 방법에 따른 주가 예측률 평균이 통계적으로 유의한 차이가 있는지 확인하고자 독립표본 T-검정을 수행하였다. 모든 게시판과 게시판의 조합에서 산출된 예측률 평균을 두 가지 번역방법간 비교로 검정하였다. <표 5>에서 볼 수 있듯이 일별(a) 주별(b) 모두 $p>0.05$ (양측검정)에 해당되기 때문에 두 가지 방법간 예측률 평균의 차이는 통계적으로 유의하다고 볼 수 없다. 즉, 한글에서 영어로 번역하거나 한글에서 일어→영어로 번역하거나, 주가 예측률에는 영향을 미치지 않는다는 것을 의미한다. 이것은 <표 3>에서 전술했듯이 주가 예측률에 두 가지 번역 방법에 따른 큰 차이가 없었던 결과와 일맥상통한다.

표 5. 두 번역 방법간 일별 주별 독립표본 T-검정 결과

(a) 일별

검정 및 집단변수	Mean	Std.	t	df	Sig.
네이버	한→일→영	56.207	-.274	18	.787
	한→영	56.531			
다음	한→일→영	55.936	.538	18	.597
	한→영	55.398			
팍스넷	한→일→영	54.581	.037	18	.971
	한→영	54.506			
네이버 + 팍스넷	한→일→영	56.946	.568	18	.577
	한→영	56.259			
네이버 + 다음	한→일→영	56.786	-.737	18	.471
	한→영	57.855			
팍스넷 + 다음	한→일→영	57.335	-.124	18	.903
	한→영	57.466			
네이버 + 다음 + 팍스넷	한→일→영	57.772	-.934	18	.363
	한→영	58.864			

(b) 주별

검정 및 집단변수	Mean	Std.	t	df	Sig.
네이버	한→일→영	57.492	.145	18	.886
	한→영	57.137			
다음	한→일→영	57.773	-.149	18	.883
	한→영	58.144			
팍스넷	한→일→영	57.639	.124	18	.903
	한→영	57.069			
네이버 + 팍스넷	한→일→영	57.938	-.177	18	.861
	한→영	58.522			
네이버 + 다음	한→일→영	59.710	-1.238	18	.232
	한→영	62.258			
팍스넷 + 다음	한→일→영	58.570	.165	18	.871
	한→영	58.148			
네이버 + 다음 + 팍스넷	한→일→영	59.366	.087	18	.932
	한→영	59.107			

<표 6>과 <표 7>은 제안 연구절차의 주가 예측률과 단순접근법으로 산출된 주가 예측률 간 평균의 차이가 있는지 검증하고자 일별, 주별 주가 예측률을 구분하여 독립표본 T-검정을 수행한 결과를 보여 주고 있다. 검정 결과 대부분의 경우 $p < 0.05$ (양측검정)에 해당되어 주가 예측률 평균의 차이가 통계적으로 유의하다는 결론을 얻을 수 있었다. 이것은 제안 연구절차와 일반적인 주가 예측방법 간의 차별성이 존재한다는 것을 의미하며, 일별과 주별의 주가 예측률 평균의 차이는 메시지의 개수에 따라 주가 예측에 있어 서로 다른 결과가 도출된다는 것으로 판단할 수 있다.

표 6. 제안 연구절차와 단순예측법 간의 일별 독립표본 T-검정 결과

(a) 단순예측법과 ‘한→일→영’ 번역방법

검정 및 집단변수	Mean	Std.	t	df	Sig.
네이버	단순예측법	50.809	2.888	18	.010
	한→일→영	56.207			
다음	단순예측법	50.809	2.768	18	.013
	한→일→영	55.936			
팍스넷	단순예측법	50.809	1.634	18	.119
	한→일→영	54.581			
네이버 + 팍스넷	단순예측법	50.809	3.139	18	.006
	한→일→영	56.946			
네이버 + 다음	단순예측법	50.809	3.030	18	.007
	한→일→영	56.786			
팍스넷 + 다음	단순예측법	50.809	3.419	18	.003
	한→일→영	57.335			
네이버 + 다음 + 팍스넷	단순예측법	50.809	3.708	18	.002
	한→일→영	57.772			

