Journal of Korean Society of Industrial and Systems Engineering
https://doi.org/10.11627/jksie.2023.46.1.032

Vol. 46, No. 1 : 32-41, March 2023
ISSN : 2005-0461 / EISSN : 2287-7975

Developing the Automated Sentiment Learning Algorithm to Build the Korean Sentiment Lexicon for Finance

Su-Ji Cho · Ki-Kwang Lee · Cheol-Won Yang†

School of Business Administration, Dankook University

재무분야 감성사전 구축을 위한 자동화된 감성학습 알고리즘 개발

조수지 · 이기광 · 양철원†

단국대학교 경영학부

Recently, many studies are being conducted to extract emotion from text and verify its information power in the field of finance, along with the recent development of big data analysis technology. A number of prior studies use pre-defined sentiment dictionaries or machine learning methods to extract sentiment from the financial documents. However, both methods have the disadvantage of being labor-intensive and subjective because it requires a manual sentiment learning process. In this study, we developed a financial sentiment dictionary that automatically extracts sentiment from the body text of analyst reports by using modified Bayes rule and verified the performance of the model through a binary classification model which predicts actual stock price movements. As a result of the prediction, it was found that the proposed financial dictionary from this research has about 4% better predictive power for actual stock price movements than the representative Loughran and McDonald's (2011) financial dictionary. The sentiment extraction method proposed in this study enables efficient and objective judgment because it automatically learns the sentiment of words using both the change in target price and the cumulative abnormal returns. In addition, the dictionary can be easily updated by re-calculating conditional probabilities. The results of this study are expected to be readily expandable and applicable not only to analyst reports, but also to financial field texts such as performance reports, IR reports, press articles, and social media.

Keywords : Text Mining; Analyst Report; Classification; Modified Bayes Rule; Sentiment Extraction

1. 서 론

빅데이터 기술의 발달과 함께 기존에 사용되지 않던 비정형자료(unstructured data)를 통해 의미 있는 정보를 추출하고자 하는 연구들이 활발해지고 있다. 가장 대표적인 비정형자료가 인간의 언어로 표현된 텍스트(text)이다. 인간은 다른 동물과 달리 언어를 통해 자신의 정보를 전달할 수 있다. 인간의 텍스트를 통해 의미 있는 정보를 추출하는 기법을 텍스트 마이닝(text mining)이라 하며, 특히 추출하고자 하는 정보가 긍정이나 부정의 감성일 때 이를 감성분석(sentiment analysis)이라 한다[5, 16].

금융(finance)은 감성분석을 적용하기 가장 좋은 분야이다. 금융분야 연구자들은 자산가격의 움직임에 주된 관심을 가지고 있으며, 특히 정보와 가격의 관계에 대한 심도 있는 연구들이 활발하게 이루어지고 있다. 텍스트에서 추출한

Received 15 December 2022; Finally Revised 10 January 2023; Accepted 10 January 2023
† Corresponding Author : yang@dankook.ac.kr

부정과 긍정의 정보가 주식가격이나 이익에 대한 예측력을 가지고, 또는 기업의 재무적 곤경이나 부도를 예측하는 데 도움을 줄 수 있음이 발견되었다[14, 27].

하지만, 텍스트에서 감성을 추출하는 방법에 있어서는 부족한 점이 많다. 지금까지 텍스트에서 감성을 추출하기 위해서는 감성사전을 사용하는 방법과 연구자의 판단에 의해 수동으로 감성이 분류된 훈련자료(manually-coded train data set)를 기초로 하는 머신러닝(machine-learning) 방법이 주로 사용되었다. 두 방법 모두 인간의 자의적인 판단에 의지하고 있다는 단점이 있다.

첫 번째 방법은 감성사전을 사용하여 텍스트에 나온 긍정어와 부정어의 수를 세는 방식이다. 이 방법은 얼마나 정확한 감성사전을 구축해서 가지고 있느냐에 따라 성과가 좌우된다. 초기 재무분야 연구에서는 하버드 대학에서 제공한 일반 감성사전(General Inquirer)을 많이 사용하였다. 하지만, Loughran and McDonald[22]은 이러한 범용 감성사전이 재무분야의 감성을 측정하는데 적합하지 않음을 보여주었으며, 자신의 판단에 의해 구축한 새로운 감성사전(LM사전)을 제시하였다. 이후 수 많은 연구들이 LM사전을 사용하여 재무분야 텍스트의 감성을 측정하고 있다. 두 번째 방법인 머신러닝을 이용한 방법도 주로 나이브 베이지안 방법(Naïve Bayesian method)을 사용한다. 여기서 중요한 것은 정확한 감성 훈련자료를 구축하는 것인데, 연구자의 판단에 의해 문장의 감성이 표시된다.

이와 같이 단어의 감성 학습 과정에 인간의 판단이 들어가면 몇 가지 문제점이 있다. 첫째, 분류를 위해 많은 시간과 노력이 들어간다. 둘째, 인간의 판단에 오류가 포함될 수 있다. 또한, 판단하는 사람에 따라 감성 분류 결과가 달라질 수도 있다.

본 연구는 자동감성학습 알고리즘을 사용하여 LM사전을 뛰어넘는 재무감성사전을 만드는 것을 목표로 한다. 우리가 제안하는 자동 알고리즘은 사전 방법과 머신러닝 방법의 장점을 결합한 하이브리드 방식이다. 각 단어에 감성을 부여하는 사전방식을 기본으로 하지만, 감성을 부여하는 과정을 베이지안 머신러닝 방법을 사용하여 자동으로 구현하였다. 우리의 방법은 기존의 방법과 같이 인간의 판단과 수작업이 필요하지 않는 장점을 지닌다. 각 단어의 긍정과 부정 감성 확률이 자동적으로 부여되기 때문이다.

좋은 감성사전을 만들기 위해서는 좋은 훈련자료가 필요하다. 이를 위해 본 연구에서는 애널리스트 보고서 자료를 사용하였다. 애널리스트는 시장과 기업에 대한 전문가이다[7]. 또한 그들이 제공하는 보고서에는 텍스트 뿐만 아니라 추천의견, 목표주가, 이익예측치 등의 수치정보도 함께 제시되어 있다. 이는 머신러닝 알고리즘을 통해 자동으로 재무감성사전을 만들기에 적합한 자료이다. 텍스트를 작성한 애널리스트 자신이 목표주가 등의 수치정보도 같이 제시하기 때문에 실제 사용한 단어들의 감성을 목표주가의 변화 등 수치자료의 변화를 통해 추출해 낼 수 있는 것이다.

