

재무제표 주석의 텍스트 분석 통한 재무 비율 예측 향상 연구¹

Financial Footnote Analysis for Financial Ratio Predictions based on Text-Mining Techniques

최 형 규 (Hyoung-Gyu Choe) 한양대학교 비즈니스인포매틱스학과²

이 상 용 (Sang-Yong Tom Lee) 한양대학교 경영대학³

ABSTRACT

Since the adoption of K-IFRS(Korean International Financial Reporting Standards), the amount of financial footnotes has been increased. However, due to the stereotypical phrase and the lack of conciseness, deriving the core information from footnotes is not really easy yet. To propose a solution for this problem, this study tried financial footnote analysis for financial ratio predictions based on text-mining techniques. Using the financial statements data from 2013 to 2018, we tried to predict the earning per share (EPS) of the following quarter. We found that measured prediction errors were significantly reduced when text-mined footnotes data were jointly used. We believe this result came from the fact that discretionary financial figures, which were hardly predicted with quantitative financial data, were more correlated with footnotes texts.

Keywords: Service innovation, Platform business, Artificial intelligence, Chatbot, Recommendation system, Transaction cost theory, Service personalization, Customer data

1) 이 논문은 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017S1A3A2066740)

논문접수일: 2020년 5월 28일; 1차 수정일: 2020년 6월 6일; 게재확정일: 2020년 6월 11일

2) 제 1저자(chlgudrbdn@hanyang.ac.kr)

3) 교신저자(tomlee@hanyang.ac.kr)

1. 서론

K-IFRS(International Financial Reporting Standards)가 2011년 의무 도입된 이후 기업 공시에서 주식이 차지하는 비중은 양적, 질적으로 커졌다. 구체적으로 K-GAAP(Korean Generally accepted accounting principles)를 따랐던 2010년에는 평균 78장이었던 반면 K-IFRS의무 도입 후 3년간 평균 102장으로 증가했으며, K-GAAP보다 충분한 정보를 제공한다는 설문 결과가 담긴 연구가 존재한다(황선필 등 2017). 그러나 동일 논문에서는 부족한 간결성과 상투적 표현 등으로 인해 핵심 정보를 파악하기 어렵다는 설문 결과가 존재한다. 특히 이용자와 외부감사인에게 이 문제가 더 크게 받아들여지는 것으로 분석됐다. 본래 재무제표 본문에 숫자로 적기 어려운 부분은 주석에 적음으로써 기업 외부의 정보이용자들은 주석을 통해 정보 획득을 할 수 있어야하나, 이처럼 핵심 정보 파악이 어렵다면 정보비대칭(Information asymmetry)이 증가할 수 있다.

정보비대칭을 줄이는 것은 투자자 입장에서 중요한 문제이다. Bushman and Smith(2001)에 따르면 공시의 질이 높을 경우 정보비대칭 수준이 완화될 수 있고, 이는 합리적인 투자 결정으로 이어질 수도 있다. 또한 경영자 입장에선 공시 수준이 증가하면 정보비대칭을 감소시켜 자본비용을 낮추는 효과가 있다고 주장했다(모예린 2018; Lang and Lundholm, 1996). 그러나 동시에 경영자는 자신에게 불리한 정보는 소극적으로 공시하거나, 사적 이익 추구를 위해 내부정보를 시장에 공개되는 것을 회피하고자 관련 공시의 수준을 감소시키려는 경향이 있다(Verrecchia 1983). 실제로 주석에서 경영자의 재량이 크게 행사될 수 있는 특수관계자 거래를 분석한 결과, 특수관계자 거래금액이 클수록 관련 주석 공시 수준이 감소하는 것으로 나타난 바 있다(김태동 등 2016). 이러한 경영자의 재량권 행사에

대한 징후들은 전문적인 지식을 갖고 다른 기업의 공시와 비교한 뒤 해당 기업에 일일이 문의하지 않는 한 일반적인 투자자는 확인하기 어렵다.

본 연구에서는 정보기술을 이용한 자동화로 이러한 문제를 해결할 것을 제안한다. 주석과 같은 정성적 데이터를 전자공시시스템(DART; Data Analysis, Retrieval and Transfer System)에서 적시에 정리하여 제공해주는 인프라가 구축된다면 자동화된 시스템은 빠르게 대량의 텍스트를 처리할 수 있으며, 한번 전문성을 축적하면 그 이후로는 사람의 피로처럼 생물학적 이유로 성과가 급변할 일이 없기 때문이다. 이 같이 비정형 데이터인 텍스트에서 유용한 정보를 추출하고 가공하는 기술을 텍스트 마이닝(Text Mining)이라 일컫는다.

텍스트 마이닝을 구현하는 데에는 여러가지 접근 방법이 있으나 기계학습(Machine learning), 특히 최근 크게 발달한 심층 신경망(Deep Neural Network; DNN)이 정성적 데이터를 처리하는 데에 크게 두각을 나타내는 중이다. 다만 기계학습에 쓰일 데이터에는 최소한의 패턴이 요구되는데, 예측에 도움될 패턴이 인공 신경망이 인식할 수 있는 형태로 존재하지 않는다면 유의미한 통찰을 추출하기 어렵기 때문이다. 때문에 본 연구에서는 신경망을 통해 주석이 예측에 미칠 영향을 확인하는 것이 목적이다. 즉, 정량적 지표만으로 예측하는 기계학습 모델에 비해 한국 재무제표의 주석을 추가적으로 입력하여 학습시킨 모델은 재무비율 예측력이 증가하는지를 확인하고자 하는 것이다.

2. 선행연구

2.1 재무비율 예측과 EPS

구체적인 예측 대상을 좁히기에 앞서 재무제표가 사용되는 과정을 살펴보면, 우선 재무제표 분석은 투

자에 있어 기본적 분석에 해당하는 과정이라는 점에서 시작해야 한다. 공신력 있는 회계 정보를 토대로 양적 그리고 질적 기본적 분석을 하여 추산한 기업가치를 바탕으로, 나름의 투자 원칙에 따라 포트폴리오를 구성하게 된다. 이때 전혀 다른 규모와 특성을 지닌 기업의 수익을 단순 비교할 수 없다. 그러나 재무 비율은 항목을 대응시켜 비율을 산정한 상대적인 값을 사용하기 때문에 기업간 수익성, 재무위험 등을 비교할 수 있다. 재무제표의 양적 데이터만 사용한 비율 분석은 예비적인 분석이라는 한계를 갖고 있어, 질적 데이터를 이용한 분석으로 이를 보완할 필요가 있다(김권중 2015). 즉, 재무제표 분석과정에서 재무 데이터와 밀접한 질적 데이터를 담은 주석은 재무비율 분석을 보완하는 데 필요한 정보가 있다고 추측할 수 있다.

재무비율에는 여러가지가 있으나 네이버 금융, 다음 금융과 같은 금융 포털 사이트 등에서 공통적으로 가장 먼저 눈에 띄는 자기자본이익률(Return On Equity, ROE), 주가수익비율(Price Earnings Ratio, PER), 주당 순이익(Earning Per Share, EPS), 주당순자산가치(Book-value per share, BPS), 부채비율(Debt to Equity Ratio)과 같은 비율이 주요 관심사라 추측할 수 있다. 이 중에서 비교적 선행 연구가 많아 예측 대상으로 중시되는 것으로 보이는 재무비율은 주당 순이익 EPS로 다음과 같이 계산된다.

$$EPS = \frac{\text{보통주에 귀속되는 당기순이익}}{\text{가중평균 유통 보통주식수}}$$

주식의 종류, 순이익의 종류, 액면 분할과 같은 주식 수 변화 조정 등 세부적인 사항에 따라 계산을 조금씩 다르게 할 수 있으나 기본적인 틀은 이를 따른다. 기업간 규모가 상이할 경우 기업간의 수익력 비교에 있어 당기순이익이나 경상이익보다는 주당순이익이 더 유용한 지표가 되며, 과거와 미래의 영업성과 예측과 투자 의사결정에 사용된다(송동섭 1998). 황선희(2006)

의 연구에서 정리한 바에 따르면 다음과 같은 용도로 쓰인다.

① 시장에서 거래되는 주가의 상대적 크기를 판단하여 투자할 주식을 선정하는 데 많이 사용되는 주가수익비율(PER) 계산의 기초가 된다.

② 기업합병에 따른 주식교환비율 산정의 기초자료로 활용되기도 한다.

