

회사채 신용등급 예측에 뉴스 텍스트 정보의 유용성*

이예진** · 노호석*** · 양철원****

〈요 약〉

본 연구는 국내 회사채 신용등급 예측에서 기존 재무변수에 추가하여 기업 뉴스와 같은 텍스트 정보를 고려하는 것의 유용성에 대해 검토하였다. 기업의 회계 결산 시점 6개월 후의 신용등급 변경을 예측하는 모형을 설정하여, 60개의 재무변수에 더하여 뉴스 텍스트 기반 변수인 뉴스지수(news index)를 추가하였다. 뉴스지수는 구글(Google)을 통해 검색한 뉴스를 한국어 버트(BERT; Bidirectional Encoder Representations from Transformers) 모형을 사용하여 긍정 및 부정 감성을 계산하였다. 뉴스지수를 추가한 모형이 기존 재무변수만을 사용한 것보다 성과를 향상시켰지만, 표본 수의 부족 등으로 인해 획기적인 성과 향상은 경험하지 못하였다. 이를 보충하기 위한 정성적 분석으로 뉴스지수를 만드는 과정에서 발견한 한계점을 6개 범주로 분류하여 자세히 기술하였다. 추가적인 연구를 통해 해당 한계점을 해결한다면 더욱 의미 있는 뉴스지수를 만들 수 있을 것으로 기대된다.

주제어 : 회사채, 신용등급, 뉴스, 감성분석, 머신러닝

논문접수일 : 2024년 03월 14일 논문수정일 : 2024년 04월 02일 논문게재확정일 : 2024년 04월 09일

* 본 논문에 대해 유익한 조언을 해 주신 익명의 심사자들에게 감사드립니다. 본 논문은 이예진의 석사학위 논문을 바탕으로 완성되었습니다. 양철원은 서울대학교 경영대학 증권금융연구소(Institute of Finance and Banking, SNU Business School; IFB SNU) 객원연구원으로 연구년 중이며 연구소의 지원에 감사드립니다.

** 제1저자, 숙명여자대학교 통계학과 대학원생, E-mail: leeyj@sookmyung.ac.kr

*** 공동저자, 숙명여자대학교 통계학과 교수, E-mail: hsnoh@sookmyung.ac.kr

**** 교신저자, 단국대학교 경영경제대학 경영학부 교수, E-mail: yang@dankook.ac.kr

I. 서 론

신용평가는 기업이 발행하는 채권의 상환 가능성을 등급으로 평가하여 투자자들에게 공시하는 과정이다. 채권은 원금, 이자, 만기가 정해져 있어 안정적으로 투자할 수 있는 상품으로 인식되고 있다. 하지만 채권 발행자의 신용도가 하락하게 된다면 채권의 가치가 떨어지고, 해당 기업이 부도가 나게 된다면 원금도 돌려받지 못할 위험이 존재한다. 그러므로 채권에 투자할 때, 전문성과 객관성을 지닌 신용평가사가 기업이 채무를 상환할 능력이 충분한지 또는 채권의 위험도가 높은 것인지 회사채 신용등급을 판단하도록 하고 있다. 이는 기업의 사업성을 평가한다는 우량도 측면과 기업이 부도날 위험이 있는지 판단하고자 하는 부도 위험도 측면의 작업으로 나누어 볼 수도 있다(황선웅, 2005). 회사채 신용등급은 전문 지식이 부족한 투자자에게 기업의 정보를 간편하게 제공하는 역할을 하여, 채권시장에서 가격이 합리적으로 형성될 수 있도록 유도하는 역할을 한다. 그러므로 미래의 회사채 신용등급을 통해 기업의 우량도와 부실성에 대한 정보를 획득할 수 있다. 이는 투자자들이 안정적으로 채권 및 주식에 투자할 수 있도록 돕는다.

기존의 신용등급 예측 연구들은 회계 및 재무제표에서 얻어지는 수치 데이터 기반으로 머신러닝, 딥러닝 등의 방법을 적용하는 연구들이 많았다(Huang et al., 2004; Park et al., 2018; Jabeur et al., 2020; Lee and Oh, 2020). 하지만 회사채 등급은 재무 상태나 경영 성과뿐만 아니라 미래 발전 가능성, 업계 현황 등이 종합적으로 고려되어 결정되기 때문에 수치 데이터만을 예측에 이용한다면 그 성과에 한계가 존재할 수밖에 없다. 예측 모형의 성능을 향상시킬 수 있는 방법을 크게 두 가지로 나눌 수 있다. 첫 번째는 기존의 변수들을 이용하여 복잡한 새로운 예측모형을 구축하는 것이다. 두 번째는 예측 모형의 문제를 잘 표현할 수 있는 새로운 변수를 모형에 추가하는 것이다. 두 번째 접근은 계산 비용이 낮고, 복잡한 모형의 구축 없이도 성능 향상을 기대할 수 있다는 점에서 선호된다. 최근 재무분야에서는 예측에 도움이 될 수 있는 비정형 데이터를 활용하여 만든 새로운 변수를 도입함으로써 예측 모형의 성능을 향상시키려는 작업들이 많이 시도되고 있다. 본 연구는 신용등급 예측에 유용한 새로운 변수를 기업의 신용등급과 관련된 뉴스를 통해 만들어 보고 그것이 신용등급 예측에 활용되었을 때 유용성 여부를 검토하였다.

본 연구는 다음과 같은 면에서 기존 문헌에 기여할 수 있으리라 생각된다. 첫째, 재무분야에서 신용등급 예측 관련된 최근 연구가 부족한 실정이며, 텍스트와 관련된 연구는 더욱 그러하다. 재무 연구자들은 주로 신용등급 공시일의 정보효과나 신용평가업에 대한 규제 및 제도와 연관하여 주로 연구하고 있다.¹⁾ 신용등급 예측모형과 연관하여서는 Pinches

and Mingo(1973), Kaplan and Urwitz(1979), Peavy and Edgar(1984)의 연구가 존재하지만, 상당한 시간이 경과했을 뿐 아니라 설명변수로 재무변수들만을 사용하고 있다. 본 연구는 최근 발달한 텍스트 정보를 추가하여 예측모형의 성과를 발전시키고자 하였다.

둘째, 한국어 뉴스를 대상으로 한 연구이다. 미국의 경우는 뉴스 텍스트를 신용등급 예측에 사용하는 시도를 한 논문이 존재한다. Lu et al.(2012)은 신용등급 변화 예측을 위한 프로빗 모형에 재무비율 변수와 함께 텍스트 정보를 추가하였으며, 그 결과 예측 성능이 향상됨을 발견하였다. 하지만, 한국에서는 신용등급 예측에 뉴스 정보를 사용한 연구는 아직 존재하지 않는다. 언어학적으로도 한국어는 영어와 구조가 다른 언어이기 때문에 영어의 결과도 한국어에서도 동일하게 발견되리라는 보장은 없다. 본 연구는 이를 실증 검증하였다. 또한 여러 사례를 통해 한국어 뉴스의 유용성과 한계점도 세밀하게 설명하고 있다.

본 연구의 주요 결과는 다음과 같다. 첫째, 기업의 재무결산부터 기업의 신용등급 변화 발표 전까지 신용등급 변화를 예측할 수 있는 뉴스 텍스트 정보가 인터넷상에 존재함을 확인하였다. 둘째, 그러한 뉴스 정보를 수치화하여 기업의 신용등급 예측 모형에서 사용하였다. 뉴스지수 정보를 재무관련 수치변수와 함께 사용하는 것이 모형의 성능향상에 기여하였다. 마지막으로, 인터넷상의 기업 관련 뉴스를 구글링하여 기업 신용등급 변화예측을 위한 뉴스지수를 만드는 접근방법의 한계를 자세하게 분석하여 향후 연구에 도움이 될 수 있도록 하였다.

본 연구는 다음과 같이 구성된다. 제Ⅱ장은 관련 문헌에 대해 소개한다. 제Ⅲ장은 연구에 사용된 자료와 변수에 대해 설명하며, 제Ⅳ장은 실증 분석 결과를 제시한다. 제Ⅴ장에서는 본 연구를 마무리한다.

Ⅱ. 문헌연구

기업의 신용평가와 관련된 국내외 연구는 회사채 신용등급 예측, 기업 부도 예측, 기업 평점 예측 등으로 정리해볼 수 있다(Park et al., 2018; Lee and Oh, 2020). 상당수 연구는 회계 및 재무제표에서 얻어진 수치 데이터 변수들을 이용하여 머신러닝, 딥러닝 기반 모형을 적합하여 예측을 진행하였다. 하지만 최근에는 텍스트와 같은 비정형 데이터를 함께 활용하여 기업의 신용등급을 예측하는 모형들이 연구되기 시작하고 있다. 본

1) 공시일의 정보효과에 대해서는 Holthausen and Leftwich(1986), Hand et al.(1992), Hite and Warga(1997), Dichev and Piotroski(2001)의 연구를, 신용평가업에 대한 규제 및 제도와 연관된 연구로는 Cheng and Neamtiu(2009), Bonsall(2014), Dimitrov et al.(2015)을 참조하시오.

연구에서는 현재까지 국내 회사채 신용등급 예측을 위해 수치 데이터와 함께 텍스트 데이터를 활용하는 것에 대해 연구가 이루어지지 않는 점에 착안하여 텍스트 데이터를 함께 활용하는 것이 국내 회사채 신용등급 예측에 있어서 가질 수 있는 유용성을 다각도로 검토하고자 한다.

1. 수치 데이터 기반 신용등급 예측 모형

Kim and Choi(2006)는 Kaplan and Urwitz(1979)의 다중선형 회귀모형과 Kim and Kim(2002)의 N-probit 모형을 비교하여 실제 기업의 신용을 평가하는 데 사용될 것으로 판단되는 유용한 재무 변수를 파악하는 연구를 진행하였다. 해당 연구에서는 회사채 등급이 존재하는 KOSDAQ 등록 기업을 표본으로 하였으며, 회사채 등급을 AA+를 15, AA를 14, ..., B-를 1로 변환한 15등급 체계를 이용하였다. Kim and Choi(2006)에서는 기업규모, 부채비율, 자본이익률, 누적이익률, 주가순자산비율 등 여러 재무 특성 및 미래의 수익성 전망을 수치화한 변수들이 신용등급 예측에 유용한 것으로 분석되었다.

Kim and Ahn(2016)의 연구에서는 한국신용정보에서 공시한 자료를 기준으로 A1, A2, A3, B, C로 나누어진 5등급 신용등급 체계에서 C를 제외한 네 개의 등급의 자료를 이용하여 신용등급 예측 모형을 적합하였고, 검증 데이터의 정분류율 측면에서 랜덤 포레스트 모형이 72.8%로 다중판별분석, 인공신경망, 다분류 SVM(support vector machine)보다 우수한 성능을 보였다. Oh et al.(2017)의 연구에서는 재무 정보, 시장 정보, 거시경제지표를 이용하여 기업부도 예측모형을 구축하였으며, 랜덤 포레스트 모형이 93.6%로 가장 높은 정확도를 보여주었다. Park et al.(2018)은 회사채 신용등급의 예측 성능이 우선시 되는 경우 랜덤 포레스트 모형을, 신용등급의 유지 및 관리 관점에서 접근할 경우에는, Elastic-net 벌점을 고려한 다중회귀모형 또는 순위 로짓, 프로빗 모형을 적용하는 것을 추천했다.

Oh et al.(2017)은 시간을 고려하는 생존분석론 방법론 중 하나인 Cox 비례위험 모형(Cox-PH model)과 인공지능방법의 하나인 RNN(Recurrent Neural Network) 모형을 사용하여 신용등급을 예측하는 것을 시도하였는데 각각의 정확도가 73.2%, 89.9%가 나와 다른 비교 대상 모형인 로지스틱 회귀모형, 랜덤 포레스트, SVM(Support Vector Machine)보다 예측력이 낮게 나타났다. Lee and Oh(2020)는 RNN 방법론 중 하나인 LSTM(Long Short-Term Memory)을 이용하여 5등급 체계로 되어 있는 기업의 신용등급을 예측하였으며 정확도, 정밀도, 재현율, F1-score는 각각 68%, 66%, 68%, 67%로 나타났다.