본 연구에서 제안한 알고리즘을 다른 금융 분야에도 적용할 수 있는 실용적 가치도 지니고 있다. 애널리스트 보고서를 기반으로 한 재무감성사전은 기본적으로 주식투자 등에 사용될 수 있다. 하지만, 더 나아가서 신용등급, 보험 등 다른 분야에도 적용 가능하다.

본 연구는 다음과 같이 구성되어 있다. 2장에서는 선행연구 결과를 제시하며, 3장에서는 자료와 방법론을 설명한다. 4장에서는 자동 알고리즘을 통해 구축한 재무감성사전의 성과를 평가한다. 5장은 결론을 제시한다.

2. 재무 분야 감성 분석 연구

최근 빅데이터 기술의 활성화와 함께 재무 분야에서 전통적 정보로서의 정량적 정보뿐만 아니라 정성적 정보 즉, 텍스트를 분석하는 연구가 수행되고 있다. 재무 분야에서 텍스트 분석은 주로 텍스트의 긍정 또는 부정 감성을 추출하여 사용하는 감성분석이 주를 이룬다. 감성분석을 위한 기초적인 방식은 단어와 각 단어의 감성이 사전에 정의된 감성사전(pre-defined dictionary)을 활용하는 것이다. 사전 정의된 감성사전은 크게 범용 감성사전과 재무 특화 감성사전으로 구분된다. 초창기 재무 분야 감성분석 연구에는 범용 감성사전 소프트웨어로서 워드넷(WordNet)사의 SentiWordNet이나 OpinionFinder, 구글(Google)의 Google-Profile of Mood State 등을 활용하거나, 또는 사전에 정의된 감성사전으로서 Harvard IV에서 제공하는 긍정어 및 부정어 리스트(General Inquirer; GI)를 활용하였다[2, 25, 26, 28]. 대표적으로 Tetlock[25]은 Wall Street Journal에 게시된 칼럼의 감성을 GI를 활용하여 추출하였으며, 부정적 어조가 익일 Dow Jones 지수의 낮은 수익률과 주식 거래량에 영향을 미친다고 결론 내렸다. 그러나 Loughran and McDonald[22]은 이 같은 범용 감성사전으로 금융 및 재무분야의 텍스트 감성을 추출하는 것은 적합하지 않음을 보였다. 그들은 1994년부터 2008년까지 미국 기업의 사업보고서에 기반한 금융 분야에 특화된 영문 단어 감성사전을 개발하였으며, 이후 대다수의 재무 분야 감성분석 연구에서 해당 감성 사전(LM 감성사전)을 활용하고 있다[6, 8, 12, 22]. Cho et al.[8]은 LM 감성사전을 바탕으로 기업 재무분석을 위한 감성 단어를 연구자 판단 하에 추가하여 한국어 감성사전 KOSELF(Korean Sentiment Lexicon for Finance)을 구축하였다. 그러나 이 같이 사전에 정의된 감성사전은 구축 과정에 상당한 시간과 비용이 요구되며, 시간의 흐름에 따른 추가 및 수정이 어렵다는 단점이 있다. 또한 감성사

전은 사전의 높은 정확도가 요구됨에도 불구하고 사전 개발자의 주관적인 판단에 의지한다는 한계가 있다.

두 번째 가능한 감성분석 방법은 머신러닝 방법론을 활용하는 것이다. 머신러닝 방법은 문서를 감성에 따라 분류하기 위하여 통계적 추론을 활용한다. Schumaker and Chen[23]은 서포트벡터회귀(Support Vector Regression)를 활용하여 온라인 뉴스와 주가 간 관계를 밝혔으며, Bollen et al.[2]은 인공신경망 모델을 통해 트위터의 감성이 다우존스지수의 변화를 예측할 수 있다고 주장하였다. 이외에도 Deng et al.[11]은 앙상블 기법으로 서포트 벡터 머신의 다양한 커널을 결합한 다중커널학습을 통해 주가를 예측하는 시스템을 개발하였다. 통계적 추론을 사용한 많은 연구는 주로 나이브 베이즈(Naïve Bayes) 알고리즘을 활용하여 각 문서의 감성을 학습시킨 뒤 분류 및 예측에 활용하였다[4, 17, 19, 20]. 대표적으로 Li[19]과 Li et al.[20]은 나이브 베이즈 알고리즘을 통해 10-K 보고서의 Management Discussion and Analysis (MD&A) 섹션에 기재된 기업의 미래 예견 관련 서술(forward-looking statements)에 대한 감성을 추출하였다. 이들은 추출한 감성이 기업의 미래 수익에 긍정적 영향을 준다는 사실을 밝혔다. 그러나 이러한 머신러닝 방법 또한 필수적으로 학습 과정에서 연구자의 판단을 필요로 한다는 단점이 있다. 예를 들어 Li[19]의 연구에서는 전체 데이터 중 학습에 사용한 데이터의 비율은 0.23%에 그쳤는데, 이는 각 문장의 감성을 수동으로 라벨링함으로써 학습하기 때문에 학습 데이터를 늘리는 것에 대한 비용 부담이 크기 때문으로 보인다. 학습 과정의 비용 문제 이외에도 머신러닝 학습 과정의 신뢰도 또한 한계점이 될 수 있다. 머신러닝 분류기의 성능은 전적으로 학습 데이터에 의존하기 때문에 이를 정확하게 학습 즉, 라벨링(labelling)하는 것은 매우 중요한 단계이다. 따라서 데이터 라벨링을 수행하는 연구자는 주로 해당 분야의 사전 지식을 보유하고 있는 다수의 전문가를 섭외하여 라벨링을 수행하는데[1], 이 경우 전문 작업자 간 라벨링 결과의 일치도가 확보되어야 한다[2][9, 17].

따라서 감성분석을 위한 학습 데이터 구축 과정에서 주관적 판단으로 인한 한계점이 분명히 존재한다. 따라서 본 연구에서는 이 같은 수동 라벨링을 통한 단어의 감성 학습 과정을 자동화할 수 있는 자동 감성 학습 알고리즘을 개발하고 이에 대한 검증을 수행하였다.