③ 새로운 주식을 발행하는 경우 주식의 발행가격 산정 자료로 활용되기도 한다.

④ 자본시장의 국제화에 따른 국제적 투자판단의 지침으로 이용될 수 있다.

특히 주식이 저평가 또는 고평가를 받는지를 현재 시장의 평가로 가늠하는 지표인 E/P ratio(Earning/Price ratio) 또는 그 역수인 주가수익비율(PER)은 시가총액에 당기순이익을 나누는 그 계산식 상의 특징으로 기업가치에 주가가 변한다는 데에 믿음을 둔 가치주 투자 전략 등지에서 자주 사용되는 것으로 보인다(장경천 2007). 계산식을 보면 알 수 있듯 당기순이익과 관련이 깊은 지표이기 때문이다. 때문에 EPS가 PER 계산에 사용되는데 그 전개식은 다음과 같다.

$$\text{주가수익률(PER)} = \frac{\text{시가총액}}{\text{당기순이익}} = \frac{\text{시가총액/주식수}}{\text{당기순이익/주식수}} = \frac{\text{주가}}{\text{주당순이익(EPS)}}$$

PER 계산에 사용되는 EPS는 지난해에 발생한 당기순이익에 기반한 EPS를 쓸 수도 있고, 직전 4분기 합계 또는 평균 EPS를 쓸 수도 있다. 주식은 과거의 성과뿐 아니라 미래의 기대 또한 가격에 반영된다는 가정하에 미래의 EPS를 예측한 뒤 이를 과거 EPS와 합하거나 평균한 값을 사용할 수도 있다. 따라서 PER이 투자에 중시되는 만큼 EPS의 예측 또한 중요한 가치를 갖게 된다.

EPS를 예측하는 데에는 여러 재무지표들이 사용될 수 있으며 다음과 같은 선행연구에서 사용된 지표들을 예시로 들 수 있다.

$$EPS_t = f\left(\frac{\text{영업활동현금흐름} - \text{현금배당}}{\text{유형자산} + \text{기타자산} + \text{유동자산} - \text{유동부채}}\right) \quad (\text{Lai and Li 2006})$$

$$EPS_t = f(\text{일별 재고자산 중 매출, 고정자산 회전을, 자본 대비 영업 이익, 자본 대비 세전 이익, 타이완 지수, 대주주 지분, 외국인 투자자 지분}) \quad (\text{Chen et al. 2019})$$

$$EPS_t = f(\text{영업활동현금흐름}_{t-1}, \text{재무활동현금흐름}_{t-1}, \text{투자활동현금흐름}_{t-1}) \quad (\text{황선희 2006})$$

$$EPS_t = \sum_{j=1}^n \alpha_j \log sig \left(\beta_{0j} + \sum_{i=1}^4 \left\{ \begin{array}{l} \beta_{1i} EPS_{t-i} + \beta_{2i} \text{재고자산}_{t-i} + \beta_{3i} \text{매출채권}_{t-i} + \\ \beta_{4i} \text{자본적지출}_{t-i} + \beta_{5i} \text{판관비}_{t-i} + \beta_{6i} \text{매출총이익}_{t-i} \\ + \beta_{7i} \text{유효세율}_{t-i} + \beta_{8i} \text{노동생산성}_{t-i} \end{array} \right\} \right) \quad (\text{Cao and Parry 2009})$$

이외에도 EPS를 예측하기 위해 다른 재무 지표를 사용한 연구들은 여럿 존재한다. 본 연구에서는 Ramya(2017)이 정리한 EPS 예측에 관련한 연구를 정리한 표를 참조하여 비교 후 Cao and Parry(2009) 연구에서 사용된 모델을 기초로 삼기로 했다. 비교적 많은 기업을 대상으로, 재무적 근거가 있는 연구에서 사용한 재무 정보를 사용했으며, 상대적으로 구현이 쉽고 개량이 용이한 신경망 모델을 통해 괜찮은 성능을 보였기 때문이다. 물론 Cao and Parry(2009)의 연구는 해외의 기업을 대상으로 하였으므로 회계기준이 다른 데서 생기는 오차가 있을 수 있다.

그러나 한국의 데이터를 대상으로 EPS 예측을 주제 삼은 논문이 드문 실정이다. 영업현금흐름과 투자현금흐름 변수가 EPS와 유의한 상관관계를 보인다 생각할 수 있으나, 계수 자체가 0에 한없이 가까우며 조정된 결정계수가 낮아 설명력이 적은 것으로 보아 이는 샘플의 수가 많기 때문에 일어난 현상으로 보인다(Lin et al. 2013). 분기 현금흐름이 그 다음 분기 당기순이익과 관련이 있을 가능성은 있으나, 해당 모델에선 총 발생액에 대한 고려가 부족하다. 본 연구에서는 해당 분기의 현금흐름 보다는 발생주의에 입각해 미래에 들어올 수 있는 돈의 규모가 미래의 EPS 예측에 더 중요하다고 판단했다. 때문에 해당 연구에서 사용된 모델은 후보선상에서 제외했다

2.2 텍스트 마이닝을 이용한 재무비율 예측

국내의 경우 텍스트와 같은 정성적 데이터를 모델에 입력해 재무적 지표 예측에 사용하는 연구는 일반적으로 감정 표현의 개수나 비율로 수치화 하여 하나의 변수로 압축해 표현하는 것이 일반적이다. 뉴스와 같은 정성적 데이터를 감성 값의 극성 분석 후 재무 비율과 함께 모델에 학습시켜, 중소 건축 기업 부도 예측에 사용하는 (조남옥·신경식 2016) 연구가 그 예시이다. 보다 정형화된 텍스트로 이뤄진 증권발행신고서의 톤 분석(Tone Analysis)을 통해 어조에 따라 최초공모가액 대비 최종 공모가격 비율에 미치는 영향을 본 연구(김용석·조성욱 2019)도 존재한다. 그러나 정작 투자자에게 있어 공신력 있을 한국 기업의 재무제표 주석을 대상으로 한 연구는 그 수가 적은 편이다. 그중 본 연구의 방향성과 밀접한 모예린(2018)의 연구는 금융업을 제외한 2011년부터 2013년 한국 기업의 사업보고서 재무제표 주석의 코사인 거리(Cosine distance)를 정보비대칭 수준을 낮추는 대리변수(proxy variable) 중 하나로 사용해 자기자본비용, 주식거래량, 이익반응계수 간의 관계를 확인했다. 특히 그 과정에서 주석의 변화는 경영자 재량권을 측정하는 지표가 될 수 있음을 주장했다.

해외의 경우 <표 1>에 정리된 것과 같이 예측에 중점을 둔 많은 연구가 이뤄져 있다. 이중 미국 제조업 기

<표 1> 해외의 공시 보고서를 이용한 선행 연구 사례

| 저자 | 주요 설명변수 | 피설명변수 |
|----------------------|---|------------------------------|
| Chen et al. (2013) | 경영자의 주관적 표현이 담긴 텍스트의 난독화 낙관적 단어 사용 | 비기대수익 (Annual earning z값 사용) |
| Lee et al. (2014) | 8K 보고서 NMF(Non-negative Matrix Factorization) 임베딩, 재무 이벤트 | 5일내 주가의 상승, 하락, 유지 여부 예측 |
| Qiu et al. (2014) | 제조업 기업의 10K, 10K405등 연례 보고서 텍스트를 BOW 임베딩 | EPS 변화, SAR 예측 |
| Karapandza (2016) | 10K 보고서 미래 시제(will, shall, going to) 빈도 기반 척도 | Abnormal Return |
| Trusov et al. (2016) | 10K 보고서 BOW 임베딩 | 재무 위험 예측(일정기간 주가 변동률 표준편차) |
| Lee et al. (2018) | 10K 보고서 1A섹션 텍스트(Risk factor 관련) 감정 분류 점수화 | IT기업 매출 예측 |
| Meier et al. (2018) | 10K, 10Q 보고서 텍스트의 감성사전 기반 지표 | ROA 예측 |
| Rawte et al. (2018) | 10-K리포트 1A섹션 텍스트 (Risk factor 관련) GloVe 임베딩 후 감성 점수 | 은행의 실패 또는 성공 여부 예측 |