2. 수치 데이터와 텍스트를 함께 이용한 신용등급 예측 모형

수치 데이터뿐 아니라 비정형 데이터인 텍스트 정보를 추가하여 신용등급 예측 성능을 높이고자 하는 연구도 최근 들어서 진행되고 있다. Lu et al.(2012)은 신용등급 변화(하락, 유지, 상승)를 예측하는 프로빗 모형에 텍스트 정보를 추가하면 예측 성능이 소폭 향상된다는 것을 보여주었다. 분기별로 뉴스 보도빈도(news coverage), 주제(topic), 감성(sentiment)을 뉴스에서 추출하여 모형에 변수로 추가하였다. 텍스트 정보 기반 변수계산을 위해 Latent Dirichlet Allocation(LDA) 모형을 이용하여 주제를 추출하고 문장 기반 감성분석 모형인 OpinionFinder를 통해 뉴스의 감성을 계산했다. 프로빗 모형에 텍스트에서 추출한 변수를 추가했을 경우, Macro-F1이 32.2%에서 38.5%까지 상승하는 결과가 나타났다.

Mai et al.(2019)은 SEC(Securities Exchange Commission)에 제출하는 10-K 연간 신고서에서 Word2Vec 방법론을 활용하여 연간 신고서별 단어 임베딩 벡터들의 평균을 구하여 수치 데이터와 함께 활용하였다. 딥러닝 모형에 텍스트 정보를 함께 활용할 경우 부도 예측 모형의 성능이 로지스틱 회귀 모형, SVM, 랜덤 포레스트보다 평균적으로 20% 향상된다는 것을 보였다.

Oh(2017)은 부도 예측과 관련하여 텍스트 자료를 사용하였다. 그는 네이버 뉴스 검색 홈페이지를 활용하여 2010년부터 2016년까지 월별 뉴스를 수집하였으며, Word2Vec 방법론을 활용하여 ‘부도’ 및 ‘상장폐지’와 관련된 변수들을 생성하였다. 첫 번째는 ‘부도’ 및 ‘상장폐지’와 기사 내 언급된 모든 단어의 유사도가 높은 상위 20개를 찾아, 해당 상위 20개의 단어가 하나라도 포함된 기사는 부도 연관 기사로 간주하였다. 정의된 부도 관련 뉴스의 비율을 계산하여 부도 기사 비율 변수를 만들었다. 두 번째는 해당 기업의 기사를 구성하고 있는 모든 단어 벡터의 평균과 ‘부도’ 및 ‘상장폐지’의 단어 벡터의 유사도를 계산하여 부도 관련 단어와 유사도 변수를 만들었다. 수치 데이터와 함께 부도 기사 비율변수와 부도 관련 단어와 유사도 변수를 추가로 신용 평가 모형에 고려했을 때, 평균 예측 정확도가 0.9033에서 0.9098로 매우 미미하게 상승하는 것으로 나타났다.

이와 같이 현재까지 신용등급 예측을 위해 텍스트 데이터를 활용하는 것에 대한 국내외 연구를 살펴보았을 때, 국내 기업 회사채의 신용등급 예측에 텍스트 데이터를 고려하는 것에 유용성에 대한 연구는 이루어지지 않았다. 회사채 신용등급이 투자자에게 하는 정보 제공 역할이나 채권시장에서 합리적 가격 형성에 미치는 역할을 생각할 때, 그에 대한 연구의 필요성은 충분하다고 판단된다. 이에 본 연구에서는 국내 회사채 신용등급 예측 모형 구축에 있어서 각 회사의 신용등급에 대한 정보가 담겨 있는 텍스트 데이터를 적절히

추출하고 가공했을 때의 유용성에 대해 검토하고자 한다.

Ⅲ. 자료

회사채 신용등급 자료는 1998년부터 2020년까지 자료 중 신용등급과 재무제표 자료, 기타 기업 관련 자료가 모두 존재하는 403개의 KOSPI 상장 기업들의 자료를 사용하여 분석을 진행하였다. 본 연구에서는 각 기업의 결산 시점의 정보를 활용하여 6개월 후의 신용등급을 예측하는 것을 고려하였다. 이는 회계 결산 후 그 정보를 3개월 이내에 공시해야 한다는 점과 공시된 정보가 신용등급에 반영되는데 3개월 정도가 소요된다고 가정된 것이다.

1. 수치 데이터

회사채 신용등급 예측과 관련하여 각 기업의 결산 시점에서의 재무제표 자료 및 기타 기업 관련 자료로부터 총 60개의 변수를 추출하였다. 이는 기존 연구에서 기업의 신용등급에 예측에 유의한 영향을 주는 것으로 고려되었던 변수들이다. 본 연구에서는 이러한 변수들을 아래와 같이 규모 지표, 비재무 지표, 생산성 지표, 수익성 지표, 안정성 지표, 현금흐름 지표, 활동성 지표로 구분했다(Kim and Choi, 2006; Kim and Ahn, 2016; Park et al., 2018).

- 규모 지표: 자산총계(X1), 매출액(X2), 자기자본(X3), 부채총계(X4), 사원수(X5)
- 비재무 지표: 베타계수(X6), 누적시장조정 수익률(X7), 업력(X8), 주가순자산비율(X9)
- 생산성 지표: 1인당 매출액(X10), 경영자본회전율(X11)
- 수익성 지표: 자기자본이익률(X12), 누적이익률(X13), EBITDA/매출액(X14), 순금융비용부담율(X15), 영업이익/총자산(X16), 이자이익/총자산(X17), 주당 순이익(X18), 금융비용부담률(X19), 금융비용대총비용 비(X20), 이자 보상배율(X21), 매출액경상이익률(X22), 수익성비율(X23), 매출액총이익률(X24), EBIT/매출액(X25), 매출액순이익률(X26), 매출원가율(X27), 판관비율(X28), 경상이익/자기자본(X29)
- 안정성 지표: 부채비율(X30), 자기자본비율(X31), 유보액대총자산비율(X32), 고정자산 구성비율(X33), 재고자산대 유동자산비율(X34), 단기차입금대총차입금비율(X35), 당좌비율(X36), 장기부채/유동부채(X37), 유동비율(X38), 고정장기적합률(X39), 순차입금비율(X40), 유동부채비율(X41), 비유동

부채비율(X42), 단기차입금의존도(X43), 차입금의존도(X44), 순자산/부채총계(X45), 순운전자본구성비율(X46)

- 현금흐름 지표: 현금흐름대총자본비율(X47), (영업활동으로 인한 현금흐름 - 현금배당)/(고정자산 + 운전자본)(X48), 영업활동으로 인한 현금흐름(X49), EBITDA/순금융비용(X50), 영업활동 후의 현금흐름(X51), OCF대부채비율(X52) (영업활동으로 인한 현금흐름(operating cash flow; OCF)), OCF대총자본비율(X53), NCF대부채비율(X54)(순영업활동 현금흐름(net cash flow from operating activities; NCF)), NCF대매출액비율(X55), 총차입금/EBITDA(X56)
- 활동성 지표: 재고자산회전율(X57), 총자산회전율(X58), 매출채권회전율(X59), 매입채무회전율(X60)

본 연구에서는 추출된 60개의 변수 전체를 사용하지 않고, 일부 변수를 제외하고 분석하였다. 제외한 변수는 총 12개로 다음과 같다: 순금융비용부담율(X15), 금융비용부담률(X19), 매출액경상이익률(X22), EBIT/매출액(X25), 매출원가율(X27), 판관비율(X28), 유보액대총자산비율(X32), 단기차입금대총차입금비율(X35), 유동비율(X38), 현금흐름대총자본비율(X47), (영업활동으로 인한 현금흐름 - 현금배당)/(고정자산 + 운전자본)(X48), 영업활동 후의 현금흐름(X51).

<표 1> 회계 및 재무 변수들 중 상관계수의 절대값이 높은 조합들

본 표는 신용등급 변경 예측에 사용할 회계변수 및 재무변수들 중에서 상관계수의 절대값이 높은 조합을 보고하였다. 변수 X1-X60에 대한 정의는 본문 3.1절의 설명을 참고하시오.

변수 1	변수 2	상관계수(Corr.)
X32	X13	1.0000
X27	X24	-1.0000
X25	X14	0.9981
X51	X50	0.9925
X19	X15	0.9925
X47	X23	0.9896
X38	X36	0.9827
X22	X14	0.9732
X28	X14	-0.9640

<표 1>은 변수가 다른 변수와의 상관계수의 절대값이 0.95 이상이 되는 조합을 보여주고

있다. 이들 조합에서 한 변수는 정보 중첩의 문제를 고려하여 제거하였다. 구체적으로 매출원가율(X27)과 매출액총이익률(X24)은 서로 선형관계에 있어서 매출원가율(X27)을 제거하였고, 자기자본비율(X31)과 유보액대총자산비율(X32)은 본 자료에서는 실질적으로 동일한 변수이어서 유보액대총자산비율(X32)을 제거하였다. 다음의 세 변수 순금용비용 부담율(X15), 단기차입금대총차입금비율(X35), (영업활동으로 인한 현금흐름 - 현금배당)/(고정자산 + 운전자본)(X48)은 전체의 약 1/3이 결측되어 있어 분석에서 제거하였다.

결과적으로 본 연구에서는 총 48개의 변수가 고려되었으며 R 패키지 mice를 이용해 결측값을 대체하여 분석에 사용하였다. 결측값 대체 후, 기존 연구와 같이 각 변수의 상·하위 5% 이상 및 이하 값들에 대해 윈저화(winsorization) 방법을 적용하였다. 윈저화 이후 분산이 큰 변수인 X1, X2, X3, X4, X5, X10, X49에 대해 분산 안정화를 위해 로그 변환을 하였으며, X49는 음수값이 존재하여 최솟값의 절댓값에 1을 더한 후 로그 변환을 진행하였다.

위 전처리를 끝낸 각 변수의 결산 시점 직전 최근 5년 데이터의 평균을 결산 시점에서의 변수 X와 대비하여 BX로 표기하고 $(X-BX)/BX$ 로 변환한 것을 신용등급 예측 모형의 최종 설명변수로 사용했다. 예를 들어 2021년 결산 기업의 재무제표 정보 및 기업 관련 변수가 X1-X60일 때, BX1-BX60은 2015년부터 2020년 사이에 존재하는 X1-X60을 각각 변수별로 평균을 내어 얻어진다. 특정 기업 관측치가 하나의 변수에 대해서라도 BX 값이 없어서 $(X-BX)/BX$ 가 정의되지 않는 경우 해당 기업 관측치 전체를 분석에 포함하지 않았다. 또한 BX가 0일 경우, 최종 설명변수 $((X-BX)/BX)$ 를 만들 때 분모가 0이 되는 것을 방지하기 위해 0.1^9 을 분모에 더하였다. $(X-BX)/BX$ 는 최근 5년의 평균값과 비교하여 지난 한 해 일어난 변화의 상대적 비율을 나타낸다. 아울러 본 연구에서는 기업의 회계 결산 시점에서의 22등급 체계상에서의 신용등급(AAA부터 D까지)을 설명변수에 추가하였다. 예측하고자 하는 반응변수는 회계 결산 시점에서 6개월 후 등급 변화(-1: 하락, 0: 유지, 1: 상승)로 하였다. 6개월 후 신용등급 자체를 예측 대상으로 하지 않은 이유는 22등급 신용등급 체계에서는 클래스별 자료 수의 불균형이 매우 심하기 때문이다. 또한, 회사채 신용등급은 등급의 변화가 있을 때 두 등급 이상 변화되는 경우가 많지 않기 때문에, 등급 변화를 예측한 것과 절대적 신용등급 자체를 예측하는 것이 유용성 측면에서 크게 다르지 않다고 판단하였다. 위 절차를 통해 수집된 최종 대상 데이터는 총 3,544개이다.

<표 2>에서는 각 자료를 결산 시점 기준으로 연도별로 정리했을 때 반응변수인 신용등급 변화의 분포를 보여주고 있다. 회계 결산 당시의 신용등급을 6개월 후에 그대로 유지한 경우는 약 85% 정도이고 15% 정도는 신용등급이 상승 또는 하락하였다.

<표 2> 관측된 신용등급 변경의 관측수

본 표는 연도별로 관측된 신용등급 변경의 관측수를 보고한다. 신용등급 변경 -1은 신용등급 하락, 0은 유지, 1은 상승을 의미한다.