(b) 단순예측법과 ‘한→영’ 번역방법

검정 및 집단변수	Mean	Std.	t	df	Sig.
네이버	단순예측법	50.809	2.943	18	.009
	한→영	56.531			
다음	단순예측법	50.809	2.474	18	.024
	한→영	55.398			
팍스넷	단순예측법	50.809	1.729	18	.101
	한→영	54.506			
네이버 + 팍스넷	단순예측법	50.809	2.907	18	.009
	한→영	56.259			
네이버 + 다음	단순예측법	50.809	3.486	18	.003
	한→영	57.855			
팍스넷 + 다음	단순예측법	50.809	3.640	18	.002
	한→영	57.466			
네이버 + 다음 + 팍스넷	단순예측법	50.809	4.180	18	.001
	한→영	58.864			

표 7. 제안 연구절차와 단순예측법 간의 주별 독립표본 T-검정 결과

(a) 단순예측법과 ‘한→일→영’ 번역방법

검정 및 집단변수	Mean	Std.	t	df	Sig.
네이버	단순예측법	46.029	2.933	18	.009
	한→일→영	57.492			
다음	단순예측법	46.029	3.222	18	.005
	한→일→영	57.773			
팍스넷	단순예측법	46.029	2.450	18	.025
	한→일→영	57.639			
네이버 + 팍스넷	단순예측법	46.029	2.846	18	.011
	한→일→영	57.938			
네이버 + 다음	단순예측법	46.029	3.753	18	.001
	한→일→영	59.710			
팍스넷 + 다음	단순예측법	46.029	3.368	18	.003
	한→일→영	58.570			
네이버 + 다음 + 팍스넷	단순예측법	46.029	3.353	18	.004
	한→일→영	59.366			

(b) 단순예측법과 ‘한→영’ 번역방법

검정 및 집단변수		Mean	Std.	t	df	Sig.
네이버	단순예측법	46.029	10.449	3.135	18	.006
	한→영	57.137	4.049			
다음	단순예측법	46.029	10.449	3.152	18	.006
	한→영	58.144	6.208			
팍스넷	단순예측법	46.029	10.449	2.439	18	.025
	한→영	57.069	9.781			
네이버 + 팍스넷	단순예측법	46.029	10.449	3.202	18	.005
	한→영	58.522	6.558			
네이버 + 다음	단순예측법	46.029	10.449	4.540	18	.000
	한→영	62.258	4.313			
팍스넷 + 다음	단순예측법	46.029	10.449	3.180	18	.005
	한→영	58.148	6.004			
네이버 + 다음 + 팍스넷	단순예측법	46.029	10.449	3.381	18	.003
	한→영	59.107	6.358			

5. 결론 및 한계점

본 연구는 국내 주식 관련 소셜 미디어 네이버, 다음, 팍스넷을 선정한 후 개별 종목에 대해 사용자들의 메시지로부터センチメント를 도출하여 해당 종목들의 주가 등락 예측이 가능한지 알아보았다. 분석 결과 주가 예측률은 55% 이상으로 주가가격의 상승과 하락 각각의 확률 50%보다 높은 것을 알 수 있었으며, 일별보다는 주별 주가 예측률이 약 2% 정도 높았다. 또한 소셜 미디어 3사의 메시지를 결합한センチメント의 주가 등락 예측률은 약 60% 이상인 것으로 파악되었으며, 이것은 메시지의 개수가 풍부할수록 사용자들의 집단지성이 좀더 뚜렷하게 분석된다는 것을 시사하는 바이다[10]. 더불어 일반적인 주가 예측 방법인 단순예측법에 비해 약 5~7% 더 높은 성과를 보여 주고 있어 본 연구에서 제안한センチメント 분석을 통한 주가 등락 예측 방법이 유용하다는 결론을 얻었다. 단, 본 연구에서 진행한 실험 결과를 확대 해석하기에는 무리가 있다. 무엇보다도 한글에 대한 감성사전 대신 영문번역을 통해センチメント 분석을 실시하였다는 점이다. 그러나 번역을 통해서도 55% 이상의 예측 성과가 나타나므로써 한국어センチメント 분석을 통한 주가 등락 예측 성과는 더 나은 결과를 도출할 것으로 사료된다. 또한 국내 소셜 미디어의 급격한 활용 증대 및 데이터 기반 연구의 중요성이 증대되고 있는 시점에서 무심하게 버려질 수 있는 소셜 미디어 데이터를 활용하여 국내 주식시장에 활용될 수 있다는 점에서 본 연구의 의의를 찾을 수 있다.