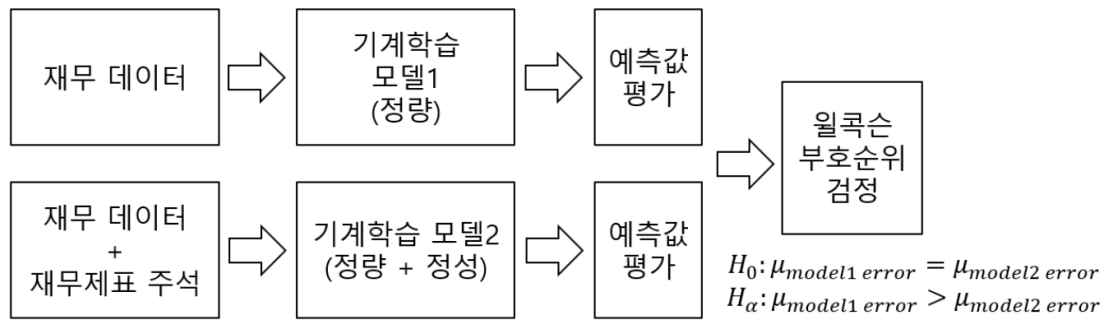
업의 10K, 10K405등 연례 보고서 텍스트로 EPS 변화량, SAR(Size Adjusted Return)이 제 1사분위 이하, 제 3사분위 이상, 그 외의 경우 중 어디에 위치 할지 예측한 Qiu et al. (2014)의 사례가 본 연구의 방향과 가장 유사하다고 본다. 물론 미국은 US-GAAP(United States - Generally Accepted Accounting Principles)라는 회계 기준을 따르고, 대상이 되는 텍스트가 주석이 아니므로 직접 비교 대상이 아니라 할 수 있다. 그러나 한국처럼 IFRS를 따르는 브라질의 재무제표에 대해서도 Qiu et al. (2014)의 연구와 유사한 방식으로 재무제표 주석을 이용해 부채비율의 변화량을 예측한 바 있다(Maeda et al. 2018). 본 연구는 선행 연구의 방향성을 참조해 한국의 재무제표 주석이 재무 비율의 예측에 미치는 영향을 예측하는 것을 차별점으로 두고자 한다. 이는 단순 사실 나열만을 목적으로 하는 주석은 감성 분석이 어렵기 때문이며, 동시에 한국의 재무제표 주석을 예측에 사용하려는 연구가 아직은 부족하기 때문이다.

3. 연구 설계 및 방법론

3.1 연구 설계

본 연구의 핵심 가설은 “재무 데이터로만 학습된 모델 보다 주석을 추가한 모델의 예측 오차가 작다”는 것이다. 이를 위해, 조남옥·신경식(2016)의 연구 방법론과 같이 주석 텍스트가 예측에 도움이 되는지 확인하기 위한 대조군으로 정량적 데이터만을 이용해 학습시킨 심층 신경망 모델 1의 성능을 확인 후, 정성적 데이터도 포함하여 학습시킨 모델 2의 성능을 실험군으로 둔 뒤 비교하는 실험방법론을 사용한다. <그림 1>은 해당 과정을 간략히 표현한 도식이다.

더불어, 별 의미가 없을 노이즈 데이터 임에도 신경망 모델 속 블랙 박스의 최적화 과정에서 우연히 성능이 좋아진 것일 수 있다. 또는 노이즈 데이터를 넣으면 정규화 효과(Regularization Effect)가 일어나 모델 성능이 개선될 수도 있다(Bishop 1995). 때문에 전혀 무관할 것으로 보이는 IMDB 영화 리뷰 텍스트(Laksh-



<그림 1> 연구 절차 도식화

mipathi N 2019)를 주식 대신 넣은 모델 3의 성능과 비교한다. 임의의 숫자를 입력 데이터로 사용해도 무방하나 텍스트가 가진 특유의 패턴이 영향을 줬을 가능성 또한 배제할 수 없기 때문에 랜덤한 숫자 대신 무관한 텍스트를 노이즈로 넣었다.

3.2 데이터 수집

본 연구에서는 크게 다음의 조건을 만족시키는 샘플을 선정하고자 한다.

- (1) 2013년부터 2018년까지 유가증권 및 코스닥 시장 상장 기업
- (2) 인수 합병 이슈가 없는 기업
- (3) 금융업 및 보험업에 속하지 않은 기업
- (4) 분기 결산이 3월 또는 9월, 반기 결산이 6월, 사업보고서 결산이 12월인 기업
- (5) 상장회사협의회가 제공하는 TS2000 데이터베이스, 그리고 DataGuide에서 재무 자료가 이용 가능한 기업

정량적인 재무 데이터는 재무 데이터 제공 어플리케이션인 DataGuide에서 취합했다. Cao and Parry(2009) 연구에서 사용된 신경망 모형은 직전 4분기 데이터를 필요로 하므로 2012년부터 2018년 사이의 유가증권 및 코스닥에 상장된 기업의 재고자산, 매출채권, 자본적 지출(CAPEX), 매출총이익, 판매관리비, 유효 세율, 매출액, 직원수, 주당순이익(EPS)을 수집했

다. 또한 규모에 따른 조정과 필터링 등을 위해 사용되 는 보통주의 수, 총자산, 산업분류 코드 같은 데이터를 수집하였다. 추가로 인수합병 이슈가 있는 기업은 제외 하기 위해 재무 데이터 베이스 TS2000에서 해당 기업 의 목록을 수집했다.

정성적인 재무제표 주석의 텍스트 데이터는 전자공 시시스템 홈페이지에 올라온 정기 공시 중 사업보고 서, 반기보고서, 분기보고서를 웹 브라우저 자동화 툴 인 selenium으로 크롤링하여 데이터를 공시 별로 저 장하였다. 식별 및 분류를 위해 법인 구분(유가증권, 코스닥), 공시대상회사, 종목코드, 보고서명, 공시 식별 번호, 정정여부, 첨부된 개별 재무제표 주식 파일 식별 번호, 첨부된 연결 재무제표 주식 식별 번호, 개별재무 제표 주식 텍스트, 연결재무제표 주식 텍스트, 접수일 자 정보들을 추가적으로 추출했다. 특히 종목코드와 보고서명의 경우 추후 정량적인 재무 데이터와 결합 시 식별하기 위해서다.

연결 재무제표가 있는 기업은 연결 재무제표 기준으 로 주식 데이터를 사용한다. 이는 재무 데이터도 마찬 가지로, K-IFRS에서 지배 회사와 종속회사의 전반의 경영 실적을 중시하는 경향을 반영하기 위해서다. 공시 에서 주석이 재무제표 본문의 별도 항목으로 쓰이거나 공시의 첨부 문서에 별도 항목으로 분리되어 있지 않 고, 재무제표 본문에 섞여 있는 등의 사유로 주석을 별 도로 추출할 수 없는 경우 샘플에서 제외했다.

국제회계기준 도입 로드맵에 따르면 상장기업에 K-IFRS가 의무 적용된 것은 2011년부터이나, 모든 상장기업이 분기 또는 반기 연결 재무제표를 작성해야 하는 것은 2013년부터이다. 또한 국제회계기준 도입 과도가이므로 그 혼란이 텍스트에 반영될 수 있기 때문에, 텍스트 데이터가 가지는 패턴의 동질성을 유지시키고 기계학습의 성능을 높이기 위해 2011년과 2012년 공시는 제외하고 2013년부터 2018년 사이 공시를 데이터가 다루는 범위로 잡았다. 따라서 2013년 2분기부터 2018년 3분기 사이의 주석데이터가 사용됐다.

같은 공시 접수일에 여러 차례 정정된 공시가 올라온 경우 가장 마지막에 접수된 공시만 남기고 그 이전의 것은 중복된 것으로 간주하고 제거했다. 또한 예측 대상 분기(t)가 지나기 전에 그 이전 분기(t-1)의 데이터로 예측해야 그 가치가 있다고 보고 예측 대상 분기가 지나기 전에 공시가 정정된 경우를 최종 샘플로 삼는다.

제조 기업이 전혀 상관없는 기업의 주식을 갖고 있다면 이것은 영업 외적인 목적으로 자산을 가진 것으로 취급될 수 있으나 금융 기업이 다른 기업의 주식을 갖고 있다면 영업을 위한 자산으로 취급될 수 있다. 이와 같이 금융업에서는 주석에서 사용되는 단어의 의미와 맥락이 다르게 적용될 가능성이 높아, 산업 분류상 금융 및 보험업(코드 64~66)에 속하는 기업 또한 기계학습의 성능을 높이기 위해 데이터에서 제외했다.