연도	-1(하락)	0(유지)	1(상승)	합계
1998	2	23	4	29
1999	11	108	31	150
2000	7	154	28	189
2001	11	155	17	183
2002	10	111	30	151
2003	3	116	17	136
2004	0	125	22	147
2005	5	127	15	147
2006	7	122	13	142
2007	1	109	23	133
2008	6	123	11	140
2009	6	117	29	152
2010	4	138	18	160
2011	7	151	10	168
2012	11	154	13	178
2013	21	162	9	192
2014	28	166	8	202
2015	19	163	5	187
2016	8	166	9	183
2017	8	171	10	189
2018	14	172	6	192
2019	13	178	3	194
합계	202	3,011	331	3,544
비율 (%)	5.7	84.96	9.34	100

IV. 실증분석 결과

1. 수치 데이터 기반 신용등급 변경 예측

재무관련 변수를 이용한 회사채 신용등급 및 기업 부도 예측 선행연구에서 랜덤 포레스트(random forest)가 예측 성능이 우수한 것으로 분석되었다(Kim and Ahn, 2016; Choi et al., 2017; Park et al., 2018). 따라서 본 연구는 선행연구를 참고하여 랜덤 포레스트를 텍스트 정보의 유용성 검정을 위한 기준 모형으로 선정하였으며 R 패키지 randomForest를 사용하여 예측 모형을 적합하였다. 설명변수로는 최근 5년의 평균값과 비교하여 지난 한 해 일어난 변화의 상대적 비율의 형태로 변환된 48개의 재무 관련 변수와 결산 시점 기준 신용등급을 사용하였으며 재무제표 결산 시점의 신용등급으로부터 6개월 뒤의 등급

변화(C)를 반응변수(-1: 하락, 0: 유지, 1: 상승)로 하는 랜덤 포레스트 모형을 구축하였다.

랜덤 포레스트의 모형을 결정하는 하이퍼파라미터(hyperparameter)는 R 패키지에서 기본값으로 제공하는 값으로 적용하였는데 랜덤 포레스트를 구성하는 의사결정 나무의 개수(ntree)는 500, 사용되는 설명변수 개수(mtry)는 $\sqrt{p} = \sqrt{49} \approx 7$ (p는 설명변수의 개수)로 지정하였다. 신용등급 변화 예측은 각 기업 자료의 결산 시점 연도를 기준으로 5년간의 데이터를 학습시켜 다음 1년에 해당하는 기업 자료에 대해 신용등급 변화를 예측하는 것으로 하였다. 회사채 등급 변화(C)의 클래스 불균형(class imbalance)을 해결하기 위해 R 패키지 scutr의 SMOTE(Synthetic Minority Over-sampling Technique) 방법론을 이용하여 5년 치 학습 데이터의 모든 클래스의 자료 수를 동일하게 맞춘 후 랜덤포레스트 모형에 적합하였다.

<표 3>은 1998년부터 2019년까지 1년씩 시간을 롤링(rolling)하는 방식으로 5년치 학습 데이터로 다음 1년치 신용등급 변화를 예측한 결과를 종합하여 작성한 혼동행렬(confusion matrix)이다. 이 혼동행렬을 기반으로 분류성능 평가척도인 정밀도(Precision), 재현율(Recall), F1-score는 모두 클래스별로 계산된다. 정밀도는 모형이 특정 클래스라고 예측한 자료 중에 실제로 그 특정 클래스인 자료의 비율이고, 재현율은 특정 클래스의 자료 중에서 예측모형에 의해 그 특정 클래스라고 예측된 자료의 비율이다. 클래스별 F1-score는 그 클래스의 정밀도와 재현율의 조화평균으로 클래스 불균형이 존재하는 분류데이터에 대해 분류성능을 평가하기에 적합한 지표이다. 이를 구체적으로 다음 식 (1)로 나타낼 수 있다.

<표 3> 수치데이터 기반 신용등급 변경 예측모형에 대한 혼동 행렬

본 표는 신용등급의 실제 클래스 대비 예측된 신용등급 클래스의 혼동행렬을 보고한다. 신용등급 변경 -1은 신용등급 하락, 0은 유지, 1은 상승을 의미한다.

		실제 클래스		
		-1	0	1
예측된 클래스	-1	37	124	0
	0	96	2207	157
	1	4	190	27

<표 4> 수치데이터 기반 신용등급 변경 예측모형의 클래스별 예측성과

본 표는 회계 및 재무변수 등의 수치데이터를 기반으로 신용등급 변경을 예측한 모형의 클래스별 성과를 보여주고 있다. 신용등급 변경 -1은 신용등급 하락, 0은 유지, 1은 상승을 의미한다. Precision, Recall, F1-score의 정의는 본문의 식 (1)을 참고하시오.

	-1	0	1	평균값
Precision	0.2298	0.8972	0.1222	0.4164
Recall	0.2701	0.8754	0.1467	0.4309
F1-score	0.2483	0.8862	0.1333	0.4226

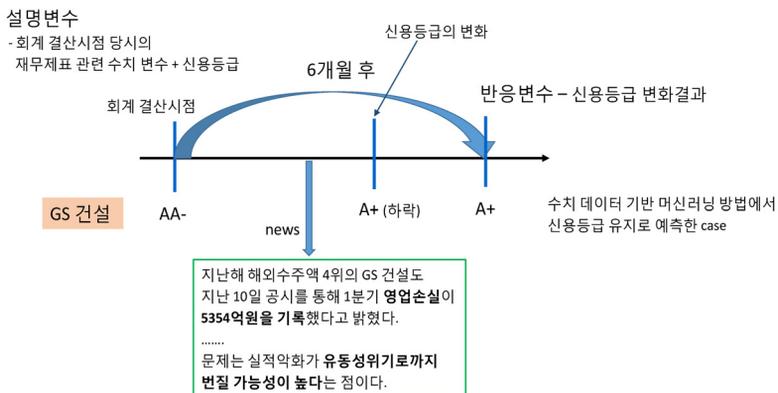
$$Precision = \frac{tp}{tp + fp}, Recall = \frac{tp}{tp + fn}, F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (1)$$

여기서, 클래스별 tp 는 해당하는 클래스를 정확하게 예측한 관측치 수, fp 는 해당하는 클래스가 아닌데 해당 클래스로 예측한 관측치 수, fn 는 실제 해당 클래스인데 다른 클래스로 예측한 관측치 수이다.

<표 4>에서 클래스 0의 tp 는 2,207이 되고 fp 는 253(=96+157), fn 는 314(=124+190)가 된다. 표의 신용등급 변화 클래스별 예측결과를 보면 수치데이터 기반 랜덤 포레스트 모형이 등급변화가 없는 경우에 대해서는 잘 예측하지만 등급이 상향 또는 하향되는 경우에 대해서는 정밀도와 재현율 모두 나쁜 것으로 나타났다. 이에 수치데이터 기반 분석의 한계를 보완해 줄 수 있는 텍스트 정보의 존재 여부를 검토하고자 하였다. 그러한 검토의 하나의 방법으로 수치데이터 기반 랜덤 포레스트 모형이 오분류한 관측치들을 정분류하는데 도움이 될 수 있는 유용한 텍스트 정보인 뉴스가 존재하는지 먼저 살펴보고자 하였다.

예측에 유용한 뉴스를 탐색하기 위한 기간으로 재무제표 결산시점과 6개월 후의 신용등급이 실질적으로 결정된 시점 사이의 기간을 고려했다. [그림 1]의 GS 건설의 예를 보면 2012년 12월 신용등급이 AA-이었지만 2013년 상반기에 신용등급이 A+로 하락되는 변화를 거쳐 2013년 6월의 신용등급이 A+가 되었다. 이러한 경우에 신용등급이 A+로 하락된 것이 어느 한 신용평가회사를 통해서라도 발표되기 시작한 이후의 뉴스는 신용등급 변화를 예측하는데 도움이 되는 뉴스라고 할 수 없다. 본 연구에서는 일정한 알고리즘을 통해(42절에 제시됨) 각 관측치의 6개월 후 신용등급이 재무제표 결산시점 이후에 실질적으로 결정된 시점을 결정하고 그 시점 이전의 뉴스 중에 수치기반 머신러닝 모형이 오분류한 관측치들을

[그림 1] 신용등급 변경 예측에 있어서 뉴스텍스트의 필요성



정분류하는 도움이 될 수 있는 뉴스를 찾고자 하였다. 실제로 [그림 1]의 GS 건설의 예에서는 2013년 6월의 신용등급이 A+ 등급(하락)은 2013년 5월 24일에 결정되었는데 2013년 4월 18일에 등급이 하락될 것을 예상할 수 있는 영업손실과 유동성 위기를 언급하는 뉴스가 있었던 것을 볼 수 있었다.

구체적으로 구글(Google)에서 재무제표 결산 시점(결산월 1일로 설정)으로부터 신용등급 결정날짜(4.2절에서 자세히 설명되어 있음) 직전까지의 기간 동안 기업명과 ‘신용등급’을 키워드로 설정하여 검토하였다. 구글 검색 결과는 관련성이 높은 순으로 정렬되므로 의미 있는 결과는 첫 페이지에 나올 가능성이 크다고 판단하여 첫 페이지에 나온 신용등급 변화와 관련된 뉴스들 중에 정분류에 도움이 될 수 있는 뉴스가 존재했는지를 정성적으로 검토하였다.

<표 5>는 수치 데이터 기반 랜덤 포레스트 모형이 오분류한 571개의 관측치와 연관하여 신용등급 변화가 잘못 예측된 자료의 등급 상향이나 하락을 올바르게 예상하는 데 도움이 되는 39개의 뉴스를 보여준다. 그러한 뉴스들의 연도별 관측치 수는 2010년 1개, 2011년 1개, 2013년 2개, 2014년 1개, 2015년 5개, 2016년 8개, 2018년 2개, 2019년 9개, 2020년 10개로 2010년 이전에는 존재하지 않으며 최근으로 가까워질수록 신용등급 예측에 유용한 텍스트 정보가 많아지는 것을 알 수 있었다. 신용등급 변화 예측에 유용한 뉴스 중에는 특정 기업을 지칭하여 작성된 뉴스도 존재했지만 여러 기업에 대해 종합적으로 다루는 뉴스도 존재하여 자동화된 수집 방법으로는 특정 기업의 신용등급 변화예측에 유용한 텍스트 정보수집이 어려워 보이는 측면이 있었다.

<표 5> 신용등급 변경 예측에 도움이 되는 뉴스

본 표는 신용등급 변경 예측에 도움이 되는 뉴스들의 내역을 보여주고 있다. 구글을 통해 검색한 기업의 신용등급 관련 뉴스를 대상으로 하여, 그 중에서 수치 데이터 기반 랜덤포레스트 모형이 오분류한 571개의 관측치와 연관하여 신용등급 변화가 잘못 예측된 자료의 등급 상향이나 하락을 올바르게 예상하는 데 도움이 된 39개 뉴스를 보여준다.