1) 이러한 과정이 매우 노동 집약적(labor-intensive)이기 때문에 최근에는 온라인 크라우드소싱 플랫폼인 Amazon's Mechanical Turk(MTurk)을 통해 데이터 라벨링 과정을 온라인에서 다수의 작업자가 수행하는 경우 또한 존재한다. 그러나, 연구에 따르면 이 경우 작업자 간 라벨링 신뢰도(reliability)가 최소 53%에서 최대 82% 범위로 나타났다[9].

2) Kim and Joh[17]의 연구에서는 연구자 본인과 다수의 석사급 연구원이 재무 분야 문장의 긍부정 감성을 학습(라벨링)한 결과 약 89%의 일치도를 보였다.

3. 자료 및 방법론

3.1 자료 수집

표본기업으로는 국내 주식시장 상장기업 중 2016년 기준 시가총액 상위 100개 기업을 선정하였다. 표본기간으로 2016년부터 2018년을 설정하고, 해당 기간동안 표본기업을 대상으로 발행된 애널리스트 보고서와 시장 자료를 수집하였다. 애널리스트 보고서는 Python 프로그래밍을 통한 웹 크롤링을 활용해 자동 수집하였으며, 한경 컨센서스[3]에 등록된 모든 애널리스트 보고서의 발행일자, 제목, 작성자(애널리스트), 증권사와 같은 일반적 정보뿐만 아니라, 제시된 추천의견, 목표주가의 정량적 정보부터 애널리스트 보고서의 본문 텍스트의 정성적 정보까지 모두 수집하였다. 시장 자료로서는 선행연구를 바탕으로 대상기업의 시가총액, 거래량, 베타, Book-to-Market ratio, 레버리지, 전년동기대비 성장률(YoY Growth), 일별 수익률 자료를 FnGuide를 통해 수집하였다[18]. 또한 시장초과수익률 산출을 위한 일별 시장 수익률(market return) 자료 또한 FnGuide를 통해 수집하였다.

본 연구의 목적으로서 애널리스트 보고서 감성을 학습하고 보고서 발행 이후 개별 종목의 실제 CAR(cumulative abnormal return) 움직임을 예측하기 위하여 수집한 애널리스트 보고서를 학습 데이터 셋(Train Data Set)과 검증 데이터 셋(Test Data Set)으로 분할하였다. 일반적인 기계학습 분류 문제에서는 학습과 검증 데이터 셋을 9:1 또는 8:2의 비율로 무작위(random) 분할하나, 이와 달리 본 연구에서의 주식 가격 분류(예측)는 시계열 데이터의 특성을 가지고 있다. 시계열 데이터의 특성 상 학습 및 검증 데이터를 무작위 추출하는 것은 시뮬레이션으로서의 의미가 적다. 따라서 예측모형의 시뮬레이션을 통해 보다 실용적인 활용능력을 검증하기 위하여 특정 시점을 기준으로 데이터 셋을 분할하되, 일반적인 분할 비율을 유지할 수 있도록 하였다. 이에 따라 2018년 3분기를 기준으로 이전 10개 분기(2016년 1분기~2018년 2분기)에 해당하는 85%의 데이터는 학습 데이터 셋으로 사용하고 이후 2개 분기(2018년 3분기~2018년 4분기)에 해당하는 15%의 데이터는 검증 데이터 셋으로 사용하였다.

3.2 감성사전 구축 알고리즘

본 연구에서는 감성사전 구축을 위해 베이즈 확률에

3) 한경 컨센서스 웹페이지: https://markets.hankyung.com/consensus.

기반한 자동학습 알고리즘을 사용하였다. 베이즈 확률은 기타 머신러닝 방법론에 비해 가벼운 연산으로도 우수한 학습성과를 나타내는 것으로 알려져 있으며, 확률에 기반하여 감성을 학습하기 때문에 학습 데이터가 추가·변경되는 경우에도 단어가 가지는 감성의 변화, 즉 확률의 갱신이 쉽고 빠르게 가능하다는 장점이 있다.

<Table 1>은 본 연구에서 감성사전 자동 학습에 사용한 학습 범주를 나타낸다. 개별 애널리스트 보고서의 특성에 따라 네 가지 범주로 구분하였으며, 애널리스트 보고서의 특성으로는 목표주가 변화율(dTPRC)의 방향과 발행시점 이후 2일 누적초과수익률($CAR_{[0,+1]}$)의 방향에 대한 조합을 사용하였다. 따라서 Class 1과 Class 2는 부정적 감성을, Class 3와 Class 4는 긍정적 감성을 학습하고 있다. 특히, 애널리스트 보고서가 제시한 목표주가 변화율의 방향과 실제 발행시점 이후 2일 누적초과수익률의 방향이 일치하는 경우(Class 1 또는 Class 4), 반대의 경우보다 긍·부정 감성의 정도가 더욱 큰 것으로 해석할 수 있다. 학습 데이터 셋 중 위 네 개 조합에 부합하지 않는 데이터는 제외하고 총 8,938개의 학습 데이터 셋에 대하여 보고서 본문의 출현 단어에 대한 감성을 학습하였다.

<Table 1> Four Classes used to Train Sentiment

Criteria	Negative Sentiment		Positive Sentiment	
	Class 1 (More Negative)	Class 2 (Less Negative)	Class 3 (Less Positive)	Class 4 (More Positive)
Direction of dTPRC	Negative	Zero (Reiteration)	Zero (Reiteration)	Positive
Direction of $CAR_{[0,+1]}$	Negative	Negative	Positive	Positive
# of Sample Reports(Train Data Set)	804 (9.0%)	3,371 (37.7%)	3,550 (39.7%)	1,213 (13.6%)