3.3 정량적/정성적 재무 데이터 전처리

Cao and Parry(2009)에서 처리한 바와 같이 채고 자산, 매출채권, 자본적 지출, 판매관리비는 보통주 수를 나눠 크기를 조정했다. 노동생산성은 매출액/직원수에 자연로그를 취했다. 또한 선행 연구에서는 유효세율을 법인세에 법인세차감전순손익을 나눠 계산했으나, 법인세가 비용이기 때문에 음수로 기록되거나 하는 등의 데이터 품질 문제가 존재하여, 별도로 계산

된 유효세율(=계속사업법인세비용/세전계속사업이익) 데이터를 썼다.

유효세율이 $\pm 10000\%$ 이상 또는 이하인 경우는 세무에서 보는 이익과 재무에서 보는 이익에 불일치가 존재한다고 하더라도 정도가 심하다 보고 해당 샘플을 제거했다. 판매관리비와 노동생산성이 음수인 나온 경우 또한 마찬가지로 이상치로 보고 제거했다. EPS는 0일 경우 보고가 누락돼 데이터베이스에서 0으로 처리한 것으로 보고 해당 샘플은 제외했다.

텍스트 마이닝을 위해선 자연어 처리 기법 중 여러 가지를 선택해야 하는 데, 우선 장문의 문서를 최소한의 의미를 가진 형태소로 분리하는 것이 선행되어야 한다. 한국어 형태소 분석기는 여러 종류가 존재하나 업데이트가 지속적으로 이뤄지며, 개발 환경의 편의상 KoNLPy에 있는 형태소 분석 태그 패키지를 선택지로 두었다. 장문 또는 단문 처리 속도와 같은 성능 차이는 존재하나 각 형태소 분석기 간의 분류 성능에 대한 절대적 차이를 증명한 결과는 없다.

Hannanum 과 Komoran 형태소 분석기는 비교적 띄어쓰기를 잘 인식하지 못하고 단순히 체언 내지는 명사로 분류해버리는 경향을 갖고 있는 것으로 보인다. 이는 다른 형태소분석기에 비해 낮게 평가될 요소이다. 그럼에도 많은 단어가 조합되어 전혀 다른 의미를 가질 수도 있는 재무 분야의 고유명사가 상당수 포함된 주석의 경우 오히려 이런 경향이 도움이 될 수 있다. 띄어쓰기 없는 문장을 잘 분리하는 다른 형태소 분석기는 일상생활에서는 잘 쓰이지 않는 재무 분야의 고유명사를 일일이 쪼개어, 전혀 다른 형태소로 나눠 예측 성능을 간접적으로 낮출 가능성이 존재하기 때문이다. 예를 들어 ‘해외사업장순투자위험회피’라는 단어가 문장에 있다면 Hannanum은 체언으로 취급하지만, 다른 형태소 분석기는 별도로 형태소 사전에 별도로 추가하지 않는 한 ‘해외/NNG; 사업장/NNG; 순/NNG; 투자/NNG; 위험/NNG; 회피/NNG;’와 같이

나눠서 인식할 가능성이 높다. 뿐만 아니라 공식적인 보고서인 이상 맞춤형이 맞을 가능성이 높기 때문에 다른 형태소 분석기처럼 알아서 띄워 쓰고, 그 속의 형태소를 나눠 인식하는 기능이 상대적으로 필요 없다. 그러나 Hannanum은 글자 처리 수 제한이 있고 숫자를 별도로 태깅하여 필터링할 방법이 없기 때문에, 본 연구에서는 차선책인 Komoran 형태소 분석기를 채택했다.

3.4 단어 임베딩 및 SVD(Single Value Decomposition)차원 축소

텍스트는 그 자체를 곧장 기계학습의 입력 데이터로 사용하기 어렵기 때문에 이를 숫자로 바꿀 필요가 있다. 간단하게는 각 형태소가 존재하는 샘플에 특정 형태소에 해당하는 열에 더미 변수와 같이 1이라 표기하고, 형태소가 존재하지 않는 공시는 0이라 표기하는 One-Hot 인코딩 방식이 있다. 또는 Word2Vec와 같이 Skip-Gram 또는 CBOW(Continuous Bag of Words) 기법으로 특정 단어로부터 문맥이 될 수 있는 단어를 예측하거나 빠진 단어를 예측하는 식으로 문장을 학습해 벡터로 표현하는 방법이 있다(Mikolov et al. 2013).

그러나 정량적인 재무지표를 동시에 모델에 입력할 때, 각 샘플당 1차원 벡터인 재무 데이터와 각 형태소 별로 1개 이상의 차원을 가지고 이를 이어서 테이블 형태로 표현될 주식 텍스트를 동시에 학습 모델에 입력하는 것은 차원이 일치하지 않아 빈 부분을 채우거나 이에 맞게 모델을 구조를 결정하는 것이 까다롭다. 형태소를 1차원의 벡터로 이어버리는 방법이나 빈 부분을 정량적 재무지표나 0으로 채우거나 하는 방법도 있지만 차원이 각 형태소의 차원의 배수만큼 커져 처리 성능이 떨어지는 것이 우려된다. 따라서 단어를 벡터로 바꾸는 접근법은 본 연구에서는 사용하지 않는다.

본 연구에서는 TF-IDF(Term Frequency - Inverse

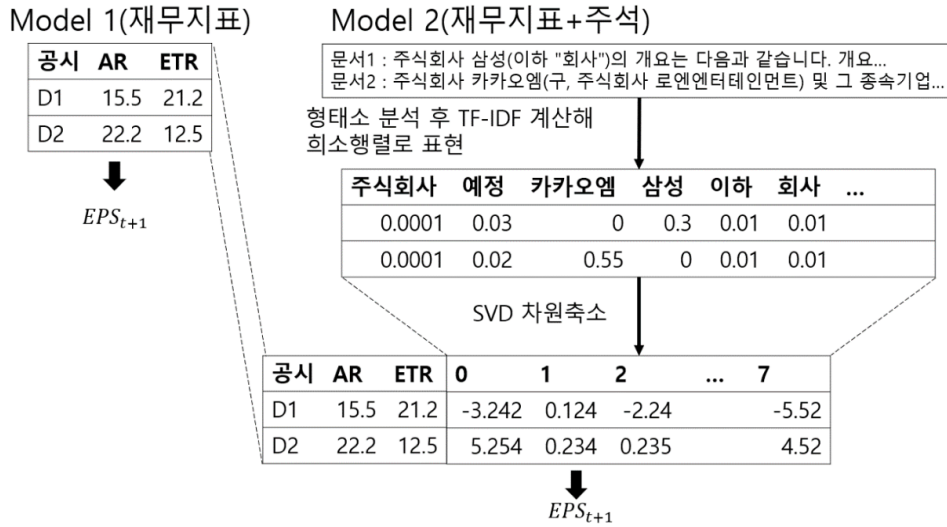
Document Frequency) 기법을 사용해 형태소를 숫자로 표현하고, 각 공시의 주식 텍스트가 담긴 문서의 특성(feature)을 표현하고자 한다. 선정 사유는 모예린(2018)의 선행 연구에서 한국 재무제표 주식의 특성을 나타내는 데 충분히 유용함을 보였기 때문이다. 예시를 들자면 각 문서별 형태소 들을 단어빈도와 역문서 빈도를 곱한 값으로 <그림 2>의 Model 2 아래의 표처럼 변화시킬 수 있다.

TF-IDF 행렬로 만든 뒤 바로 입력하기에는 행렬의 크기가 32684×70133로, 연산을 하기에는 너무 크다. 또한 형태소의 부재를 의미하는 0이 많은 희소 행렬이므로 설명력 높은 정보를 보존하면서도 이를 적절한 수준으로 차원 축소할 필요가 있다.

차원 축소의 방법으로는 절단된 특이값 분해(Truncated Singular Value Decomposition, Truncated SVD)를 사용한다. TF-IDF 기법은 단어의 의미를 고려하지 못한다는 단점이 있으나 SVD를 거치면, 설명력 낮은 정보는 삭제하고 높은 정보는 남기면서 잠재적 의미를 끌어낼 수 있어 잠재의미분석(Latent Semantic Analysis, LSA)에 사용 되는 방법이기도 하다.

문서와 단어에 대한 $m \times n$ 크기의 행렬 A 를 연구자가 사전에 지정한 k 만큼 절단하여 특이값 분해하면 TF-IDF 행렬을 줄일 수 있으며, 본 연구에서 k 값은 적어도 95% 수준까지 각 feature들의 분산을 설명할 수 있는 8로 정했다. 절단된 형태의 우향의 행렬을 다시 곱하면 원본 A 와는 다르나 필요한 특징은 여전히 가진 축소된 형태의 행렬이 계산된다.