기업	이전 신용 등급	바뀐 신용 등급	신용등급 변화 날짜	뉴스 발표 날짜	핵심 뉴스 내용
삼성SDI	AA-	AA	2010-06-29	2010-02-01	국내 신용평가사가 부여한 삼성SDI의 회사채 신용등급은 AA-. 등급전망은 ‘긍정적’이다.
LS	A+	AA-	2011-06-24	2011-05-23	LS그룹, 1분기 실적 호조...계열사들 선전 “눈에 띄네”
동양	BB	BB-	2013-06-14	2012-12-14	동양그룹, M&A 전략 차질... 결정타는 건설 부진
GS건설	AA-	A+	2013-05-24	2013-04-18	지난해 해외수주액 4위의 GS건설도 지난 10일 공시를 통해 1분기 영업손실이

기업	이전 신용 등급	바뀐 신용 등급	신용등급 변화 날짜	뉴스 발표 날짜	핵심 뉴스 내용
한진해운	BBB+	BBB-	2014-03-13	2014-01-16	5,354억 원을 기록했다고 밝혔다... 문제는 실적악화가 유동성 위기로까지 번질 가능성이 높다는 점이다. 조선·해운 관련 업종(STX엔진, 한진 중공업, STX팬오션, SK해운, 현대상 선, 한진해운)의 업종이 전년에 이어 2013년에도 등급하향을 주도한 것으로 조사됐다.
GS건설	A+	A	2015-02-25	2014-12-19	지난 15일 나이스신평은 GS건설(A+ 급)과 태영건설(A급) 회사채 신용등급 을 부정적 관찰대상에 올렸다
태영건설	A	A-	2015-04-20	2014-12-19	지난 15일 나이스신평은 GS건설(A+ 급)과 태영건설(A급) 회사채 신용등급 을 부정적 관찰대상에 올렸다.
삼성중공업	AA	AA-	2015-05-21	2014-12-31	한신평과 NICE신평은 내년 조선업에 대해 '부정적', 한기평은 '보수적'으로 평가한다고 밝혔다.
삼성엔지니어링	A+	A	2015-05-22	2015-04-29	... 삼성중공업은 삼성엔지니어링과의 합 병을 준비하며 3,152억 원의 자사주를 매입해 재무 부담이 확대됐다. 한신평은 또 삼성엔지니어링과 SK건 설의 등급 전망을 '안정적'에서 '부정적' 으로 동반 하향 조정했다.
SK하이닉스	A+	AA-	2015-05-27	2015-01-28	SK하이닉스 등 전자업종 옐로우칩, 지 난해 '날았다'
두산	A	A-	2016-02-19	2016-02-12	'대규모 손실에 등급 강등 위기'의 두산 그룹, 돌파구는 있다
LS	AA-	A+	2016-04-26	2016-03-30	최근 크레딧시장에서 LS그룹의 재무 부담 확대 우려와 신용등급 조정 가능 성이 거론되고 있다.
삼성중공업	A+	A-	2016-05-27	2016-05-05	한국기업평가는 현대중공업과 삼성중 공업의 신용등급을 똑같이 'A+(부정 적)'로 평가하고 있다.
현대미포조선	A	A-	2016-06-08	2016-05-05	현대미포조선의 신용등급은 한 단계 낮 은 'A(부정적)'다.
한라홀딩스	A+	A	2016-05-31	2016-03-10	NICE신용평가는 지난 8일 한라홀딩스 신용등급을 A+로 유지하며 하향 검토 대상에 등재했다.
CJ CGV	AA-	A+	2016-06-15	2016-04-07	NICE신용평가는 CJ CGV의 수시평가 를 실시해 장·단기신용등급을 하향검 토 등급감시 대상(Credit Watch)에 등 재했다고 7일 밝혔다.

기업	이전 신용 등급	바뀐 신용 등급	신용등급 변화 날짜	뉴스 발표 날짜	핵심 뉴스 내용
한진해운	BB+	CCC	2016-06-17	2016-05-25	한진해운의 경우에도 Credit Event 발생 가능성을 감안한 수준(B- 부정적검토)으로 등급조정을 완료했기 때문이다.
한미약품	A	A+	2016-05-09	2015-12-25	먼저 한미약품은 대규모 기술수출로 수익성과 재무안전성이 개선될 것으로 평가되면서 현재 신용등급(A)이 상향 검토될 예정이다.
페이퍼코리아	B-	B	2018-01-25	2017-12-14	전주페이퍼·페이퍼코리아 등 제지회사는 △신문용지 수요의 감소 △원재료가 상승 등에 따른 경영난을 이유로 각각 15일과 20일부터 신문용지 가격을 통당 10%씩 인상하겠다는 방침을 신문사들에 통보했다
메타랩스	B	B-	2018-04-20	2018-04-19	지난 1월 한국기업평가가 이 회사 신용등급(B)을 하향 검토 대상에 올린 데 이어 이달 13일 나이스신용평가가 등급 전망을 '안정적'에서 '부정적'으로 변경했다.
두산	A-	BBB+	2019-05-10	2019-02-14	두산그룹 주요 계열사 신용등급 하향 검토...두산건설 대규모 손실 여파
삼화페인트	A-	BBB+	2019-05-21	2019-04-08	특히 삼화페인트는 하향 트리거 대부분을 충족하며, 신용사 3곳 가운데 2곳으로부터 부정적 아웃룩을 부여 받았다.
롯데정밀화학	A+	AA-	2019-04-11	2019-01-28	롯데정밀화학, 펀더멘털 개선...AA급 진입 청신호
GS건설	A-	A	2019-04-29	2018-12-27	우선 GS건설은 올해 신용등급이 A-로 유지됐고 등급 전망이 '안정적'에서 '긍정적'으로 상향됐다
웅진	BBB+	BBB-	2019-04-01	2019-03-28	지난달 27일 'BBB(하향검토)'로 떨어진 웅진의 신용등급은 더 내려갈 수 있다.
롯데쇼핑	AA+	AA	2019-05-03	2019-04-01	롯데쇼핑이 지난해 대규모 당기순손실을 기록한 가운데 차입금과 부채비율이 증가하는 등 재무지표 마저 전반적으로 악화되고 있다. 올해도 유의미한 실적 개선이 쉽지 않다는 전망이 이어지면서, 'AA+' 신용등급도 반납 위기에 처했다.
두산중공업	BBB+	BBB	2019-02-25	2018-12-26	지난 19일 한국신용평가는 두산중공업의 신용등급을 BBB+안정적에서 BBB+부정적으로 내렸다.
신세계건설	A-	A	2019-06-12	2019-05-14	신세계건설은 앞서 신세계백화점, 스타필드 등 유통, 상업시설을 중심으로 주거, 레저 등 다양한 곳에서 개발사업을 수행하며 시공 우수성을 보여 NICE(나

기업	이전 신용 등급	바뀐 신용 등급	신용등급 변화 날짜	뉴스 발표 날짜	핵심 뉴스 내용
롯데제과	AA+	AA	2019-05-14	2019-04-05	이스)신용평가에서 SK건설, 대우건설, KCC건설 등과 같은 신용등급인 A-(안정적)로 평가 받은 바 있다. 3월말 기준 신용등급 전망 ‘부정적’ 기업은 현대카드, 현대캐피탈, 기아자동차, 현대자동차, 롯데카드, 롯데쇼핑, 롯데제과, 롯데칠성음료, LG디스플레이, LG하우시스, 두산, 두산중공업, 두산건설, 한국항공우주산업, 지화이브, 버추얼텍,신원, 금호HT, 에이스테크놀로지, 현진소재, 삼화페인트공업, 페이퍼코리아, DGB생명보험, JT캐피탈 등이다.
두산	BBB+	BBB	2020-06-15	2020-04-09	한신평 “두산그룹, 강도 높은 자구책 필요… 주요 계열사 신용등급 하향 가능성 높아”
금호전기	B	B-	2020-05-12	2020-02-23	신용평가사들은 금호전기가 자산 매각 및 지분정리에도 재무부담이 과중한 것으로 평가한다. 미흡한 영업 수익성과 높은 경쟁 강도 등을 고려하면 내부현금창출력에 의한 재무안정성 개선 가능성은 제한적이며 과중한 재무 부담이 지속할 것이란 전망이다.
홍아해운	B-	CCC	2020-03-11	2019-12-26	이달 초 한국신용평가는 사업기반 약화와 유동성 위험 확대를 근거로 홍아해운 무보증사채 신용등급을 ‘B+(안정적)’에서 ‘B(부정적)’으로 낮췄다.
OCI	A+	A	2020-02-28	2020-02-05	수익성 부진 탈출구 안 보이는 OCI, 커지는 신용도 위험
계룡건설	BBB	BBB+	2020-05-15	2020-04-07	계룡건설산업은 BBB 등급을 유지하면서 지난해 긍정적 등급전망이 달렸다
두산중공업	BBB	BBB-	2020-06-15	2020-03-20	신용평가사들은 두산중공업의 신용등급을 ‘BBB(부정적)’로 하향 조정했다.
엔씨소프트	AA-	AA	2020-06-10	2020-05-19	역대급 실적을 토대로 국내 신용평가사의 등급상향 요건을 훌쩍 넘어서고 있다.
CJ CGV	A+	A	2020-05-22	2020-05-10	CJ CGV가 지속되는 재무구조 약화로 인해 신용등급 관리에 비상이 걸렸다.
한미약품	A+	A	2020-05-27	2020-01-17	올해 한미약품의 신용등급(A+, 부정적) 사수가 위태롭다.
이마트	AA+	AA	2020-02-12	2019-12-03	현재 한기평이 부여한 이마트의 신용등급은 AA+, 등급전망은 부정적이다.

2. 뉴스텍스트 감성 지수 제안

4.1절에서는 수치 데이터 기반 머신러닝 방법에 의해 오분류된 관측치를 정분류할 수 있도록 도와주는 텍스트 뉴스가 존재함에 대해 확인하였다. 본 절에서는 이러한 이해를 바탕으로 기계적으로 기업의 신용등급과 연관한 뉴스 텍스트를 수집하고 자연어 처리 기법을 활용해 신용등급 예측에 도움이 되는 새로운 수치 변수를 만드는 방법을 제안하고 그 유용성과 한계에 대해 검토하고자 한다. 이를 위해 먼저 본 연구에서 각 관측치의 6개월 후 신용등급이 실질적으로 결정된 날짜를 확정하는 데 사용한 알고리즘을 설명하고자 한다.

1) 신용등급 결정 날짜 알고리즘

본 연구에서 회계 결산시점으로부터 6개월 뒤의 신용등급을 예측하는 데 사용하고자는 텍스트는 각 기업의 신용등급과 관련하여 인터넷상에서 수집할 수 있는 다양한 뉴스이다. 회계결산 시점으로부터 6개월 뒤 신용등급이 확정되기 전까지 인터넷상의 각 기업의 신용등급과 연관한 뉴스들의 감성(긍정, 부정, 중립)을 의미 있게 분석해낸다면 신용등급의 변화를 예측하는데 좋은 정보를 추출할 수 있다고 판단하였다.

각 기업의 채권의 신용등급은 분기별로 평가되고 발표되는 것이 아니라 각 기업의 채권에 대한 신용평가 기관의 분석 결과를 바탕으로 비정기적으로 발표된다. 따라서 우리가 예측하고자 하는 6개월 후 신용등급이 결정되는 시점이 자료마다 다를 수 있다. 예를 들어 12월 결산기업인 한진해운의 2013년 12월 신용등급이 BBB+였는데 2014년 4월 17에 BBB-로 변화되어 2014년 6월 신용등급이 BBB-가 되었다면 2014년 6월의 신용등급은 실질적으로 2014년 4월 17일에 결정된 것이라고 볼 수 있다. 본 연구에서는 결산 시점으로부터 6개월 후 신용등급이 발표되기 전에 각 기업의 신용등급 변화를 예측할 수 있는 텍스트를 찾되 실질적인 등급 결정 날짜 이전으로 텍스트 검색 범위를 줄여 예측 용도로 사용할 수 있는 텍스트 정보의 추출 가능성을 확인하고자 했다. 각 기업별 신용등급 결정 날짜는 다음에 기술되는 방법을 따라 설정하였다.

각 기업 자료별 신용등급 결정 날짜를 확정하기 위해 두 가지를 고려하였다. 첫 번째는 세 개의 신용평가회사(한국신용평가, 한국기업평가, NICE평가정보)가 평가한 신용등급을 종합한 것을 추정 대상 신용등급으로 선정하였다. 두 번째는 예측 용도로 사용할 수 있는 텍스트 정보 추출 가능성을 엄격하게 확인하고자 신용등급 결정 날짜의 후보가 여러 개인 경우 가장 빠른 날짜를 선택하였다. 이러한 점들을 고려하여 신용등급 결정 날짜를 아래에

주어진 세 가지 case로 나누어 결정하였다.

Case 1: 결산 시점 이후 6개월 동안 신용등급이 평가되지 않았다면 결산월로부터 6개월 후 마지막 날로 설정하였다. 예를 들어 12월 결산 기업의 경우 6월 30일로 설정하였다.

Case 2: 결산 시점 이후 6개월 동안 같은 등급으로 평가가 지속되었고 그것이 6개월 후 최종 신용등급이 되었다면 그 등급이 첫 번째로 평가된 날짜로 설정하였다.