애널리스트 보고서 본문 단어에 대한 감성은 다음과 같이 학습하였다. 먼저 애널리스트 보고서 본문 전처리(preprocessing) 과정으로서 불용어(영문, 숫자, 특수문자)를 제거하고 경제·금융분야 자연어 처리를 위한 한국어 형태소 분석기인 eKoNLPy(Korean NLP Python Library for Economic Analysis)를 사용하여 형태소 분석 후 명사만을 추출하였다. 추출한 명사(단어)는 문맥에 따른 감성 변화를 반영하기 위하여 n-gram으로 변환하였으며, 이를 토큰(token), 즉 감성 학습을 위한 기본 입력단위로 정의하였다. N-gram은 n개의 연속된 단어 나열을 의미하며, '최저임금 상승(2-gram)', '수익성 악화 지속(3-gram)' 등 개별 단어(1-gram)로 반영하기 어려운 감성 변화를 반영하기 위하여 사용되는 방법이다. 본 연구에서는 문맥의 의미를 반영할 수 있도록 최대 3-gram 토큰을 사용하였다. 보고서 본문을 토큰화(tokenizing)한 이후에는 각 n-gram 토큰이 4개의 감성 범주 각각에서 출현한 횟수를 세어 특정 n-gram 토큰 k($token_k$)가 특정 감성 범주 내(inclass)에서 나타날 확률($p_{k,indass}$)과, 특정 감성 범주 외(outclass)에서 나타날 확률($p_{k,outdass}$)을 계산한다. 계산 과정에서 아래와 같은 조건부확률을 사용하였다. 이후 두 조건부확률의 비율을 통해 토큰 k가 특정 범주 외 출현 대비 특정 범주 내에서 출현한 비율을 산출한다. 마지막으로 토큰 k의 절대적인 출현비중을 보정하기 위하여 전체 샘플 중 토큰의 k의 출현횟수(f_k) 대비 특정 범주에서의 출현횟수($f_{k,indass}$)를 곱하여 토큰 k가 특정 범주에서 출현할 가능도($Likelihood_{k,indass}$)를 산출한다.

$$p_{k,indass} = p(token_k | class_{indass}) \quad (1)$$

$$p_{k,outdass} = p(token_k | class_{outdass}) \quad (2)$$

$$Likelihood_{k,indass} = \frac{f_{k,indass}}{f_k} \frac{p_{k,indass}}{p_{k,outdass}} \quad (3)$$

<Table 2>는 4개 감성 학습 범주에 따라 학습된 주요 n-gram 토큰을 나타낸다.

<Table 2> Example of 1 to 3-gram Tokens Assigned as High Likelihood for Each Class

Class 1	Production schedule (생산일정), return (반환), full-time (정규직), downgrading (하향), temporary workers (비정규직), donations (후원금), corruption (비리), similar (대동소이), early redemption (조기상환), sharing (분담), illegality (불법), sluggishness (지지부진), tax refund (세금환급), changes in industry conditions (업황변화), downgrade target price (목표가격 하향), contract return (계약 반환), temporary and regular workers (비정규직 정규직), launch delay (출시 연기), planning paid-in capital increase (유상증자 계획), reduction earnings (수익률 하향), inevitable reduction (하향 불가피), delayed reflection (지연 반영), return of rights (권리 반환), work disruption (조업 차질), hiring corruption (채용 비리), business downturn (업황 침체), demand dispersion (수요 분산), minimum wage increase (최저임금 상승), conversion to full-time employees (정규직 전환), discount rate expansion (할인율 확대), payment refund (계약금 반환), adjustment inevitable (조정 불가피), concerns realized (우려 현실화), business loss (사업 손실), lowered target price (목표가격 하향 조정), lowered performance reflected (실적 하향 반영), lower than before (종전 대비 하향), negative operating environment (부정적 영업 환경), conservative approach required (보수 접근 필요), slowing earnings momentum (실적 모멘텀 둔화), inevitable profitability decline (수익성 하락 불가피)
Class 2	Republican Party (공화당), connection fee (접속료), monthly rent (월세), net neutrality (망중립성), usage fee (사용료), receivables (미수금), lawsuit (피소), appeal trial (항소심), ordinary loss (경상 손실), store sale (점포 매각), sales slowdown (매출 둔화), owner family (오너 일가), cost deterioration (원가

	악화), burden increase (부담 증대), court ruling (법원 판결), market deterioration (시장 악화), industrial rate reorganization (산업 요금 개편), continuing deterioration in profitability (수익성 악화 지속), performance Short-term sluggishness (실적 부진 단기), expected loss reduction (손실 축소 예상), sluggish profitability improvement (부진 수익성 개선), rising cost price (원가 상승 가격)
Class 3	Internalization (내실화), bonus issue (무상증자), revenue ranking (매출순위), partnership (파트너쉽), superiority (능가), production share (생산비중), capital strength (자본력), welfare (복지), consolidation (결합), shareholder profit (주주이익), reform (개혁), new growth engine (신성장동력), plan expression (계획 표명), return to shareholder profit (주주 이익 환원), strengthening new growth engine (신성장동력 강화), consumer trust (소비자 신뢰), business normalization (영업 정상화), steady performance (실적 꾸준), improvement plan (개선 계획), improved operating profit record (개선 영업이익 기록), continued foreign buying (외국인 매수 지속), product spread recovery (제품 스프레드 회복), differential share price increase (차별 주가 상승), rate competition eased (요금 경쟁 완화), sales decline offset (매출 감소 상쇄)
Class 4	Upward (상향), tender offer (공개매수), high-end (고급화), net interest income (순이자이익), earnings surprise (어닝서프라이즈), pipeline (파이프라인), boom phase (호황국면), target price increase (목표가격 상향), upward top pick (상향 탑픽), highest expected price (예상 최고가), reserve of capacity (여력 보유), future growth potential (미래 성장성), realization of return on investment (투자수익률 실현), demand pull (수요 견인), multiple increase (멀티플 상향), deficit improvement (적자 개선), maintenance target price increase (유지 목표가격 상향), application of the highest expected price (예상 최고가 적용), increase in operating profit forecast (영업이익 전망 상향), reflection of profitability improvement (수익성 개선 반영)), portion of high value-added products (고부가 제품 비중), brisk sales expected (판매 호조 예상), subsidiary profit increase (자회사 이익 증가)

애널리스트 보고서 본문에 나타난 모든 단어의 감성 점수로서 가능도를 산출하여 사전화한 이후, 실제 애널리스트 보고서에서 사용된 개별 단어들의 가능도를 통합하여 보고서 자체의 감성 점수를 산출하여야 한다. 대부분의 선행연구에서는 긍정 범주와 부정 범주로 단어 사전을 구성한 이후, 문서에 긍정어가 출현할 시 +1점을 부여하고 반대로 부정어가 출현할 시 -1점을 부여하는 방식으로 문서의 감성을 산출한다. 그러나 본 연구에서는 특정 단어가 네 개 감성 범주에 속할 가능도를 통해 단어 사전을 구성하기 때문에, 특정 애널리스트 보고서가 네 개 감성 범주 각각에 속할 가능도를 산출할 수 있다. 이를 문서 d(doc_d)의 특정 감성범주에 대한 membership($Membership_{d,indass}$)으로 정의하고, 다음과 같이 산출하였다. $Inclusion(token_{k,}doc_{d})$은 문서 d에 특정 토큰 k($token_k$)가 출현하면 1의 값을, 출현하지 않으면 0의 값을 반환하는 지시함수(indicator function)이다. 최종적으로는 normalized membership을 사용하였는데, 각 보고서의 길이에 따라 계산된 membership의 편차를 보정하기 위하여 표준화한 값이다. $Class$는 4개 감성범주의 집합으로 정의하였다.