즉, 본 연구에서 각 샘플 별로 주식데이터가 추가된다는 것은 데이터 세트에 8차원까지 축소된 벡터가 붙는 것을 의미한다. 각 벡터는 주식의 축소된 특징을 나타내며 유사한 특징을 가진 주식 문서는 가까운 공간에 위치할 가능성이 높고, 이는 주식이 가진 방향성과 잠재적인 의미를 표현한다고 볼 수 있다.



<그림 2> 모델에 입력되는 데이터 비교

4. 모델 구축

4.1 신경망 모델

신경망 모델 또는 인공신경망(Artificial Neural Network, ANN)은 두뇌에서 착안해 비선형적 구조로 정보를 전달하는 통계학적 학습 네트워크로 EPS 예측을 연구한 선행 연구에서 전반적으로 양호한 성능을 보여 기반 모델로 자주 사용 된 것을 볼 수 있다 (Ramya 2017). 다른 기계학습 모델이 일반적으로 분류 모델링 성능에 최적화된 것에 비해 연속적인 값을 예측하는 데에 적합하기 때문에 현업에서 애용되기 때문으로 짐작된다.

본 연구는 EPS 예측력이 높은 모델을 만드는 것이 아닌, 주석이 EPS 예측에 미치는 영향을 보는 것이 주목적이므로 Cao and Parry(2009)의 선행연구에서 사용한 Back propagation을 이용한 신경망 예측 모델에서 성능상의 이유가 있는 경우를 제외하고는 바꾸지 않았다.

Hidden layer 수는 1개, Hidden layer의 노드 수 또한 선행연구에서 세운 아래의 기준에 따라 결정했다.

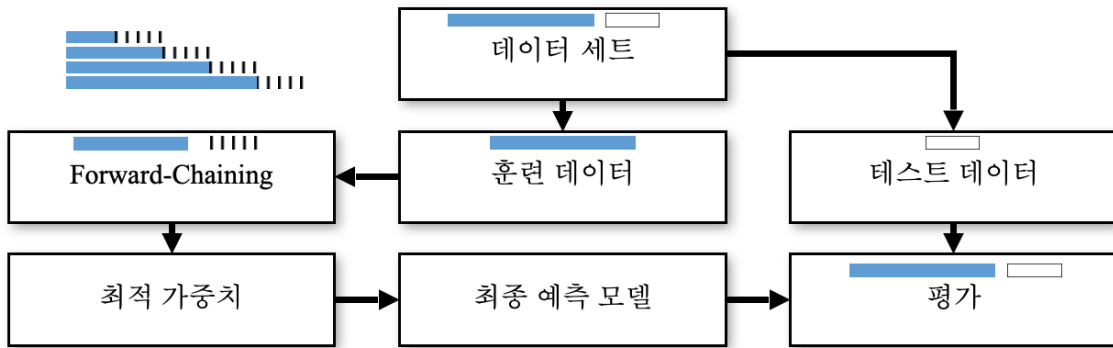
$$\text{노드 수} = 1/2 \times (\text{입력 변수의 수} + \text{출력 변수의 수}) + \sqrt{\text{훈련 패턴의 수}}$$

활성 함수는 선행연구에서 사용한 Sigmoid 대신 ReLU(Rectified Linear Units)를 사용했다. 함수가 간단하여 연산비용이 크지 않고 학습이 원본의 Sigmoid 함수보다 6배는 더 빨라지기 때문이다 (Krizhevsky et al. 2012).

4.2 모델 구축 절차

전반적 모델 구축 과정은 중첩 교차 검증(Nested Cross Validation)과 비슷하게 전개된다. 주요 차이점은 미래의 정보가 예측 모델에 반영되는 데이터 누수(Data Leakage) 문제를 피하기 위해 훈련 데이터를 k 개로 쪼개 k-fold 교차 검증을 하는 대신, 전방연쇄(Forward Chaining)를 사용한다는 점이다. Forward Chaining이란 시계열 데이터가 있으면 성능 테스트에 사용할 몇개의 미래 데이터는 테스트 데이터로 두고, 바로 이전 시계열은 검정 데이터로 두고 검정 데이터 이전 샘플은 모두 훈련데이터로 하는 방식이다(Bergmeir and Benítez 2012; Hjorth and Hjort 1982). 이를 도식화한 것이 <그림 3>이다.

Forward Chaining 검증을 통해 검정 데이터에서 가장 낮은 오차값을 얻어낸 최적의 가중치로 이뤄진 모델을 얻으면, 이제껏 학습에 사용하지 않기 위해 최



<그림 3> 순서도 원본 출처: 안드레아스 뮐러-세라 가이드(2017) 서적

초에 분리한 테스트 데이터를 예측에 사용한다. 테스트 데이터는 2016년 2분기부터 2018년 3분기까지 총 10개 분기이다. 10개 분기에 대한 예측 값을 실제 값과 비교해 나온 오차 값의 평균을 해당 모델이 보인 성능으로 본다. 이때 예측 값은 t분기에 유상증자, 액면 분할과 같이 주식수에 변화를 주는 이벤트가 있는 경우 (t-1분기 보통주 수/ t분기 보통주 수)를 곱해 그 크기를 조정한다. 이렇게 할 경우 예측된 t분기 EPS의 분모는 t-1분기가 아닌 t분기의 변경된 주식 수로 계산 된다. 재무 데이터만으로 구축된 모델 1과 재무 데이터와 차원 축소된 주식의 데이터가 같이 입력된 모델 2 모두 같은 절차를 통해 구축한다.

4.3 모델 평가 및 비교

$$RMSE(\text{Root mean square error}) = \frac{1}{N} \sqrt{\sum (y_i - \hat{y}_i)^2}$$

$$MAE(\text{Mean Absolute Error}) = \frac{1}{N} \sum |y_i - \hat{y}_i|$$

$$MAPE(\text{Mean Absolute Percentage Error}) = \frac{100}{N} \sum \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

$$MSPE(\text{Mean Squared Percentage Error}) = \frac{100}{N} \sum \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2$$

\hat{y} : predicted value, y : True value, n : size of sample

피예측변수는 연속적인 재무비율이므로 예측 성능 평가에는 연속적인 숫자형 데이터를 예측오차를 확인하는 데 자주 사용되는 Root Mean Square

Error(RMSE), MAE(Mean Absolute Error), MAPE(Mean Absolute Percentage Error), MSPE(Mean Squared Percentage Error)를 척도로 사용한다. 해당 오차 값이 크면 클수록 그 모델의 성능은 낮다는 것을 의미한다. 반대로 오차 값이 작으면 그 모델의 성능이 우수하다 할 수 있다. RMSE와 MAE는 흔히 오차를 측정하기 위한 척도로 자주 쓰이나, Outlier에 민감하며 크기 의존적 에러(Scale-dependent Errors)가 생길 수 있다. 예를 들어 실제 EPS 값이 100임에도 90이라 예측한 것과 10인데도 5라 예측한 것은 RMSE와 MAE 관점에선 전자가 더 큰 오차라고 보나, 이는 MAPE와 MSPE의 관점에서 본다면 전자의 오차는 예측해야 할 값에 비하면 10% 수준으로 오차가 난다고 본다. 물론 MAPE와 같이 크기의 의존적 에러 문제를 해결한 척도라도 실제 값이 0에 가까울 경우 예측 값이 크게 차이가 없더라도 너무 크게 오차 값이 나오는 단점이 있다. 이처럼 각 척도마다 일장일단이 있어 현상을 좀 더 세세히 이해하기 위해 여러 척도를 사용했다. 다만 비중은 MAPE와 MSPE에 둔다. 기반을 둔 Cao and Parry(2009)의 선행연구에서 사용된 오차 계산법과 거의 같기 때문이다.

모델의 성능 비교는 다른 두 집단간 평균의 차이가 유의한지 비교하는데 사용할 수 있는 윌콕슨 부호순위 검정(Wilcoxon's signed rank test)을 사용한다. 이때

두 집단이란 모델 1과 모델 2의 오차 척도의 표본이다.

두 집단이 서로 다르다고 한 이유는 모델1은 정량적 데이터로만 학습됐고, 모델2는 정량적 데이터에 정성적 데이터까지 합쳤기 때문에 비록 공통되는 부분이 있더라도 그 성격이 상이한 데이터에 기반해 학습된 모델에서 나온 것으로 보기 때문이다.