예를 들어 사례 1의 한진해운은 12월 결산 기업이고 2013년 12월 결산 당시 신용등급은 BBB+이다. 그 후 아래 <사례 1>과 같은 등급 평가를 거쳐 2014년 6월의 신용등급이 BBB-로 평가되었는데 이러한 경우, 2014년 한진해운의 6월 신용등급 결정 날짜는 2014-03-13으로 설정된다.

<사례 1>

발표 날짜	발행주체	종합	한신평	한신정	한기평
2014-03-13	한진해운	BBB-	BBB+	BBB+	BBB-
2014-03-20	한진해운	BBB-	BBB	BBB+	BBB-
2014-04-17	한진해운	BBB-	BBB	BBB	BBB-

Case 3: 결산 시점 이후 6개월 동안 등급 변화가 여러 번 있는 경우, 관측치를 신용등급 발표 날짜와 등급을 기준으로 군집(cluster)을 만들어 신용등급 결정 날짜를 아래와 같은 과정으로 결정하였다.

3-1: 신용등급 발표 날짜 차이가 한 달 이내 이거나 발표된 신용등급이 동일하면 한 군집으로 묶는다.

3-2: 한 개의 군집인 경우는 그 군집 내에서 신용등급 첫 번째 결정날짜를 신용등급 발표 날짜로 설정한다.

3-3: 두 개 이상의 군집일 때 시간상으로 가장 마지막에 있는 군집에 속한 신용등급 발표 날짜에서 처음 발표된 날짜를 신용등급 결정날짜로 설정한다.

예를 들어 아래 <사례 2>의 12월 결산 기업 한진해운의 2016년 6월 신용등급 결정 날짜는 군집 2의 첫 번째 신용등급 발표 날짜인 2016-06-17로 설정된다.

<사례 2>

발표 날짜	발행주체	종합	한신평	한신정	한기평	군집
2016-03-22	한진해운	BB	BB+	BB	BB+	1
2016-03-30	한진해운	BB-	BB	BB	BB-	1
2016-04-22	한진해운	B-	B-	B-	B-	1
2016-06-17	한진해운	CCC	CCC	B-	B-	2
2016-06-20	한진해운	CCC	CCC	CCC	CCC	2

2) 텍스트 데이터 수집 및 전처리

본 연구에서는 앞부분에서 기업의 신용등급과 관련된 구글 검색 결과에서 유용한 뉴스를 발견한 것에 착안하여 각 기업 자료에 대한 재무제표 결산 시점부터 신용등급 결정날짜 전까지 기간의 신용등급 관련 구글 검색 결과 첫 페이지를 활용하여 뉴스 관련 변수를 만들고자 했다. 구글 검색 결과를 수집하기 위해 동적 웹 크롤링(Web Crawling) 툴인 셀레니움(Selenium)을 이용하여 자료를 구축하였다. 4.1절에서와 같이 기간과 키워드를 설정하여 얻은 구글 검색 첫 페이지의 HTML 자바스크립트(Java Script)에서 각 뉴스의 뉴스 제목, 뉴스 요약, 날짜를 추출했다([그림 2] 참조). 연속적으로 뉴스 검색 결과를 스크래핑하면 구글로부터 차단당할 수 있어서 이를 방지하기 위해 각 페이지에 접속할 때마다 1부터 10까지 난수를 뽑아 난수만큼의 초 간격을 두고 접속하여 스크래핑하였다.

위 방법으로 구글 검색을 하면 해당 기업의 신용등급과 관련된 뉴스가 나오는 경우도 있지만, 회사채 신용등급 예측과는 관계없는 뉴스들도 검색 결과에 포함되는 경우도 존재했다. 예를 들어 회사채 신용등급이 아닌 해당 기업 제품 소비자들의 신용등급이 언급된 뉴스, 모기업의 신용등급 검색 시 계열사의 신용등급이 언급된 뉴스, 기업들의 신용등급이 나열된 뉴스 등이 있었다. 이러한 뉴스들을 필터링하기 위해 세 가지 전처리를 했다. 첫 번째는 모기업과 관련된 뉴스의 추출과 연관하여 계열사 관련 뉴스를 필터링하기 위해 기업명 뒤에 조사와 명사를 붙여 해결하였다. 예를 들어 두산그룹과 같은 대기업들은 두산중공업, 두산건설, 두산인프라코어 등 여러 계열사가 존재한다. 모기업인 두산그룹의 신용등급과 관련된 뉴스를 수집하기 위해 ‘두산의’, ‘두산은(는/이/가)’, ‘두산그룹’, ‘두산 그룹’, ‘두산’처럼 언급된 뉴스 제목 또는 뉴스 요약들만 가져왔다. 두 번째로는 ‘신용등급’, ‘신용’, ‘등급’과 같은 단어들이 언급된 뉴스 제목 또는 뉴스 요약만 가져왔다. 세 번째로 각 뉴스 안에 함께 제시되는 뉴스 요약의 여러 문장 중 완전한 문장의 형태로 주어진 요약만을 분석 대상으로 했다.

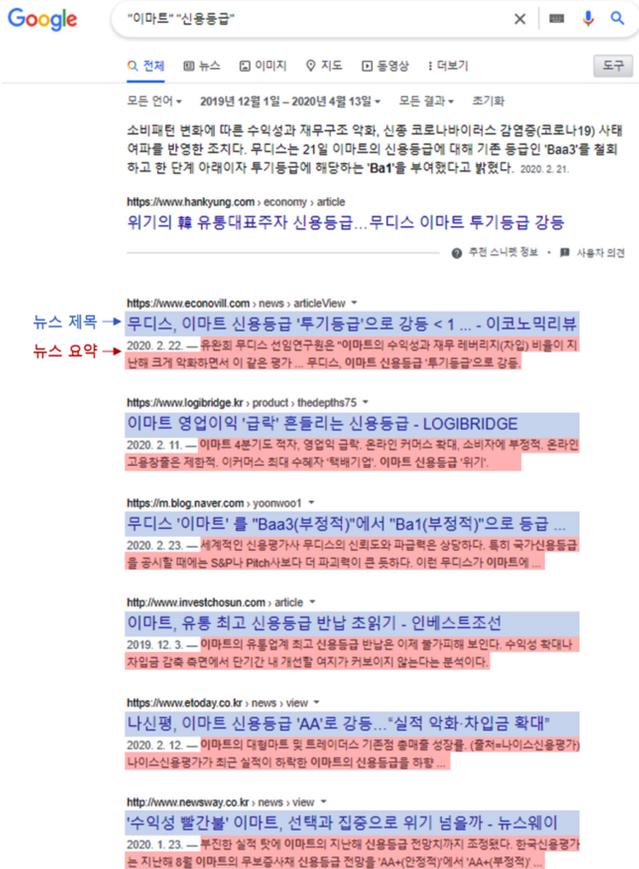
[그림 2]의 세 번째 뉴스의 뉴스 요약은 ‘세계적인 신용평가사 무디스의 신뢰도와 파급력은

상당하다. 특히 국가신용등급을 공시할 때는 S&P나 Pitch사보다 더 파괴력이 큰 듯하다. 이런 무디스가 이마트에...'로 주어지 있는데, 본 연구에서는 '세계적인 신용평가사 무디스의 신뢰도와 파괴력은 상당하다.', '특히 국가신용등급을 공시할 때는 S&P나 Pitch사보다 더 파괴력이 큰 듯하다.'와 같은 완전한 문장만을 뉴스 요약으로 가져와서 연구에 이용하였다.

앞의 과정을 거쳐 추출된 각 뉴스의 뉴스 제목, 뉴스 요약들에 대해 아래의 추가적인 전처리 과정을 수행하였다.

- (1) '[]', '()', '< >' 괄호 안 텍스트와 특수문자를 제거했다.
- (2) 영어가 있는 경우 대문자를 모두 소문자로 변경했다.
- (3) 다중 공백일 때 한 개의 공백으로 변경했다.

[그림 2] 한 기업 자료에 대한 구글 검색 결과 예시.



3) 뉴스 텍스트를 이용한 수치 변수 생성

이번 절에서는 전처리한 텍스트 데이터로부터 신용등급 예측에 활용하기 위해서는 수치 데이터를 생성해 내는 과정에 대해 설명하고자 한다. 이를 위해 사용되는 자연어 처리 모형인 BERT와 텍스트 마이닝 기법인 감성분석에 대해 설명하고자 한다.

2018년 구글이 개발한 트랜스포머(Transformer) 기반 언어모형 버트(BERT; Bidirectional Encoder Representations from Transformers)는 문장의 양방향으로 단어들을 참고하여 문장의 문맥을 반영한 단어의 수치적인 벡터 표현인 임베딩(Embedding)을 얻어내는 자연어 처리 방법론이다. BERT와 같은 사전학습(Pretrained) 언어모형은 레이블이 없는 대용량 텍스트 데이터로부터 언어 구조 및 문법을 학습한 모형으로 여러 기업 또는 연구소에서 사전 훈련한 결과를 공유하고 있다. 이러한 사전 훈련된 모형을 가지고 하이퍼파라미터 재조정 및 레이블이 있는 데이터로 추가 훈련을 진행하면 여러 자연어 관련 task에 우수한 성능을 보여줄 수 있다는 사실이 잘 알려져 있으며, 이 과정을 파인 튜닝(Fine-tuning)이라고 부른다.

감성분석(Sentiment Analysis)은 텍스트에 들어있는 의견, 감정, 평가, 태도 등의 주관적인 정보를 긍정, 중립, 부정으로 분류하는 분석 방법이다. 본 연구는 서울대학교 언어학과 연구실에서 개발한 금융 도메인 기반 한국어 감성분석 버트 모형인 KR-FinBert-SC (<https://huggingface.co/snunlp/KR-FinBert-SC>)를 이용하였다. KR-FinBert-SC는 약 44만 개의 뉴스 제목과 본문, 11만 개의 애널리스트 리포트를 이용해 추가학습을 한 후 5만 개의 감성 분류 레이블 데이터로 파인 튜닝한 모형이다. KR-FinBert-SC 모형은 주어진 금융 기반 텍스트가 각 감성(긍정, 중립, 부정)으로 분류될 확률을 반환한다.

본 연구는 KR-FinBert-SC 모형을 이용하여 뉴스 기반 수치 변수를 만들었으며 방법은 다음과 같다. 가장 높은 확률을 가진 감성이 각 뉴스의 감성(sentiment)이 되며 부정은 -1, 긍정은 1, 중립은 0으로 변환하여 처리했다. 뉴스 t 의 각 감성의 확률을 $P_{neg}(t)$, $P_{pos}(t)$, $P_{neu}(t)$ 일 때 감성 점수(sentiment score) $S(t)$ 는

$$sentiment(t) \times \max(P_{neg}(t), P_{pos}(t), P_{neu}(t))$$

로 계산한다. 만약 어떤 기업 자료의 구글 검색 첫 번째에 n 개의 신용등급 관련 뉴스 제목 $t_1, t_2, t_3, \dots, t_n$ 이 존재한다면 $\sum_{i=1}^n S(t_i)$ 를 그 기업 자료에 대한 뉴스지수(news index)로 정의한다. 검색 결과상에 표시된 뉴스이지만 해당 뉴스의 제목이 전처리되어 분석 대상이 안 되는 경우라도 해당 뉴스의 요약이 완전한 문장의 형태로 전처리를 통과한 경우, 그

요약 문장의 감성 점수를 뉴스의 감성 점수로 삼았다. 또 그러한 요약 문장이 두 개 이상 존재하는 경우는 각 요약 문장의 감성 점수의 평균을 뉴스의 감성 점수로 삼았다.

<표 6>은 기업과 관련된 뉴스가 있어 뉴스지수가 만들어진 경우에 대한 요약통계를 보여준다. 표에서 알 수 있듯이 과거에는 기업의 신용 상황에 대한 인터넷상의 뉴스가 별로 많지 않아 뉴스지수를 함께 고려할 수 있는 관측치의 수가 매우 제한적임을 알 수 있다. 하지만 최근 들어서는 뉴스지수를 함께 고려할 수 있는 관측치의 개수가 많아졌음을 알 수 있다. 뉴스지수를 계산할 수 있는 기업의 수가 상당히 증가한 시점인 2013년부터 2019년까지 계산된 뉴스지수는 상당히 음의 값에 치우쳐져 있다. 이는 기업의 신용등급에 관련된 기사가 날 경우에 대부분 신용등급 하락과 관련된 부정적인 내용이 포함되어 있을 가능성이 크다는 것을 의미하고 있다.