$$Membership_{d,indass} = \sum_{k=1}^{n} Inclusion(token_{k,}doc_{d}) \bullet \mathrm{Ln}(Likelihood_{k,indass}) \quad (4)$$

$$NormalizedMembership_{d,indass} = \frac{Membership_{d,indass}}{\sum_{indass \in Class} Membership_{d,indass}} \quad (5)$$

또한 감성변수 간 비교를 위해 재무·금융 분야에서 가장 우수한 성능을 보이고 있는 LM사전을 사용하였으며, LM사전으로 측정한 감성변수($Tone_{LM}$)를 다음과 같이 산출하였다.

$$Tone_{LM} = \frac{Pos_{LM} - Neg_{LM}}{Pos_{LM} + Neg_{LM}} \quad (6)$$

Pos_{LM}과 Neg_{LM}은 각각 LM사전으로 측정한 애널리스트 보고서 내 긍정어와 부정어 수를 의미한다. LM 감성사전의 긍정어 및 부정어 리스트는 구글 번역기를 통해 한글로 번역하여 사용하였다.[4)]

<Table 3>은 학습 이후 LM 사전으로 측정한 감성변수($Tone_{LM}$)와 본 연구의 감성사전으로 측정한 감성변수(MEM1, MEM2, MEM3, MEM4)에 대한 학습 데이터 셋의 기술 통계량을 보고하고 있다.

<Table 3> Descriptive Statistics of Sentiment Variables from Train Data Set Measured by LM Dictionary and Membership (1to3-gram) Model

	N	Mean	Median	Std.	1Q	3Q
Tone_LM	8,938	-0.010	0.000	0.333	-0.231	0.217
MEM1	8,938	0.171	0.169	0.014	0.162	0.177
MEM2	8,938	0.310	0.310	0.016	0.300	0.320
MEM3	8,938	0.315	0.316	0.016	0.305	0.326
MEM4	8,938	0.203	0.200	0.018	0.193	0.210

4) 일반인에게 공개되어 사용 가능한 한국어 감성사전으로서 대표적으로 서울대에서 개발한 KOSAC(Korean Sentiment Analysis Corpus), 군산대에서 개발한 KNU-한국어 감성사전 등이 있으나, 이는 범용 감성사전이기 때문에 재무 분야 텍스트의 감성 추출에 적합하지 않다. 현 시점에서 LM 감성사전은 재무 분야 특화 감성사전으로서 공개된 이후 재무 감성분석에서 지배적인 역할을 하고 있기 때문에 다수의 선행연구에서 이를 각국의 언어로 번역하여 연구에 활용하고 있다[1, 13].

$Tone_{LM}$은 평균 -0.01로 애널리스트 보고서가 긍정적 감성보다 부정적 감성을 더욱 나타낼 것으로 예상된다. 하지만 본 연구에서 개발한 감성변수로서 MEM1(0.171), MEM2(0.310), MEM3(0.315), MEM4(0.203)의 평균치를 살펴보면, 긍정 감성이 51.8%로 부정 감성(48.1%)보다 더 많이 분포하는 것으로 나타나 반대되는 결과를 보이고 있다. 이는 LM사전의 긍·부정 단어 구성비율 중 부정단어가 긍정단어의 약 6.5배 등록되어 있어 LM사전이 부정적 감성을 검출하기 용이한 구조로 설계되어 있기 때문으로 해석할 수 있다.

3.3 분류기 성능의 평가

본 연구에서는 자동으로 학습한 감성변수에 주식 가격 예측력이 존재하는지 여부를 판단하기 위하여 보고서 발행 이후 누적초과수익률을 예측하는 예측모형을 설계하였다. 다수의 선행연구들은 애널리스트 보고서 텍스트를 포함한 재무분야의 질적 정보가 시장 반응과 주식 가격에 영향을 미친다는 사실에 일반적으로 동의하고 있다[3, 10, 15, 21]. 그러나 이 같은 연구들은 회귀계수의 통계적 유의성과 회귀모형의 설명력을 바탕으로 결론내리고 있어 실제 인과관계(causal effect)에 대한 검증은 이루어지지 않았다. 모형의 인과관계는 예측력에 대한 검증으로서, 학습 데이터에 사용되지 않은 전혀 새로운 데이터에서 예측된 y값의 성능을 의미한다[24]. 따라서 본 연구의 예측 모형은 수익률 상승과 하락으로 구분된 아래와 같은 이진 분류(Binary Classification)를 목표로 한다. 예측 모형은 이진 분류에 널리 사용되는 로지스틱 회귀 모형(Logistic Regression Model)을 사용하였다.

$$y = \begin{cases} 1, & \text{if } CAR_{[0,+1]} \geq 0 \\ 0, & else \end{cases} \quad (7)$$

예측 모형에 대한 성과 평가는 분류 모형에 널리 사용되는 아래 네 가지 평가 지표를 사용하였으며, 검증 데이터로서 18년도 3,4분기 표본 외 검증 데이터(Out-of-Sample Test Set)에 대하여 상승 또는 하락 여부를 예측하였다.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

$$F1 - score = \frac{2 \times \text{Precision} \times Recall}{(\text{Precision} + Recall)} \quad (11)$$

*TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative

이후 각 평가 지표를 산출하고 각 모형의 예측 성과를 비교·분석하였다. 여기서 '양성(Positive)'과 '음성(Negative)'은 모형이 예측한 범주를 나타내며, '참(True)'과 '거짓(False)'은 모형이 예측한 범주가 실제와 동일한지 여부를 나타낸다.