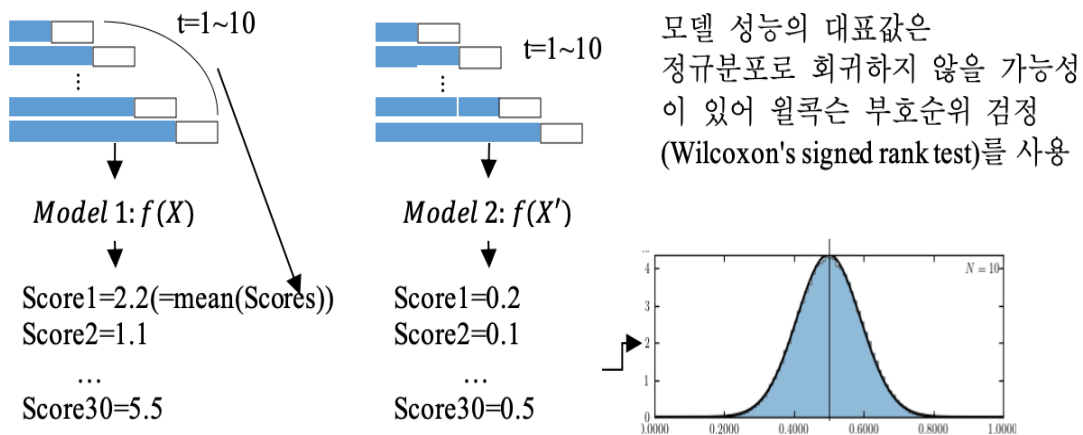
이러한 비교를 하는 이유는 신경망 모델과 같이 난수에 기반해 최적값을 찾아가는 모델의 경우 최초에 주어진 난수 값에 따라 최적해에 도달하는 과정에서 국소 최적점(Local optimum)에 머물 가능성이 있기 때문이다. 각각 모델 1, 모델 2에서 도출해낸 최적 가중치가 우연으로 서로 차이 나는 오차 값을 낸 것이 아님을 알기 위해, 난수의 seed 값을 30회가량 변경했다. 즉, 여기서 한 개의 표본은 10개 분기에 대한 예측 오차의 평균 값을 의미한다. 두 모델의 30회 시도 후 나온 오차 측정 값들의 평균이 유의하게 차이가 나는지 검정하면 모델의 성능 차이가 우연이 아님을 확인할 수 있다. 이때 모델 오차 측정 값들은 거듭해 실험할수록 평균에 정규분포로 수렴한다고 가정한다. <그림 4>는 이를 표현한 도식이다.

5. 분석 및 결과

5.1 모델 비교 결과

<표 2>에서처럼, RMSE와 MAE는 귀무가설이 채택됐으나, 증시되는 오차 측정 척도인 MAPE와 MSPE는 95%수준에서 유의하게 귀무가설을 기각했다. 즉, 주석을 포함시킨 모델은 EPS의 MAPE, MSPE 오차 측정값이 유의하게 떨어진 것을 볼 수 있다. 재량적 발생액은 비교적 주당순이익(EPS)이 큰 샘플들에 대해서 비교적 잘 설명하여 RMSE나 MAE는 달리 유의하게 감소하지 않았지만, MAPE나 MSPE는 유의하게 감소한 것으로 본다.

모델링에 사용된 샘플과 겹치는 재무 전문가들의 컨센서스를 같은 방법으로 측정한 오차와 비교해 보면 본 모델1은 비교적 관찮은 수준의 성능을 보이는 것으로 보여, 비록 MAPE가 실제 값 크기의 100%이상 틀릴지라도 사용된 재무 데이터와 모델은 한국의 EPS 예측에도 어느 정도는 유용한 것으로 보인다. 다만 <표 3>은 평균으로만 비교한 결과이며 각 분기별로 나타난 측정오차를 보면 전반적으로 컨센서스의 오차가 작다. 이는 본 연구에서 상정된 예측 시점(공시일)보다 45일정도 더 늦은 시기 이후에 3개월간 발표된 E3 예



<그림 4> 모델 성능 평가 과정

측 값이므로 시간적 이득(Time Advantage)이 존재하고, 재무전문가가 사용할 수 있는 내부 정보들과 직관을 생각해보면 당연한 결과이다.

추가로 상관없는 IMDB 영화 리뷰 텍스트를 같은 방식으로 전처리 한 뒤 모델 1의 재무 데이터와 같이 학습시켜 나온 모델 3의 오차 측정 값을 비교한 결과는 <표 4>과 같다. 여기에서는 모든 오차 척도가 유의한 차이를 보이지 않고 있음을 알 수 있다. 이는 우연히 또는 정규화 효과에 의해 주석의 텍스트가 유의한 영향

을 미친 것이 아니라는 사실을 지지하는 것으로 해석할 수 있다.

5.2 분석 결과 해석

주석에 입력되는 정보는 여러가지인 만큼 여러가지 설명이 있을 수 있다.

주석에 기재되는 항목인 스톡옵션(Stock option) 부여시 매출 증가 경향이 있다는 조영석(2005)의 연구를 참고하면, 스톡옵션을 주는 것이 주석에 기록되고 근

<표 2> 모델 1과 모델 2의 오차의 척도별 평균값 비교

$$H_0: \mu_{model1\ error} = \mu_{model2\ error}$$

$$H_a: \mu_{model1\ error} > \mu_{model2\ error}$$

| H _a : Error1-Error2 > 0 | W-statics | p-value |
|------------------------------------|-----------|---------|
| RMSE | 251 | 0.352 |
| MAE | 249 | 0.367 |
| MAPE | 346 | *0.010 |
| MSPE | 335 | *0.018 |

<표 3> 모델 1 오차, 모델 2 오차 그리고 컨센서스 척도별 평균과 비교

| Mean | Error1 | Error2 | Consensus |
|------|------------|------------|------------|
| RMSE | 2054.279 | 2036.960 | 1991.957 |
| MAE | 418.510 | 417.611 | 527.293 |
| MAPE | 411.923 | 379.525 | 436.748 |
| MSPE | 517119.628 | 356147.678 | 441195.438 |

<표 4> 모델 1과 모델 3의 오차의 척도별 평균값 비교

$$H_0: \mu_{model1\ error} = \mu_{model3\ error}$$

$$H_a: \mu_{model1\ error} > \mu_{model3\ error}$$

| H _a : Error1-Error3 > 0 | W-statics | p-value |
|------------------------------------|-----------|---------|
| RMSE | 209 | 0.686 |
| MAE | 204 | 0.721 |
| MAPE | 305 | 0.068 |
| MSPE | 274 | 0.197 |

로의욕을 고취시켰기 때문에 당기순이익이 증가했다고 생각할 수도 있다. 그러나 해당연구에서 주당순이익은 증가는 하나 유의한 상관관계를 보이지 않아 이 추측은 참고만 되는 정도이다.

또는 파생상품을 헷지(Hedge) 목적으로 사용 시 EPS 예측력에 영향을 미친다는 송재성(2011)의 연구를 참고해, 헷지 목적 파생상품에 대한 언급이 주석에 존재하는 기업은 그만큼 정보를 잘 제공하기 때문에 그에 따른 특징이 반영되어 단순 재무 데이터만 이용한 경우보다 예측력이 향상된 것으로 볼 수도 있다.

마지막으로 재무제표 주석이 재량적 발생액과 관련이 있어, 이에 대한 정보가 EPS(주당순이익)예측에 영향을 줬다고 보는 관점이다. 모예린(2018)의 선행연구에 따르면 주석의 변화는 기업 고유의 재량적 발생액과 유의한 관계가 존재한다는 실증분석 결과가 존재한다.