<표 6> 뉴스지수의 요약통계량

본 표는 뉴스지수의 연도별 요약통계량을 보고하고 있다. 구글을 통해 검색한 기업의 신용등급 관련 뉴스를 대상으로 하여, 서울대학교 언어학과 연구실에서 개발한 금융 도메인 기반 한국어 감성분석 버트 모형인 KR-FinBert-SC(<https://huggingface.co/snunlp/KR-FinBert-SC>)를 이용하여 뉴스지수를 계산하였다. KR-FinBert-SC 모형은 주어진 금융 기반 텍스트가 각 감성(긍정, 중립, 부정)으로 분류될 확률을 반환한다. 가장 높은 확률을 가진 감성이 각 뉴스의 감성이 되며 부정은 -1, 긍정은 1, 중립은 0으로 변환하여 처리했다.

연도	기업수	관측수	Mean	S.D.	Q1	Median	Q3
1998	0	0	-	-	-	-	-
1999	1	1	0.9893	-	0.9893	0.9893	0.9893
2000	1	1	0.9998	-	0.9998	0.9998	0.9998
2001	1	1	0.9997	-	0.9997	0.9997	0.9997
2002	0	0	-	-	-	-	-
2003	2	2	-0.0003	1.4135	-0.5000	-0.0003	0.4995
2004	1	1	0.9999	0.9999	0.9999	0.9999	0.9999
2005	0	0	-	-	-	-	-
2006	3	3	0.9282	0.1238	0.8925	0.9997	0.9997
2007	1	1	0.5752	-	0.5752	0.5752	0.5752
2008	4	4	-0.9947	0.0074	-0.9995	-0.9976	-0.9928
2009	6	6	-0.2602	0.8369	-0.9775	-0.4559	0.2893
2010	5	5	0.1775	0.8047	0.0000	0.0000	0.8829
2011	7	7	0.7140	0.4877	0.4994	0.9998	0.9998
2012	7	7	0.2335	0.7870	-0.1918	0.0000	0.9985
2013	15	22	-0.2955	1.2086	-0.9984	-0.9250	0.9964
2014	6	7	-0.1616	0.9755	-0.9838	-0.4913	0.7459
2015	18	22	-0.3413	0.7531	-0.9983	-0.5726	0.0000
2016	22	26	0.0333	0.8871	-0.9977	0.2874	0.9133
2017	17	22	-0.3273	0.8365	-0.9952	0.0000	0.0000
2018	35	51	-0.2531	1.1802	-0.9995	-0.9747	0.9969
2019	56	102	-0.3125	1.2952	-0.9997	-0.5660	0.9840
전체	208	291	-0.1606	1.0736	-0.9989	0.0000	0.9943

3. 뉴스지수를 포함한 신용등급 변경 예측

본 절에서는 4.2절에서 정의한 뉴스지수를 4.1절에서 사용했던 수치형 변수들과 함께 사용하였을 때 신용등급 변화 예측 결과를 살펴봄으로써 뉴스지수의 유용성을 검토하고자 한다. 뉴스지수가 있는 관측치는 총 291개(208개의 기업에 대한 관측치)로 다소 관측치의 개수가 적은 관계로 자료를 시간순으로 나열하여 시간순서상 앞부분 80%를 훈련데이터로 뒷부분 20%를 검증데이터로 하여 분석결과를 산출해 보았다. 예측모형으로는 4.1절과 동일하게 랜덤 포레스트(random forest) 모형을 사용하였으며 수치데이터만 활용하여 예측하였을 때와 수치데이터와 뉴스지수를 함께 사용하여 예측하였을 때의 결과를 비교하여 보았다.

<표 7>과 <표 8>은 수치데이터만 가지고 신용등급 변화를 예측하였을 때 혼동행렬과 클래스별 예측결과 값이고, <표 9>, <표 10>은 수치데이터와 뉴스지수를 함께 사용하여 신용등급 변화를 예측하였을 때의 혼동행렬과 클래스별 예측결과 값이다. 결과를 비교하여 보면 신용등급 유지인데 신용등급 하락으로 잘못 예측한 자료 하나를 추가적으로 정분류되었다. 정정된 자료 하나는 유지(0)인데 하락(-1)으로 오분류되었던 것이 뉴스지수를 추가하였을 때 유지(0)로 정분류되었다. 이로 인해 하락(-1) 표본에서는 Precision이, 유지(0) 표본에서는 Recall 값이 향상되어, 결과적으로 두 표본 모두에서 F1-score가 모두 상승하였다. 하지만, 전체적으로 보았을 때 이것 외에는 분류결과상의 큰 차이는 없다. 4.1절에서 확인된 인터넷상의 기업관련 뉴스가 가지는 신용등급 변화 예측과 연관된 정보력에 비해서는 다소 실망스런 결과이다.

이는 일차적으로 연구 표본수가 작은 것이 원인일 수 있다. 전체 표본수가 적는데 신용등급 상승(하락)과 관련한 긍정적인(부정적인) 뉴스 표본의 비중도 적었다면 뉴스분석을 통한 예측력이 낮은 것도 당연한 결과가 될 것이다. 이런 정량적인 분석의 한계를 고려하여 이후 부분은 구체적인 뉴스 사례를 개념적으로 분류하는 정성적인 분석도 추가하였다.

<표 7> 수치데이터 기반 신용등급 변경 예측모형에 대한 혼동 행렬

본 표는 뉴스지수가 존재하는 표본을 대상으로 신용등급의 실제 클래스 대비 예측된 신용등급 클래스의 혼동행렬을 보고한다. 신용등급 변경 -1은 신용등급 하락, 0은 유지, 1은 상승을 의미한다.

		실제 클래스		
		-1	0	1
예측된 클래스	-1	1	1	0
	0	6	33	1
	1	0	0	0

<표 8> 수치데이터 기반 신용등급 변경 예측모형의 클래스별 예측성과

본 표는 뉴스지수가 존재하는 표본을 대상으로 회계 및 재무변수 등의 수치데이터를 기반으로 신용등급 변화를 예측한 모형의 클래스별 성과를 보여주고 있다. 신용등급 변경 -1은 신용등급 하락, 0은 유지, 1은 상승을 의미한다. Precision, Recall, F1-score의 정의는 본문의 식 (1)을 참고하시오. 예측모형에서 클래스 1(상승)로 예측한 관측치가 없어 클래스 1에 대한 Precision, F1-score는 제시하지 않았다.

	-1	0	1	평균값
Precision	0.5000	0.8250	NaN	-
Recall	0.1429	0.9706	0.0000	0.3712
F1-score	0.2222	0.8919	NaN	-

<표 9> 수치데이터와 뉴스지수를 함께 사용하여 신용등급 변경을 예측했을 때 혼동 행렬

본 표는 뉴스지수가 존재하는 표본을 대상으로 회계 및 재무변수와 뉴스지수를 함께 사용하여 신용등급의 실제 클래스 대비 예측된 신용등급 클래스의 혼동행렬을 보고한다. 신용등급 변경 -1은 신용등급 하락, 0은 유지, 1은 상승을 의미한다.

		실제 클래스			
		-1	0	1	
예측된 클래스	-1	1	0	0	
	0	6	34	1	
	1	0	0	0	

<표 10> 수치데이터와 뉴스지수를 함께 사용하여 신용등급 변경을 예측했을 때 클래스별 예측결과

본 표는 뉴스지수가 존재하는 표본을 대상으로 회계 및 재무변수와 뉴스지수를 함께 사용하여 신용등급 변화를 예측한 모형의 클래스별 성과를 보여주고 있다. 신용등급 변경 -1은 신용등급 하락, 0은 유지, 1은 상승을 의미한다. Precision, Recall, F1-score의 정의는 본문의 식 (1)을 참고하시오. 예측모형에서 클래스 1(상승)로 예측한 관측치가 없어 클래스 1에 대한 Precision, F1-score는 제시하지 않았다.

	-1	0	1	평균값
Precision	1.0000	0.8293	NaN	-
Recall	0.1428	1.0000	0.0000	0.3809
F1-score	0.2500	0.9066	NaN	-

다음 정성적 분석 부분에서는 인터넷상 기업관련 뉴스를 기반으로 뉴스지수를 만들었을 때 기대했던 바와 달리 신용등급 변화 예측력의 유의미한 성과개선으로 이어지지 못하는 이유를 살펴보고, 추후 신용등급 변화 예측에 유용한 뉴스지수 계산을 위해 고려할 점이 무엇인지 생각해보고자 한다.

4. 제안한 뉴스지수가 신용등급 변경 예측에 가지는 한계점 분석

제안한 뉴스지수가 존재하는 관측치의 개수는 총 208개로 양수 값을 갖는 관측치는 80개,

음수 값을 갖는 관측치는 99개이다. 뉴스지수가 양수인 관측치 중에 실제 클래스가 상승(1)인 관측치는 총 4개 존재하였으며 이들은 모두 랜덤 포레스트 모형에 의해 유지(0)로 오분류된 관측치이다(<표 11> 참조). 이들 관측치는 수치 데이터만을 이용하여 구축한 머신러닝 모형은 오분류했지만 뉴스지수에 의해 정분류될 수 있는 관측치들이다.

<표 11> 뉴스지수가 양수인 표본 중 신용등급이 상승한 관측치

본 표는 뉴스지수가 양수인 표본 중 신용등급이 실제 상승한 관측치를 보고하고 있다. 이들 4개 표본은 모두 랜덤 포레스트 모형에 의해 유지(0)로 오분류된 관측치들이다. 즉, 예측값은 0이지만, 실제값이 1로 신용등급이 상승하였다. 뉴스지수는 구글을 통해 검색한 기업의 신용등급 관련 뉴스를 대상으로 하여, 서울대학교 언어학과 연구실에서 개발한 금융 도메인 기반 한국어 감성분석 버트 모형인 KR-FinBert-SC(<https://huggingface.co/snunlp/KR-FinBert-SC>)를 이용하여 계산하였다. KR-FinBert-SC 모형은 주어진 금융 기반 텍스트가 각 감성(긍정, 중립, 부정)으로 분류될 확률을 반환한다. 가장 높은 확률을 가진 감성이 각 뉴스의 감성이 되며 부정은 -1, 긍정은 1, 중립은 0으로 변환하여 처리했다.

연도	월	기업명	이전 신용 등급	바뀐 신용 등급	예측값	실제값	뉴스지수
2019	6	롯데정밀화학	A+	AA-	0	1	0.9994
2013	6	현대차	AA+	AAA	0	1	0.9998
2007	6	S-Oil	AA	AA+	0	1	0.7852
2005	6	현대상선	BBB	BBB+	0	1	0.9998

표본에서 뉴스지수가 양수인 나머지 76개중 73개의 관측치가 실제 신용등급을 유지한 관측치, 3개는 하락한 관측치이다. 실제 하락한 관측치 3개에 관련된 정보는 <표 12>에 정리했다. 이마트 관련하여 수집된 뉴스는 총 3개로 모두 신용등급 하락 가능성을 내포하고 있는 부정적 뉴스이다. 하지만 사용하는 감성분석 모형에 의해 첫 번째 뉴스는 ‘중립’, 두 번째 뉴스는 ‘긍정적’, 세 번째 뉴스는 ‘부정적’으로 분류되었다. 세 개 뉴스의 감성 점수가 합쳐지면서 뉴스지수가 양수인 0.0012가 되었다. 롯데쇼핑과 현대상선 관련 뉴스는 신용등급과 연관하여 어떤 성격의 뉴스인지 알기 힘들지만, 감성분석 모형은 긍정적 뉴스로 분류된 경우이다. 이러한 상황들은 현재 감성분석 모형이 금융 텍스트의 감성을 파악하는데 가지는 한계로 보인다. 이러한 한계는 추후 감성분석 모형에 대한 추가학습이 이루어지면 개선될 수 있는 여지가 있다고 보인다.