예측에는 본 연구의 감성사전이 기존의 재무 분야 감성사전보다 우수한 예측력을 보이는지 여부를 판단하고자 LM 감성사전을 통해 산출한 감성변수와 본 연구의 membership 감성사전을 통해 산출한 감성변수를 각각 사용하였다. 또한 본 연구의 membership 감성사전은 1-gram부터 2-gram, 3-gram 각각을 사용한 모형과 측정한 n-gram 전체(1to3-gram)를 사용한 모형으로 구분하여 n-gram별 예측력에 차이가 있는지 비교·분석하였다.

4. 결과 및 해석

본 연구에서 개발한 자동 학습 감성사전을 통해 개별 보고서가 지니는 감성을 추출한 이후, 추출한 감성 변수를 통해 실제 주식 가격의 상승과 하락을 예측하였다. <Table 4>는 서로 다른 감성변수를 사용한 CAR 예측모형의 분류 결과를 나타낸다.

<Table 4> Out-of-Sample Binary Classification Results by Using Different Sentiment Variables (10-fold Cross Validation)

Metric	LM	Membership			
		1-gram	2-gram	3-gram	1to3-gram
Accuracy	0.57	0.56	0.58	0.61	0.59
Precision	0.54	0.53	0.55	0.58	0.55
Recall	0.88	0.89	0.75	0.73	0.85
F1-score	0.67	0.66	0.63	0.64	0.67

분석 결과 본 연구에서 제안하는 membership 감성변수를 사용한 모형이 LM사전의 감성변수를 사용한 모형에 비하여 전체적인 정확도(Accuracy)가 더 높게 나타났으며, 최소 1%에서 최대 4%의 정확도 차이를 보였다(1-gram만 사용한 예측 모형 제외). 정밀도(Precision) 측면에서도 마찬가지로 LM사전의 감성변수를 사용한 모형보다 membership 감성변수를 사용한 모형이 더 우수한 성능 지표를 보였는데(최소 1%에서 최대 4%), 이는 주가예측 시 실용적 측면에서 재현율(Recall)에 비하여

<Table 5> Student's t-test Result of Average $CAR_{[0,+1]}$ from Predicted Classes after 10-fold Cross Validation

Predicted Class by Classifier	(1) LM		Membership								$CAR_{[0,+1]}$ Diff.	t-value (p-value)
			(2) 1-gram		(3) 2-gram		(4) 3-gram		(5) 1to3-gram			
	Avg. $CAR_{[0,+1]}$	N	Avg. $CAR_{[0,+1]}$	N	Avg. $CAR_{[0,+1]}$	N	Avg. $CAR_{[0,+1]}$	N	Avg. $CAR_{[0,+1]}$	N		
Positive $CAR_{[0,+1]}$	0.208	12,316	0.131	12,648							-0.077	-1.802 (0.072)
					0.360	9,920					0.152	3.332 (0.001)
							0.511	9,130			0.303	6.563 (0.000)
									0.316	11,604	0.108	2.474 (0.013)
Negative $CAR_{[0,+1]}$	-1.056	2,984	-0.847	2,652							0.209	2.125 (0.034)
					-0.840	5,070					0.216	2.675 (0.007)
							-0.913	5,860			0.143	1.781 (0.075)
									-1.142	3,696	-0.086	-1.144 (0.253)
$CAR_{[0,+1]}$ Diff.	1.264		0.978		1.200		1.424		1.458			
t-value (p-value)	18.084 (0.000)		13.305 (0.000)		20.278 (0.000)		24.995 (0.000)		22.873 (0.000)			

정밀도가 더욱 중요하기 때문에 보다 의미를 가진다. 이는 주가예측에서 '위음성(False Negative)'의 비용보다 '위양성(False Positive)'에 대한 비용이 더 크다는 사실을 의미한다. 즉, 향후 본 모형의 예측결과를 바탕으로 실제 주식 투자를 수행할 경우, 모형이 특정 종목의 수익률 상승을 예측했을 때에(모형이 '양성'이라고 분류했을 때) 실제로 해당 종목 수익률이 상승하는지 여부(분류 결과가 '참'인 경우)가 투자자에게는 더욱 중요한 사실이기 때문이다.

N-gram 변수 간 예측 성능 지표를 살펴보면, 1to3-gram이 가장 다양한 긍·부정 표현들을 고려하도록 설계되었음에도 불구하고 Membership 모델 중 가장 높은 성능을 보이지는 않았다. 이는 만약 1-gram만으로 예측한 주가의 등락이 틀렸을 경우 이를 포함하는 1to3-gram 모형에서는 오히려 예측에 있어 일종의 노이즈(noise)로 작용하였기 때문인 것으로 해석할 수 있다.

이 같은 맥락에서, <Table 5>는 각 모형의 예측(분류) 결과로서 두 개 집단(양의 CAR 집단과 음의 CAR 집단)에 대한 실제 $CAR_{[0,+1]}$의 평균차이검정(t-test) 결과를 나타낸다. 즉, 모형이 특정 종목의 수익률 상승을 예측한 경우(양의 CAR 집단) 실제 CAR이 높다면, 모형의 유용성이 크다고 해석할 수 있다. 반대로 모형이 특정 종목의 수익률 하락을 예측한 경우(음의 CAR 집단) 실제 CAR이 낮아야 할 것이다. 개별 모형이 예측한 양의 CAR 집단과 예측된 음의 CAR 집단 간 평균차이는 모두 통계적으로 유의하게 나타났다. 절대적인 평균차이는 1 to 3-gram membership을 이용한 모형(모형 5)이 가장 큰 차이를 보였는데, 예측된 양의 CAR 집단이 예측된 음의 CAR 집단보다 실제 1.458% 큰 수익률을 나타냈다. 또한 각 모형이 예측한 양의 CAR 집단에 대해서 실제 평균 CAR에 대한 모형 간 평균차이검정을 실시한 결과, 최소 0.108%에서 최대 0.303% 차이로 membership 감성변수를 사용한 모형이 LM사전의 감성변수를 사용한 모형보다 더 나은 예측결과를 보였다. 반대로 각 모형이 예측한 음의 CAR 집단에 대해서 동일한 방식으로 평균차이검정을 실시한 결과, 최소 0.143%에서 최대 0.209% 차이로 LM사전의 감성변수를 이용한 모형이 더 나은 예측 결과를 보였다. 다만 1 to 3-gram membership을 사용한 모형 (모형 5)은 LM사전의 감성변수를 이용한 모형보다 더 나은 예측결과를 보였으나, 이러한 차이는 통계적으로 유의하지 않았다(t=-1.144). 종합하여 해석하면, LM 감성사전의 경우 부정적 감성에 대해서는 국내 증권사의 애널리스트 보고서에서 사용한 감성 단어보다 더 나은 예측 결과를 보였으나 긍정적 감성에 대해서는 본 연구의 membership 감성변수가 효과적인 예측결과를 나타내었다.