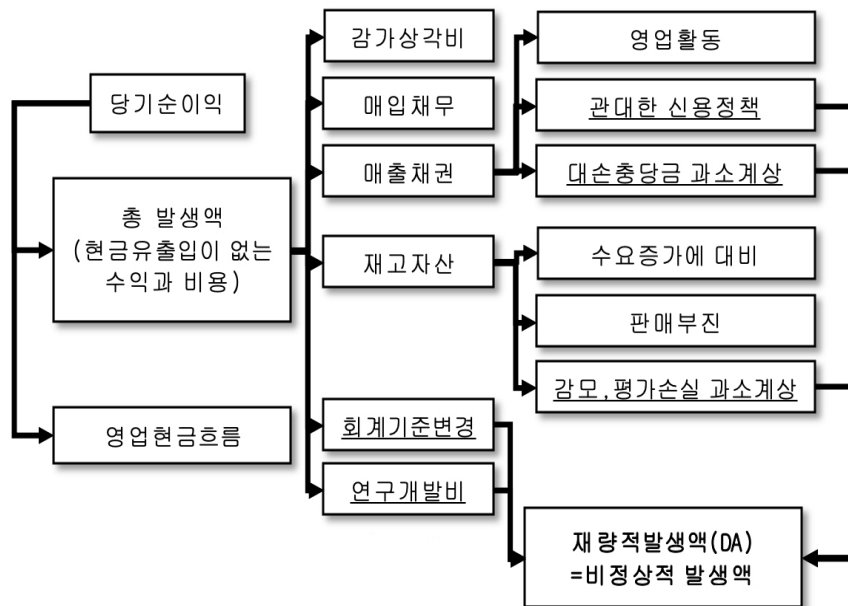
EPS의 분자인 당기순이익은 수익을 보는 관점에 따라 크게, 당장 들어온 현금만을 본 영업현금흐름과 그 외에 현금 유출입 없이 앞으로 들어올 것을 수익으로 본 총 발생액으로 나눌 수 있다. 이 총 발생액에서 정

량적인 재무 지표로 설명이 안되는 부분이 재량적 발생액으로 이는 사후적으로 추정(Jones 1991)하는 것이 보통이다. 재량적 발생액 대부분 회계 기준을 변경하여 어느정도 순이익을 조정할 경우 생기며 도식화하면 아래의 그림과 같다.

실제로 MAPE, MSPE는 유의하게 측정 오차가 줄은 반면 RMSE가 그렇지 못하다는 것은 당기순이익이 비교적 크거나 작은 기업에 대해 예측을 잘했다는 의미이고, 당기순이익이 비교적 크거나 작은 이유는 재량적발생액이 있기 때문일 수 있다.

구체적으로 어떤 양상으로 영향을 끼쳤는지는 신경망 모델이 블랙박스이고, 주석을 차원축소 하는 과정에서 각 형태소마다의 설명력을 구분하는 것이 어렵기에 명확한 설명이 쉽지 않다. 그러나 몇 가지 추측은 가능하다.

첫째, 주석에 관례적으로 사용하던 회계 기준이 변경된다는 공시를 한다면, 기계학습 모델은 다른 유사한 산업의 기업에선 사용하지 않는 형태소가 사용된 것으로 변화가 인식 될 것이다. 이러한 변화는 재량적



<그림 5> 재량적 발생액이 나타나는 부분 (FnGuide 2012)

발생액을 늘리려는 징조로 해석할 수 있다. 재량적 발생액을 변화시키는 경향은 금융감독원으로부터 감리지적을 받지 않는 한 다음 분기에도 이어질 수 있다. 변화된 재량적 발생액은 당기순이익에 영향을 미칠 것이고, 이에 기반한 EPS 예측에도 영향을 미쳤을 것으로 보인다.

둘째, 학습 데이터에 특정 회계기준으로 변경했을 경우 재량적 발생액의 증가로 당기순이익이 늘어나는 사례를 모델이 학습하고, 이후 다른 기업에서도 재량적 발생액을 변화시키기 위해 회계기준을 변경하는 것을 의미하는 형태소가 등장했을 경우 재량적 발생액이 증가한 것으로 간주하고 EPS를 좀 더 높이거나 낮춰 계산했을 것이다.

셋째, 모예린(2018)의 연구에 따르면 t-1시점에서 t시점으로 주석의 변화가 t시점 재량적 발생액과 상관있었다고 보았다. 발생액은 미래에 유입 또는 출입할 현금을 현재의 수익으로 본 것이기 때문에, 실제로 현금화 될 가능성이 높고 그것이 실현 될 시기가 1분기 뒤일 경우 이는 현금으로 들어올 것이다. 모델은 이렇게 영업활동 결과에서 정상적으로 발생하는 재량적 발생액의 경우 주석이 EPS 예측에 영향을 미친다고 학습했을 수 있다.

모델 설계의 특성과 재량적 발생액의 성격상 두 번째 추측이 보다 사실에 근접해 보인다. 어느 쪽의 추측이 맞건 주석은 경영자의 재량적 발생액을 늘리려는 의도를 감안해 EPS 예측에 반영하고, 예측 오차를 줄이는 데에 쓸 수 있다는 점에는 변화가 없다.

6. 결론

6.1 연구 요약

본 논문에서는 한국 재무제표 주석이 재무비율 그 중에서도 EPS 예측에 미치는 영향을 확인하는 방향으

로 접근했다. 연구를 위해 기존 EPS 예측을 다른 문헌을 살펴보고, 텍스트 데이터를 이용해 다른 재무지표를 예측하거나 그 관계를 확인하는 연구를 비교해보았다.

그 중에서 Cao and Parry(2009) 연구에서 만든 모델을 기반으로 재무 데이터만으로 학습된 기계학습 모델과 재무 데이터에 주석을 차원 축소시킨 값을 학습시킨 기계학습 모델을 비교하였다. 난수 Seed를 바꿔가며 모델이 낼 수 있는 오차의 표본 값을 추출한 뒤 월록슨 부호순위검정을 실시하였고 그 결과 오차가 유의하게 줄어드는 것을 확인했다.

이는 재무제표 주석이 재량적 발생액과 상관관계가 존재(모예린 2018)한다는 실증연구에 비춰 봤을 때, 과거의 재무제표에서 나타난 특징이 미래의 재량적 발생액과 관련하여 예측 오차가 감소한 것으로 보인다.

이 논문은 이제껏 자연어 처리 기술의 부족으로 재무 비율 예측 연구 대상으로 잘 쓰이지 않은 재무제표 주석을 사용한 주당순이익 예측 결과를 보여줌으로써, 숫자로 된 재무 데이터 외에 주석의 텍스트 데이터 또한 주당순이익 예측에 사용하여 자동화에 활용할 수 있음을 보여준다. 앞으로 주석 텍스트 데이터 분석을 고도화시켜 자동화에 응용하면, 로보어드바이저와 같은 서비스의 비용절감으로 이어질 것으로 기대한다.

6.2 연구의 의의

본 연구는 숫자로만 예측할 수 있는 영역을 넘어 경영자의 재량적인 영역까지 예측에 넣어 이에 대한 시장 반응을 예측하고 투자 포지션을 결정하는 데 사용할 수 있다. 가치주 투자 전략을 쓰는 이들 중 보수적인 관점에서 수익을 바라보는 이들은 재량적 발생액의 비중이 커 이익의 질이 낮은 경우를 안 좋게 보고 재량적 발생액까지 감안한 예측 모델을 꺼려할 수 있겠지만, 주식시장이 기대에 의해 형성되며 가치주 투자 전략을 사용하는 이들의 기대 또한 기대의 일종으로 보고 계산에 넣는 투자자에게 유용한 정보가 될 수 있다

본다.

증권가 안팎에서 무형자산이 가치 창출에 핵심적 역할을 하고 있는 신흥 기업의 재무제표가 가진 한계를 지적하는 목소리가 커진다는 기사가 존재한다(임근호 2018). 해당기사에서 전문가들은 이에 대해, 무형의 자산이 가진 가치를 측정하기 어려운 것은 어쩔 수 없으나 주석 등의 공시를 통해서라도 상세한 정보를 투자자에게 전달할 필요가 있다는 의견을 보였다. 실제로 재무분석가는 무형자산 비중이 높거나 특허권 취득한 기업은 이익예측 정확성이 높다는 연구가 있다(전성일·이기세 2012). 앞으로 지식 정보화 사회가 될수록 기업이 가진 무형의 서비스 전략, 네트워크, 전문가에게 종속된 암묵지(Tacit knowledge) 등이 기업의 가치에 직결될 경향이 더욱 커질 것이고, 전문가들의 의견에 따라 성실하게 이를 공시한다면 그만큼 주석이 가질 수 있는 가치는 증대할 것으로 예상된다. 또한 시간이 지나 충실한 주석 문서가 늘어날수록 본 연구에서 제안한 모델의 성능이 발전될 것으로 기대 된다. 신경망 모델의 특성상 데이터가 많을수록 성능이 점차 증대되는 경향을 갖고 있기 때문이다.