표본 중 뉴스지수가 음수인 관측치 중에 실제 신용등급이 하락한 관측치는 18개이고, 이 중 <표 13>에 제시된 12개의 관측치는 수치 데이터 기반 랜덤 포레스트 모형에 의해 유지(0)로 오분류된 관측치이고 나머지는 정분류되었다. 이들 <표 13>의 12개의 관측치는 뉴스지수를 활용하여 정분류될 수 있는 가능성이 있는 관측치이다. 나머지 뉴스지수가 음수인 81개의 관측치는 모두 신용등급이 유지된 관측치였다.

<표 12> 뉴스지수가 양수인 표본 중 신용등급이 하락한 관측치

본 표는 뉴스지수가 양수인 표본 중 신용등급이 실제로는 하락한 관측치를 보고하고 있다. 뉴스지수는 구글을 통해 검색한 기업의 신용등급 관련 뉴스를 대상으로 하여, 서울대학교 언어학과 연구실에서 개발한 금융 도메인 기반 한국어 감성분석 버트 모형인 KR-FinBert-SC(<https://huggingface.co/snunlp/KR-FinBert-SC>)를 이용하여 계산하였다. KR-FinBert-SC 모형은 주어진 금융 기반 텍스트가 각 감성(긍정, 중립, 부정)으로 분류될 확률을 반환한다. 가장 높은 확률을 가진 감성이 각 뉴스의 감성이 되며 부정은 -1, 긍정은 1, 중립은 0으로 변환하여 처리했다.

연도	월	기업명	이전 신용 등급	바뀐 신용 등급	예측 값	실제 값	뉴스 텍스트	감성 점수	뉴스지 수
2020	6	이마트	AA+	AA	0	-1	이마트24의 지분 100%를 보유하고 있는 모회사 이마트의 경우 회사채 신용등급은 AA+, 아웃룩은 '부정적'이다.	0	0.0012
							이마트(139480)와 LG디스플레이(034220) 신용등급이 또 다시 강등됐다. 한국기업평가는 지난해 5월 이마트의 신용등급 전망을 기존 '안정적'에서 '부정적'으로 끌어내린 데 이어 연말에는 추가 강등 가능성을 시사하기도 했다.	0.9997	
2019	6	롯데쇼핑	AA+	AA	0	-1	롯데쇼핑, 국내·외 신용등급 평가간극 8계단차로 '최대' - CEO ...	-0.9984	0.4716
2014	6	현대상선	BBB+	BB+	-1	-1	특히 한진해운과 현대상선의 경우 재무개선작업의 이행과 성과 등이 신용등급의 방향에 직접적인 영향을 미칠 전망이다.	0.9991	0.9991

<표 13> 수치데이터 기반 예측모형에서 오분류된 관측치 중 뉴스지수가 음수이고 신용등급이 하락한 관측치

본 표는 수치데이터 기반 예측모형에서 오분류된 관측치 중 뉴스지수가 음수이고 신용등급이 실제로는 하락한 관측치를 보고하고 있다. 뉴스지수는 구글을 통해 검색한 기업의 신용등급 관련 뉴스를 대상으로 하여, 서울대학교 언어학과 연구실에서 개발한 금융 도메인 기반 한국어 감성분석 버트 모형인 KR-FinBert-SC(<https://huggingface.co/snunlp/KR-FinBert-SC>)를 이용하여 계산하였다. KR-FinBert-SC 모형은 주어진 금융 기반 텍스트가 각 감성(긍정, 중립, 부정)으로 분류될 확률을 반환한다. 가장 높은 확률을 가진 감성이 각 뉴스의 감성이 되며 부정은 -1, 긍정은 1, 중립은 0으로 변환하여 처리했다.

연도	월	기업명	이전 신용 등급	바뀐 신용 등급	예측값	실제값	뉴스지수
2020	6	두산중공업	BBB	BBB-	0	-1	-2.9669
2020	6	한미약품	A+	A	0	-1	-0.9997
2020	6	OCI	A+	A	0	-1	-0.9989
2020	6	CJ CGV	A+	A	0	-1	-0.8644
2019	6	두산	A-	BBB+	0	-1	-0.9997
2019	6	삼화페인트	A-	BBB+	0	-1	-0.9963
2019	6	롯데지주	AA+	AA	0	-1	-0.9930
2018	6	대한해운	BBB	BBB-	0	-1	-1.9184
2016	6	LS	AA-	A+	0	-1	-0.9991
2016	6	CJ CGV	AA-	A+	0	-1	-0.9985
2013	6	STX	BBB+	B-	0	-1	-0.9785
2013	6	동양	BB	BB-	0	-1	-0.3836

이상의 논의에서 뉴스지수 활용을 통해 상승으로 정분류될 수 있는 관측치가 4개, 하락으로 정분류될 수 있는 관측치가 12개 나타남을 알 수 있다. 이러한 관측치 개수는 수치 데이터 기반 머신러닝 모형이 상승과 하락으로 정분류한 관측치들의 개수에 대한 상대적 비율(상승: 14.8%, 하락: 32.4%)을 고려할 때 적지 않은 비율이다. 하지만, 이러한 유용성은 뉴스지수의 값이 양수나 음수인 관측치 중에 실제 클래스가 유지(0)인 관측치가 다수 존재한다는 사실로 인해(양수 중에 73개, 음수 중에 81개) 의미가 없게 되는 측면이 있다.

본 절에서는 실제 클래스가 유지된 관측치들(실제값=0)에 대해 뉴스 제목이나 뉴스 요약의 내용을 통해 계산된 뉴스지수가 가지는 한계의 원인을 크게 총 여섯 개의 그룹으로 분류하여 파악하고자 했다. 회사채 신용등급이 유지된 관측치들에 대한 뉴스 제목이나 뉴스 요약의 내용에 따른 분류 전체 결과를 정리한 표는 Lee(2023)의 부록에서 확인할 수 있다. 우리의 연구에서 신용등급 변화 예측을 위해 구글링하여 얻어진 뉴스 텍스트는 ‘회사채 신용등급의 문맥이 아니거나 알기 힘든 경우’, ‘계열사에 대한 뉴스’, ‘신용등급 유지한다는 뉴스’, ‘텍스트 정보가 미래에 반영’, ‘텍스트에 있는 정보가 실제로 일어나지 않음’, ‘회사채 신용등급과 무관한 뉴스’로 총 여섯 그룹으로 분류할 수 있다.

(1) 신용등급을 유지한다는 뉴스

기업의 회사채 신용등급이 변동 없이 유지된다는 뉴스가 20개 존재했으며, 이는 실제 유지된 관측치의 약 11%를 차지한다. 해당 뉴스는 “렌탈 전문기업 AJ네트웍스는 한국신용평가의 정기평가를 통해 장기신용등급 전망이 BBB+ ‘부정적’에서 ‘안정적’으로 상향됐다고 2일 밝혔다.”, “OCI, 회사채 등급 ‘AA-/안정적’유지”와 같이 신용등급이 ‘안정적’, ‘유지’와 같은 키워드를 가진 뉴스들이 존재했다. 실제 해당 내용의 뉴스를 가진 관측치들 모두 6개월 동안 신용등급을 유지했으며, 머신러닝 기반 예측에서도 유지됐다. 하지만, 신용등급이 유지된다는 뉴스의 감성이 긍정 또는 부정으로 분류됐을 때의 상황이 존재하여 해당 문제에 대한 논의가 필요하다.

(2) 뉴스 정보가 시차를 두고 미래에 반영되는 경우

기업의 신용등급 변화를 내포하는 뉴스 텍스트가 실제 신용등급 변동에 반영되는 시간이 6개월이 아닌 1년 이상 걸린 경우도 존재했다. 관련된 뉴스들은 총 16개 존재했으며, 이는 실제 유지된 관측치의 약 8%이다. 예를 들어, SK하이닉스 같은 경우, ‘SK하이닉스는 설립 후 처음으로 투자적격 지위를 얻었다.’라는 뉴스가 나오고 나서 약 1년 뒤 신용등급이 상승했다. 동국제강 같은 경우, ‘신용등급이 A0 등급으로 떨어진 상태이고, 회사채 만기가 도래하고

있기 때문이다.'라는 내용의 뉴스 발표가 나고 약 1년 뒤 신용등급이 하락했다. 또한 머신러닝 기반 모형에서도 하락을 예측한 것으로 보아, 재무제표와 텍스트 지표는 하락을 의미해도 반영되는 시차가 존재한다는 것을 알 수 있다.

(3) 회사채 신용등급과 관련된 문맥이 아닌 경우

뉴스를 크롤링하기 위해 '기업명'과 '신용등급' 두 개의 키워드를 이용하였으나, 해당 기업의 회사채 신용등급과 관련된 내용의 뉴스가 수집되지 않은 경우가 존재했다. 관련된 뉴스들은 총 15개 존재했으며, 이는 전체 유지된 관측치의 약 7%이다. 1) 그룹명의 동음이의어가 키워드로 들어간 뉴스, 2) 회사채 신용등급이 아닌 다른 의미인 '신용등급'이 키워드로 들어간 뉴스가 있다. 예를 들어, 대상 기업에서 수집된 뉴스 텍스트 중 하나는 '신용평가 대상은 일반적으로 채무자(Issuer)와 개별 채무(Issue)로 나누어 볼 수 있다.'와 같이 대상 기업이 아닌 동음이의어인 '대상(對象)'의 키워드로 들어갔다. 회사채 신용등급과 무관한 다른 의미인 '신용등급'이 키워드로 들어간 뉴스들은 "삼성엔지니어링의 안전관리직무에 지원하려면 오픈 II 등급과 토익스피킹 5등급을 충족하면 된다.", "금호타이어, EU 타이어 라벨링 최고 등급 타이어 개발"과 같다. 이를 통해 '기업명'과 '신용등급'을 키워드를 넣고, 기업명과 '신용', '등급'이 정확하게 언급된 뉴스 텍스트를 가져오더라도 회사채 신용등급 변동과 관련된 뉴스를 선별하기가 어렵다는 것을 알 수 있다.

(4) 계열사에 대한 뉴스

기업 지주회사와 계열사의 경영 성과는 상호관계를 맺고 있어 신용등급에 간접적으로 영향을 줄 수 있다. 하지만 정확한 분석을 위해 계열사를 제외한 지주회사와 관련된 뉴스만 수집하고자 했으나 완전히 제외하는 것은 기술적 한계가 있었다. 해당 뉴스는 전체 유지된 관측치의 약 3%이며, 총 6개이다.

(5) 신용등급에 대한 사실을 단순히 기술하는 뉴스

회사채 신용등급과 관련된 사실만 담았고 향후 신용등급이 어떻게 변할지 예측하기 어려운 뉴스도 존재했다. 관련 뉴스는 전체 유지된 관측치의 약 55%로 크롤링된 절반에 가까운 뉴스가 각 기업의 신용등급에 대한 사실을 단순히 기술하는 뉴스였다. 1) 과거 기업의 신용등급과 관련된 사실을 적시한 뉴스, 2) 국제 신용평가사와 관련된 뉴스, 3) 신용등급과 관련된 기업의 계획을 포함한 뉴스, 4) 현재 신용등급을 언급한 뉴스, 5) 신용등급에 대한 뉴스이지만 등급변화의 방향에 대한 정보가 없는 뉴스가 있다. 예를 들면, "LG화학은

AA+(긍정적) 등급까지 획득하며 AAA 등급 획득 눈앞에까지 갔다가 좌절된 바 있다.”와 같이 LG화학의 과거 신용등급과 관련된 사실을 언급하지만 현재 신용등급의 가망은 알기 힘든 뉴스이다. 국제 신용평가사가 평가한 뉴스는 “글로벌 신용평가기관인 무디스(Moody’s)는 KT의 신용등급을 Baal ‘Positive(긍정적)’에서 A3 ‘Stable(안정적)’로 상향 조정했다고 24일 밝혔다.”와 같다. 그러나 국제신용평가사와 국내신용평가사의 신용등급 체계와 평가 기준이 다르기 때문에 국제신용평가사의 신용등급이 국내 신용등급에 영향을 미친다고 보기 어렵다. 또한 기업의 신용등급 상향을 위한 계획과 현재 신용등급을 언급한 뉴스들의 예시는 다음과 같다. “신용등급 A까지 올리겠다”…아시아나항공, 재무구조 개선에 가속’과 ‘현재 국내 일반 비금융기업 중 AAA 등급을 보유하고 있는 곳은 SK텔레콤과 KT 등 통신사, 그리고 현대자동차까지 딱 3곳이다’. 위와 같은 뉴스는 신용등급과 관련된 사실이 포함되어 있었지만, 회사채 신용등급의 향후 변동성을 알기 어렵다. 신용등급에 대한 뉴스이지만 문맥을 읽기 힘든 뉴스들은 “E1의 신용등급을 감안하면 이례적인 일로 회자됐다.”, “신용평가업계는 두산인프라코어의 신용등급 변동 가능성에 주목하고 있다.”와 같이 신용등급 변동 가능성은 언급하지만 어떻게 변동될지에 대해 파악하기 힘든 뉴스들이 수집됐다. 해당 뉴스들을 제거할 수 있는 전처리 방법은 추후 연구에서 다루고자 한다.