이러한 결과는 다음의 두 가지 관점에서 해석할 수 있다. 첫째, LM사전은 긍정어보다 부정어의 비율이 절대적으로 많기 때문에 보고서 본문에 나타난 부정어를 더욱 잘 검출하고, 때문에 보고서에 나타난 부정적 감성을 증

폭시키는 경향이 있다. 둘째, 학습 데이터 셋보다 검증 데이터 셋의 CAR가 음의 수익률에 치우쳐 있다. 이는 <Table 6>의 기술통계량에서 학습 데이터 셋과 검증 데이터 셋 간 CAR의 분포 차이를 확인할 수 있다. 즉, 2018년 3분기~4분기 동안의 시장 상황이 다소 부정적인 상황이었음을 알 수 있다. 이 같은 검증 데이터 셋 기간의 특성을 고려하면, 분류 모형의 예측 결과는 합리적인 수준이라고 해석할 수 있다.

<Table 6> Descriptive Statistics of two-days Cumulative Abnormal Returns ($CAR_{[0,+1]}$) from Train and Test Data Set

Data set	N	Mean	Median	Std.	1Q	3Q
Train (CY16Q1 ~CY18Q2	8,958 (85%)	0.29	0.19	3.58	-1.66	2.19
Test (CY18Q3 ~CY18Q4)	1,532 (15%)	-0.04	-0.12	3.46	-2.16	2.09

5. 결 론

본 연구는 애널리스트 보고서의 본문 텍스트 정보를 통해 자동 감성 학습이 가능한 감성사전을 개발하고 실제 주가 변화에 대한 예측력을 검증하였다. 국내 시가총액 상위 100개 기업을 대상으로 2016년부터 2018년까지 발행된 애널리스트 보고서를 수집하여 보고서 본문에 사용된 개별 단어에 대한 감성을 학습하였다. 특히, 감성 학습 과정에서 기존 기계학습에서와 같이 연구자의 수동 학습(라벨링) 과정을 필요로 하지 않는 자동 감성 학습 알고리즘을 개발하였다. 이후 실제 학습된 감성에 대한 주가 변화 예측력을 검증하기 위하여 보고서 발행 이후 주가 상승 또는 하락을 예측하는 이진 분류 모형을 활용하였다. 본 연구에서 제안하는 감성사전은 현 시점 재무 분야 감성사전에서 지배적이라고 할 수 있는 LM 감성사전에 비하여 실제 주가 변화에 대하여 1%~4% 우수한 예측 성능을 보였다. 따라서 본 연구에서 제안한 감성 분석 알고리즘은 LM 감성사전과 유사한 성능을 보임과 동시에 비용 측면에서 효율적이라고 해석할 수 있다. 실제 예측된 상승과 하락 집단의 실제 평균 수익률 차이를 검증한 결과, 수익률 상승 집단에 대해서는 본 연구에서 제안한 감성사전을 활용한 모형이 우수한 성능을 보였으며, 반대로 수익률 하락 집단에 대해서는 기존 LM 감성사전을 활용한 모형이 우수한 성능을 보였다.

본 연구는 다음과 같은 측면에서 기존의 연구와 차별점을 가진다. 첫째, 본 연구에서 제안하는 감성 학습 과정은 별도의 수동 라벨링 과정이 필요하지 않기 때문에 학습 데이터의 규모와 관련 없이 효율적인 학습이 가능하다. 둘째, 본 연구는 제안한 감성사전을 통해 실제 주식 가격의 예측력을 검증함으로써 기존 선행연구들이 텍스트 정보의 설명력을 검증한 결과와 차별화된다. 셋째, 본 연구에서 제안한 감성 사전은 베이즈 확률에 기반하여 자동 학습하기 때문에 확률의 갱신 주기를 최소화할 수 있다. 즉, 추가적인 학습 데이터가 주어질 경우 실시간으로 그러한 정보를 사전이 반영하고 투자자의 의사결정에 도움을 줄 수 있다. 넷째, 본 연구에서 제안한 감성 학습 알고리즘은 애널리스트 보고서 뿐만 아니라 실적 보고서, IR 보고서, 언론기사 또는 소셜미디어 등의 재무 분야 텍스트와 텍스트 발행 이후 실제 시장수익률을 비교함으로써 손쉽게 확장·적용이 가능하다.

그럼에도 불구하고 본 연구는 다음의 한계점을 가진다. 첫째, 국내 주식시장에서 규모가 큰 종목에 대해 발행한 애널리스트 보고서를 감성 학습에 사용하였기 때문에 개별 종목이나 개별 종목이 속한 산업이나 섹터별 특성을 반영하지 않는다. 둘째, 본 연구에서는 애널리스트 보고서가 제공하는 질적 정보에 집중하여 본문 텍스트의 감성만을 수익률 변화 예측에 사용하였으나, 실제로 투자자가 본 연구의 분석 결과를 보다 효과적으로 활용하기 위해서는 개별 주식이나 시장 상황에 대한 기술적(technical) 또는 기본적(fundamental) 분석이 함께 이루어져야 할 것이다.

본 연구결과는 향후 다음과 같은 후속 연구를 수행함으로써 확장 및 보완할 수 있다. 첫째, 향후 감성 사전의 학습을 산업별로 차별화함으로써 각 산업의 특성을 사전에 반영할 수 있다. 예를 들어, '달러 강세', '유가 하락' 등의 단어는 산업에 따라 긍정적이거나 반대로 부정적 의미를 내포할 수 있기 때문이다. 둘째, 본 연구에서는 감성변수의 예측력을 검증하기 위하여 선형 모형인 로지스틱 회귀모형을 사용하였으나, 기타 비선형 분류 모형으로서 랜덤 포레스트, 서포트 벡터 머신, 또는 신경망 등의 다양한 예측모형을 적용하고 분석 결과를 비교할 수 있을 것으로 기대된다.