본 연구의 한계점으로는 메시지 개수의 제약으로 인해 더 많은 주식 종목을 표본으로 사용하지 못하고 10개의 표본만 사용했다는 점이다. 향후 연구에서

는 더 많은 종목을 기반으로 성과를 도출하여 통계적으로センチメント 분석의 유의성을 검증할 것이다. 또한 본 연구에서는 단순히 예측성과만을 살펴보았으나, 더 나아가 투자전략 개발을 위한センチメント 분석 모델을 개발하고자 한다. 즉, 긍정 및 부정을 판단하는 기준에 임계치를 도입하여, 주식 매수, 매도, 보유의 세 가지 형태의 의사결정을 판단하여 투자 전략 개발이 가능할 것이다. 더불어 t 시점 예측을 위해 사용될 메시지 기간(예를 들어 t-1, t-2, t-3)을 고려하여 정보전이효과(Information Spillover)를 분석하고자 한다. 마지막으로 국내 주식시장 관련 감성사전을 구축하여 제안 연구 절차와의 비교 분석을 통해 더 나은 주가 예측 프로세스를 개발하고자 한다.

참고문헌

- [1] 김승우, 김남규, 2014. “오피니언 분류의 감성사전 활용효과에 대한 연구”, *지능정보연구*, 제20권, 제1호, pp.133-148.
- [2] 김유신, 김남규, 정승렬, 2012. “뉴스와 주가: 빅데이터 감성분석을 통한 지능형 투자의사결정모형”, *지능정보연구*, 제18권, 제2호, pp.143-156.
- [3] 박경미, 박호건, 김형근, 고희동, 2011. “SNS에서 오피니언마이닝 연구”, *정보처리학회지*, 제18권, 제6호, pp.68-78.
- [4] 백우진, 경명현, 민경수, 오혜란, 임차미, 신문선, 2007. “주식 가격 변동 예측을 위한 다단계 뉴스 분류시스템”, *정보관리학회지*, 제24권, 제2호, pp.123-141.
- [5] 송치영, 2002. “뉴스가 금융시장에 미치는 영향에 관한 연구”, *국제경제연구*, 제8권, 제3호, pp.1-34.
- [6] 안성원, 조성배, 2010. “뉴스 텍스트 마이닝과 시계열 분석을 이용한 주가예측”, *2010 한국컴퓨터종합학술대회 논문집*, 제37권, 제1호(C), pp.364-369.
- [7] 유은지, 김유신, 김남규, 정승렬, 2012. “주가지수 상승 예측을 위한 주제지향 감성사전 구축 방안”, *한국지능정보시스템학회 2012년 추계학술대회*, pp.42-49.
- [8] 이근영, 2006. “북한 핵관련 뉴스가 국내주식 및 외환시장에 미치는 영향”, *동북아경제연구*, 제18권, 제1호, pp.61-90.
- [9] 이신영, 고일주, 2013. “소셜 미디어에서 사용되는 한국어 정서 단어의 정서가, 활성화 차원 측정”,