(6) 텍스트에 있는 정보내용이 실제 현실에서 발생하지 않음

국내 신용평가사, 증권사 애널리스트, 또는 금융 전문가가 기업의 신용등급 전망을 ‘긍정적’ 또는 ‘부정적’으로 발표했으나 실제 신용등급이 상향 또는 하향으로 이어지지 않는 경우가 35개로 약 17% 존재했다. 예를 들어, “계속되는 ‘테라 효과’…하이트진로, 신용등급 전망 상향”과 같이 화이트진로의 신용등급이 상향한다는 뉴스가 존재했으나 실제 신용등급 상승까지 이어지지 않았다. 이처럼 한 등급 내에서 ‘부정적’, ‘안정적’, ‘긍정적’과 같은 세 단계가 존재하지만, 해당 단계는 변화 가능성을 높일 뿐 실제 반영되지는 않는다는 것을 알 수 있다.

(1) 신용등급을 유지한다는 뉴스, (2) 뉴스 정보가 시차를 두고 미래에 반영되는 뉴스는 전체의 약 20%로 신용등급 변동과 관련된 정보가 담겨 있는 뉴스이다. 하지만 (1) 신용등급을 유지한다는 뉴스에도 한계점이 존재했다. 기업의 회사채 신용등급 추이에 따라 신용등급이 유지한다는 뉴스의 의미가 중립이 아닌 긍정적/부정적으로 나누어진 경우가 존재했다. 본 연구의 뉴스지수는 뉴스를 긍정, 부정, 중립으로 나누어 신용등급의 상승, 하락, 유지를 분류하고자 하는 방식이기 때문에 문제가 존재한다. 이는 감성분석 모형의 추가학습을

통해 해결될 수 있는 가능성이 존재한다. (2) 뉴스 정보가 시차를 두고 미래에 반영되는 뉴스는 본 연구의 분석틀과 실제 상황과의 차이에서 발생하는 한계이다. 앞서 말했듯이 실제 신용평가사는 사람의 주관적 결정도 포함되기 때문에 신용등급 변화와 관련된 사건이 발생해도 즉각 반영되지 않을 수 있다. (3) 회사채 신용등급과 관련된 문맥이 아닌 경우, (4) 계열사에 대한 뉴스, (5) 신용등급에 대한 단순한 사실 기술 뉴스는 크롤링과 관련된 문제로 정교하게 수집할 수 있다면 해결될 수 있을 것이다. (6) 텍스트에 있는 정보내용이 실제로 발생하지 않음으로 분류된 뉴스는 주로 신용등급이 상향 또는 하락으로 예측했지만 신용등급이 유지된 경우들이었다. 이는 뉴스를 작성한 사람의 예상이 빗나간 것으로 사람도 기업의 신용등급을 예측하기 어렵다는 것을 내포한다.

추후 연구를 통해 (3), (4), (5), (6)으로 분류된 뉴스를 제외하고 (1)과 (2)와 같은 의미 있는 뉴스를 수집할 수 있다면, 뉴스지수를 활용하여 신용등급 예측 모형 성능을 높일 수 있을 것으로 기대된다.

V. 결 론

본 논문은 기업 신용등급과 연관한 뉴스 텍스트를 이용하여 만든 감성변수가 기존의 머신러닝 모형이 오분류한 관측치를 정분류할 수 있도록 도움을 주는지, 그 유용성과 한계에 대해 검토하였다. 기업의 신용등급 예측에 도움이 되는 뉴스가 점차 증가하고 있지만 여전히 기업들의 뉴스 노출도(coverage)가 매우 적고, 뉴스가 하나의 포털 사이트에 정리되어 있지 않고 여러 개의 사이트에 흩어져 있는 점 등은 신용등급 예측에 텍스트를 활용하는 것을 어렵게 한다. 본 논문에서는 이를 해결하기 위해 구글 검색을 통한 뉴스의 제목이나 요약 추출을 시도하였고, 만들어진 뉴스지수를 기존의 재무관련 수치데이터와 함께 신용등급 변화 예측에 사용하였다. 그 결과 분류성능이 향상될 수 있는지, 뉴스지수의 가능성에 대해 확인하였다.

그러나 신용등급 변화 예측을 위해 텍스트 정보를 활용하는 것과 관련하여 해결해야 하는 여러 한계점이 존재한다. 첫 번째는 제목이나 요약이 아닌 뉴스의 본문 속에 텍스트나 도표 속에 중요한 정보가 들어있는 경우이다. 각 포털 사이트별 본문을 크롤링하는 것은 기술적인 한계로 구현하지 못했다. 두 번째는 정확하게 신용등급과 연관한 뉴스 추출이 어려웠다. 기업명이 일반적으로 사용되는 단어와 동음이의어인 경우와 기업의 고객 신용등급과 연관된 뉴스가 수집되는 경우가 존재했다. 세 번째는 크롤링된 뉴스가 신용등급과 관련된 뉴스이지만 변화 방향에 대한 정보가 없는 뉴스들이었다. 마지막으로 신용등급이

유지한다는 의미를 가진 뉴스의 감성이 자연어 처리의 기술적 한계 등으로 인해 긍정적 또는 부정적으로 분류될 수 있는 한계가 존재한다. 기업의 신용등급 추세에 따라 등급 유지의 의미가 긍정적/부정적인 의미를 가질 수 있기 때문이다. 마지막으로 본 연구의 분석틀은 현실적인 상황과 차이가 존재한다. 현실적인 유용성을 생각할 때는 등급 변화 예측은 수시로 이루어져야 하지만 모형의 한계로 인해 특정한 시차인 결산 후 6개월만을 정하여 분석하였다. 이런 점들이 개선된다면 향후 더 정확한 신용등급 변화 예측이 가능할 것이다.

참 고 문 헌

- 오세경, 최정원, 장재원, “빅데이터를 이용한 딥러닝 기반의 기업 부도예측 연구”, KIF working paper, 8, 2017, 1-113.
- 황선웅, “국내 및 국제간 기업신용평가 비교연구”, 재무관리논총, 제11권 제1호, 2005, 61-99.
- Bonsall, S. B., “The Impact of Issuer-pay on Corporate Bond Rating Properties: Evidence from Moody’s and S&P’s Initial Adoptions,” *Journal of Accounting and Economics*, 57, (2014), 89-109.
- Cheng, M. and M. Neamtiu, “An Empirical Analysis of Changes in Credit Rating Properties: Timeliness, Accuracy and Volatility,” *Journal of Accounting and Economics*, 47, (2009), 108-130.
- Dichev, I. D. and J. D. Piotroski, “The Long-n Stock Returns Following Bond Ratings Changes,” *The Journal of Finance*, 56, (2001), 173-203.
- Dimitrov, V., D. Palia, and L. Tang, “Impact of the Dodd-frank Act on Credit Ratings,” *Journal of Financial Economics*, 115, (2015), 505-520.
- Hand, J. R. M., R. W. Holthausen, and R. W. Leftwich, “The Effect of Bond Rating Agency Announcements on Bond and Stock Prices,” *Journal of Finance*, 47, (1992), 733-752.
- Hite, G. and A. Warga, “The Effect of Bond-rating Changes on Bond Price Performance,” *Financial Analysts Journal*, 53, (1997), 35-51.
- Holthausen, R. W. and R. W. Leftwich, “The Effect of Bond Rating Changes on Common Stock Prices,” *Journal of Financial Economics*, 17, (1986), 57-89.
- Huang, Z., H. Chen, C. J. Hsu, W. H. Chen, and S. Wu, “Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study,” *Decision Support Systems*, 37(4), (2004), 543-558.
- Jabeur, S. B., A. Sadaoui, A. Sghaier, and R. Aloui, “Machine Learning Models and Cost-sensitive Decision Trees for Bond Rating Prediction,” *Journal of the Operational Research Society*, 71(8), (2020), 1161-1179.
- Jang, E., H. Choi, and H. Lee, “Stock Prediction Using Combination of BERT Sentiment Analysis and Macro Economy Index,” *Journal of the Korea Society of Computer and Information*, 25(5), (2020), 47-56.

- Kaplan, R. S. and G. Urwitz, "Statistical Models of Bond Rating: A Methodological Inquiry," *Journal of Business*, 52, (1979), 231-261.
- Kim, J. S. and Y. M. Choi, "Development of a Bond Rating Prediction Model Based on Financial and Stock Price-based Variables, Study on Accounting," *Taxation & Auditing*, 43, (2006), 185-217.
- Kim, K. and J. Kim, "Development of Bond Rating Prediction Model for Effective Interest Rate Estimation," *Korean Accounting Journal*, 11, (2002), 81-100.
- Kim, S. J. and H. Ahn, "Application of Random Forests to Corporate Credit Rating Prediction," *The Journal of Business and Economics*, 32, (2016), 187-211.
- Lee, S. and S. Oh, "LSTM-based Deep Learning for Time Series Forecasting: The Case of Corporate Credit Score Prediction," *The Korea Association of Information Systems(Kais)*, 29(1), (2020), 241-265.
- Lee, Y., "A Study on the Usefulness of Considering Text Information for Forecasting Corporate Bond Credit Rating," Master's Thesis, 2023.
- Lu, H., F. Tsai, H. Chen, M. Hung, and S. Li, "Credit Rating Change Modeling Using News and Financial Ratios," *ACM Transactions on Management Information Systems*, 3(3), (2012) 1-30.
- Mai, F., S. Tian, C. Lee, and L. Ma, "Deep Learning Models for Bankruptcy Prediction Using Textual Disclosures," *European Journal of Operational Research*, 274, (2019), 743-758.
- Park, H., Kang, J., Heo, S., and Yu, D., "Comparative Study of Prediction Models for Corporate Bond Rating," *The Korean Journal of Applied Statistics*, 31(3), (2018), 367-382.
- Peavy, J. and S. Edgar, "An Expanded Commercial Paper Rating Scale: Classification of Industrial Issuers," *Journal of Economics, Finance and Accounting*, 11(3), (1984), 397-407.
- Pinches, G. E. and K. A. Mingo, "A Multivariate Analysis of Industrial Bond Ratings," *Journal of Finance*, 28, (1973), 1-18.

THE KOREAN JOURNAL OF FINANCIAL MANAGEMENT
Volume 41, Number 2, April 2024

The Usefulness of News Text Information for Forecasting Corporate Bond Credit Rating*

Yejin Lee** · Hosuk Noh*** · Cheol-Won Yang****

〈Abstract〉

This research investigates the efficacy of incorporating textual data, specifically company-related news, alongside established financial and accounting factors in forecasting Korean corporate bond ratings. We introduce a novel variable, news index, based on news text, aiming to enhance the accuracy of a credit rating prediction model within six months following a firm's fiscal year-end. During the news index development, we observe the value of news in credit rating prediction, yet identify certain limitations. Addressing these limitations through additional research is anticipated to yield a more meaningful news index.

Keywords : Corporate Bond, Credit Rating, News, Sentiment Analysis, Machine Learning

* We express our gratitude to the anonymous reviewers for providing valuable feedback on this paper. This paper is based on Yejin Lee's master's thesis.

** First Author, Graduate Student, Department of Statistics, Sookmyeong Women's University,
E-mail: leeyj@sookmyung.ac.kr

*** Co-Author, Professor, Department of Statistics, Sookmyeong Women's University,
E-mail: hsnoh@sookmyung.ac.kr

**** Corresponding Author, Professor of Finance, School of Business Administration, Dankook University,
E-mail: yang@dankook.ac.kr