Acknowledgement

This paper was supported by the research fund of the National Research Foundation of Korea (NRF-2019S1A5A2A03038389).

References

[1] Bannier, C., Pauls, T. and Walter, A., Content Analysis

of Business Communication: Introducing a German Dictionary, *Journal of Business Economics*, 2019, Vol. 89, No. 1, pp. 79-123.

[2] Bollen J., Mao H., and Zeng X., Twitter Mood Predicts the Stock Market, *Journal of Computational Science*, 2011, Vol. 2, No. 1, pp. 1-8.

[3] Brockman, P., Li, X. and Price, S.M., Conference Call Tone and Stock Returns: Evidence from the Stock Exchange of Hong Kong, *Asia-Pacific Journal of Financial Studies*, 2017, Vol. 46, No. 5, pp. 667-685.

[4] Buehlmaier, M.M. and Whited, T.M., Are Financial Constraints Priced? Evidence from Textual Analysis, *The Review of Financial Studies*, 2018, Vol. 31, No. 7, pp. 2693-2728.

[5] Cambria E., Schuller B., Xia Y., and Havasi, C., New Avenues in Opinion Mining and Sentiment Analysis, *IEEE Intelligent Systems*, 2013, Vol. 28, No. 2, pp. 15-21.

[6] Chen, H., De, P., Hu, Y.J. and Hwang, B.H., Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media, *The Review of Financial Studies*, 2014, Vol. 27, No. 5, pp. 1367-1403.

[7] Cho, S.J., Kim, H.K. and Lee, K.K., Optimization of Investment Decision Making by Using Analysts' Target Prices, *Journal of Society of Korea Industrial and Systems Engineering*, 2020, Vol. 43, No. 4, pp. 229-235.

[8] Cho S.J., Kim H.K. and Yang C.W. Building the Korean Sentiment Lexicon for Finance(KOSELF), *Korean Journal of Financial Studies*, 2021, Vol. 50, No. 2, pp. 135-170.

[9] Conley, C. and Tosti-Kharas, J., Crowdsourcing Content Analysis for Managerial Research, *Management Decision*, 2014, Vol. 52, No. 4, pp. 675-688.

[10] Das S., and Chen M., Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web, *Management Science*, 2007, Vol. 53, No. 9, pp. 1375-1388.

[11] Deng, S., Mitsubuchi, T., Shioda, K., Shimada, T. and Sakurai, A., Combining Technical Analysis with Sentiment Analysis for Stock Price Prediction, *In Dependable, Autonomic and Secure Computing (DASC), 2011 IEEE Ninth International Conference*, 2011, pp. 800-807.

[12] Garca, D., Sentiment During Recessions, *The Journal of Finance*, 2013, Vol. 68, No. 3, pp. 1267-1300.

[13] Guo, H., Wang, Y., Wang, B. and Ge, Y., Does Prospectus AE Affect IPO Underpricing? A Content Analysis of the Chinese Stock Market, *International Review of Economics and Finance*, 2022, Vol. 82, pp. 1-12.

[14] Heidari M. and Felden, C., Financial Footnote Analysis: Developing a Text Mining Approach, *In Proceedings of International Conference on Data Mining (DMIN)*, 2015, pp. 10-16.

[15] Huang, A.H., Zang, A.Y. and Zheng, R., Evidence on the Information Content of Text in Analyst Reports, *Accounting Review*, 2014, Vol. 89, No. 6, pp. 2151-2180.

[16] Kim, H.S. and Kim, C.S., An Analysis for IT Proposal Evaluation Results using Big Data-based Opinion Mining, *Journal of Society of Korea Industrial and Systems Engineering*, 2018, Vol. 41, No. 1, pp. 1-10.

[17] Kim, Y., and Joh, S.W. Text Analysis for IPO Firms in Korea: Analysis of Korean Texts in Registration Statements via Machine Learning, *Korean Journal of Financial Studies*, 2019, Vol. 48, No. 2, pp. 215-235.

[18] Lee, E., and Park, C.G., Does Adoption of K-IFRS Increase Upward Bias in Analysts' Earnings Forecasts?, *The Korean Journal of Financial Management*, 2019, Vol. 36, No. 1, pp. 179-205.

[19] Li F., The Information Content of Forward-Looking Statements in Corporate Filings-A Naïve Bayesian Machine Learning Approach, *Journal of Accounting Research*, 2010, Vol. 48, No. 5, pp. 1049-1102.

[20] Li, F., Lundholm, R. and Minnis, M. A Measure of Competition Based on 10-K Filings, *Journal of Accounting Research*, 2013, Vol. 51, No. 2, pp. 399-436.

[21] Liang D., Pan Y., Du Q. and Zhu L., The Information Content of Analysts' Textual Reports and Stock Returns: Evidence from China, *Finance Research Letters*, 2022, Vol. 46, Part. B, pp. 1-6.

[22] Loughran T., and McDonald B., When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks, *The Journal of Finance*, 2011, Vol. 66, No. 1, pp. 35-65.

[23] Schumaker, R. P. and Chen, H., A Quantitative Stock Prediction System Based on Financial News, *Information Processing & Management*, 2009, Vol. 45, No. 5, pp. 571-583.

[24] Shmueli, G., To Explain or to Predict? *Statistical Science*, 2010, Vol. 25, pp. 289-310.

[25] Tetlock, P. C., Giving Content to Investor Sentiment: The Role of Media in the Stock Market, *The Journal of Finance*, 2007, Vol. 62, No. 3, pp. 1139-1168.

[26] Tetlock, P.C., Tsechansky, M.S. and Macskassy, S., More Than Words: Quantifying Language to Measure Firms' Fundamentals, *The Journal of Finance*, 2008, Vol. 63,

No. 3, pp. 1437-1467.

[27] Yang, C.W., Information Content of Analyst Report Title: Focusing on the Tone of Text, *The Korean Journal of Financial Management*, 2021, Vol. 38, No. 3, pp. 1-38.

[28] Yu, J.D. and Lee, I.S., A Prediction of Stock Price Through the Big-data Analysis, *Journal of Society of Korea Industrial and Systems Engineering*, 2018, Vol. 31, No. 3, pp. 154-161.

ORCID

Su-Ji Cho | https://orcid.org/0000-0003-1511-5348

Ki-Kwang Lee | https://orcid.org/0000-0003-2291-8376

Cheol-Won Yang | https://orcid.org/0000-0001-9023-5089