- 한국감성과학회지, 제16권, 제2호, pp.167-176.
- [10] Antweiler, W. and Frank, M.Z., 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards", *The Journal of Finance*, Vol.59, No.3, pp.1259-1294.
- [11] Bollen, J., Mao, H. and Zeng, X.J., 2011. "Twitter mood predicts the stock market", *Journal of Computational Science*, Vol.2, No.1, pp.1-8.
- [12] Capelle-Blancard, G. and Monjon, S., 2010. "Socially responsible investing: It takes more than words", CEPII Working Paper, No.2010-15.
- [13] Das, S.R. and Chen, M.Y., 2007. "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web", *Management Science*, Vol.53, No.9, pp.1375-1388.
- [14] Dubinko, M., Kumar, R., Magnani, J., Novak, J., Raghavan, P. and Tomkins, A., 2007. "Visualizing tags over time", *ACM Transaction*, Vol.1, No.2, pp.139-202.
- [15] Fu, T., Lee, K.K., Sze, D., Chung, F.L. and Ng, C.M., 2008. "Discovering the correlation between stock time series and financial news", *IEEE/WIC/ACM International Conference on the Web Intelligence and Intelligent Agent Technology*, pp.880-883.
- [16] Hristidis, V., Ruiz, E.J., Castillo, C., Gionis, A. and Jaimes, A., 2012. "Correlating financial time series with micro-blogging activity. *Web Search and Data Mining*.
- [17] Hu, M. and Liu, B., 2004. "Mining and summarizing customer reviews", *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp.168-177.
- [18] Kirchhoff, K., Turner, A.M., Axelrod, A. and Saavedra, F., 2011. "Application of statistical machine translation to public health information: a feasibility study", *Research and applications*, Vol.18, No.4, pp.473-478.
- [19] Li, N. and Wu, D.D., 2010. "Using text mining and sentiment analysis for online forums hotspot detection and forecast", *Decision Support System*, Vol.48, No.2, pp.354-368.
- [20] Loughran, T. and McDonald, B., 2011. "When is a Liability not a Liability", *Journal of Finance*, Vol.66, No.1, pp.35-65.
- [21] Mittermayer, M.A. and Knolmayer, G., 2007. "Text mining systems for predicting market response to NEWS", *Proceedings of the IADIS European Conference on Data Mining*, pp.164-169.
- [22] Pang, B. and Lee, L., 2008. "Opinion Mining and Sentiment Analysis", *Foundations and Trends in Information Retrieval*, Vol.2, No.1-2, pp.1-135.
- [23] Zhou, Y., Guan, X., Zhang, Z. and Zhang, B., 2008. "Predicting the tendency of topic discussion on the on-line social networks using a dynamic probability model", *Proceedings of the hypertext 2008 workshop on Collaboration and collective intelligence*, pp.7-11.

저자소개



김영민
(Kim, Youngmin)

국민대학교 비즈니스IT학과에서 학사를 마쳤으며, 현재 연세대학교 정보산업공학과에서 박사과정에 있다. 주요 연구 관심분야는 빅데이터 분석을 활용한 금융시장 분석 및センチメント 분석, 인공지능을 활용한 트레이딩 시스템 개발 등이다.

E-mail: wealth38@gmail.com

Tel: +82-2-415-4141



정석재
(Jeong, Sukjae)

연세대학교 정보산업공학과에서 박사학위를 취득하고 현재 광운대학교 경영학부 교수로 재직중이다. 주요 관심 분야는 SCM, 물류 최적화 모델링, 시뮬레이션 등이다.

E-mail: sjjeong@kw.ac.kr.com

Tel: +82-2-940-5294



이석준
(Lee, Sukjun)

연세대학교 정보산업공학과에서 박사학위를 취득하고 현재 광운대학교 경영학부 교수로 재직중이다. 주요 관심 분야는 빅데이터 분석을 활용한 금융시장 분석, 텍스트 마이닝, 데이터 마이닝 등이다.

E-mail: sjlee@kw.ac.kr

Tel: +82-2-940-8103

◇ 이 논문은 2014년 10월 16일에 접수하여 2014년 11월 3일 1차 수정을 거쳐 2014년 11월 5일에 게재확정되었습니다.