더불어 로보어드바이저(Robo-advisor)와 같은 인공지능 재무 분석 서비스에도 응용할 수 있을 것으로 예상된다. 로보어드바이저는 로봇으로 빅데이터 분석, 포트폴리오 구성 등을 제공하는 자산관리서비스이다. 물론 추세만 보고 기술적 분석 후 투자하는 기존의 시스템 트레이딩과 별 차이가 없는 서비스를 로보어드바이저라 하는 경우도 있다(박선후 2016). 그러나 시스템 트레이딩이 가진 이례적 사건이나 구조적 변화 등에 약한 면모를 극복해야 로보어드바이저가 가진 본 의미에 더 가까워질 것이다. 그러기 위해 빅데이터의 3V(Volume, Velocity, Variety)로 불리는 특징 중 다양한 정보(Variety)를 다룬다는 점을 살려 투자에 도움될 이례적 사건과 구조적 변화를 제어하면 분석 성능을 보완하는 데 본 연구가 도움이 될 것을 기대한다.

더불어 2019년 4월 부로 로보어드바이저의 직접 펀드 운용이 허용됐다. 본 연구를 기반으로 이제 인간이 처리해왔던 정성적 데이터를 컴퓨터가 일부 처리하여 프로세스의 자동화 비율을 늘리고 생산성을 향상시킬 수 있을 것이다. 금융정보 분석의 자동화로 향상된 생산성은 비용절감을 가능케 한다는 것은 금융정보 스타트업 켄쇼(Kensho)의 성공 사례 등을 통해 미뤄 짐작할 수 있다. 비용절감은 서비스 가격 인하의 유인이 되고, 간접적으로 펀드 고객들의 수수료 부담이 줄어들 가능성이 높아질 것을 기대할 수 있다.

추가적으로 주석의 텍스트 데이터와 같은 정성적 데이터를 재무 분석에 전면적인 활용에 앞서, 투자자들이 일일이 크롤링 후 전처리하는 수고 등을 줄일 필요가 있다. 때문에 전자공시시스템에서 기업 보고용 국제 표준화 체계인 확장성 경영보고 언어(eXtensible Business Reporting Language, XBRL)의 주석과 관련된 항목을 정비해 API(Application Programming Interface)와 같은 인프라지원을 할 필요가 있어 보이며, 본 연구는 이의 준거가 될 것을 기대한다.

6.3 연구의 한계점

본 연구에는 현실적인 문제로 인해 몇 가지 한계점을 가질 수밖에 없었다.

첫째, 신경망 모델은 복잡한 함수들의 네트워크로 구성되어 있기 때문에 완벽한 인과관계를 증명하기 어렵다. 설령 인과를 완벽히 설명하는 독립변수를 입력에 사용했다더라도 구체적으로 어떤 양상으로 영향을 미쳤는지에 대한 해석이 회귀분석만큼 용이하지 않다.

둘째, 주석의 허위 기재 여부에 대해서는 고려하지 못했다. 금융감독원 감리는 어디까지나 일부 기업을 대상으로 할 뿐 전체 기업에 대해 실시되지 않기 때문에 샘플이 된 기업의 재무제표 주석이 허위 기재를 하지 않았다는 증거를 제시하기가 어렵다. 이는 주석이

아닌 다른 숫자로 된 재무 데이터에도 적용되는 일로, DataGuide에서 모은 데이터가 허위 또는 잘못 기재됐을 가능성이 존재한다. 일례로 특정 기업의 법인세 항목을 확인했을 때 음수가 나온 샘플이 있는데, 비용으로 판단해 음수로 표기한 것이 그대로 데이터베이스에 반영해서 그런 것으로 추측된다. 특히 4분기 데이터가 다소 신뢰하기 어려운 경향이 있다. 이는 분기에 대한 정보가 공시되지 않아 사업보고서에 기재된 값에 9월 분기 보고서의 반기까지 누적 값을 빼는 등의 계산이 들어가기 때문으로 보인다. 데이터 정제 과정에서 문제될 것으로 보이는 부분은 최대한 제거했으나 완벽하다 보기는 어려운 면이 있다.

세째, 4분기 정보가 담긴 사업보고서는 공시 규정상 90일 이내, 보통 마지막 날인 다음해 4월 1일 즈음에 접수되고, 이를 통해 해당 연도의 1분기 정보를 예측하게 되는데 이미 1분기가 하루 지난 시점에선 공시되지 않은 정보라도 예측의 의미가 일부 퇴색할 수 있다.

넷째, 데이터 전처리가 끝난 후 발표된 연구(정건용 2019)에 따르면 회계용어로 이뤄진 텍스트에 기존 분석기 보다 더 나은 성능을 보이는 형태소 분석기를 만들었다. 당시 해당 연구가 있었다면 전처리에 활용할 가치가 높았으나, 시기가 맞지 않아 사용하지 못했다. 해당 분석기를 사용시 예측 모델링 성능 향상에 기여될 것으로 보인다.

6.4 향후 연구 과제

더불어 보다 진보된 임베딩, 차원축소 기법과 학습 모델과의 앙상블 등을 통해 성능을 높일 수 있을 것으로 기대한다. SVD는 적은 데이터로도 잘 작동하나 문서 전체 통계에 기반하므로 단어가 등장하는 세밀한 맥락은 고려 못한다는 단점이 존재하는 데, SVD 보다 후에 등장한 Doc2vec와 같이 문서를 축약하는 기법(Dai et al. 2015)을 이용해 보다 심화된 문맥을 읽어낼

것을 기대할 수 있다.

또한 성능 좋은 OCR(Optical character recognition; OCR) 소프트웨어를 이용하면 다소 무질서하게 기록되어 있는 주석 속 숫자 데이터도 처리 대상으로 삼아 맥락과 결합하여 예측 모델링에 추가로 활용할 수 있을 것으로 기대한다. 비슷한 맥락으로 정보보호 공시(박재영·정우진 2019)와 같이 몇몇 검증된 키워드에 한정해 관련 텍스트를 같은 방식으로 추가해 모델 성능을 높이는지 확인할 때 사용 할 수 있을 것으로 본다.

본 연구에서는 주당수익률만 예측 모형으로서 활용하였는데, 향후에는 다른 재무 비율(예, 자기자본비용, price earning ratio) 등에도 본 연구의 방법론을 적용해볼 수 있을 것이다.

참고 문헌

[국내 문헌]

1. 김권중 2015. “K-IFRS 재무제표분석과 가치평가,” 서울:창민사, pp 6.
2. 김용석, 조성욱 2019. “한국어 텍스트 분석과 적용,” *한국증권학회지* (48:2), pp 215-235.
3. 김창범, 최현섭 2016. “공기업의 높은 부채비율과 회계처리방법이 투자의 효율성에 미치는 영향,” *국제회계연구* (70), pp 27-50.
4. 김태동, 이윤아, 배창현 2016. “특수관계자 거래와 기업의 공시수준 - 재무제표 주석을 중심으로 -,” *경영학연구*(45:3), pp 761-793.
5. 박재영, 정우진 2019. “기업의 정보보호 공시가 기업가치에 미치는 영향,” *지식경영연구* (20:4), pp 39-55,
6. 송동섭 1998 “주당이익과 주식수익율의 관련성에 관한 연구,” *생산성논집* (12:3), pp 361-379.
7. 송재성, 박종원, 박광주 2011. “파생상품 사용이 한국기업의 정보비대칭에 미치는 영향,” *재무관리연구* (28:1), pp 165-196.
8. 안드레아스 뮐러, 세라 가이도 2017. 파이썬 라이브러리를 활용한 머신러닝, 서울:한빛미디어, pp 322
9. 장경천, 김연권 2007. “가치주와 성장주의 투자성과분석,” 대한경영학회 학술발표대회 발표논문집, pp 193-222.
10. 전성일, 이기세 2012. “코스닥 신규상장 기업의 특성에 따른 재무분석가의 이익예측력에 관한 연구,” *지식경영연구* (13:5), pp 1-13
11. 정건용, 윤승식 2019. 재정정보 활용을 위한 텍스트 마이닝 기반 회계용어 형태소 분석기 구축. *한국정보시스템학회지:정보시스템연구* (28:4), pp

155-174.

12. 조남옥, 신경식 2016. “Bankruptcy Prediction Modeling Using Qualitative Information Based on Big Data Analytics,” *지능정보연구* (22:2), pp 33-56.
13. 조영석 2005. “스톡옵션 부여의 공시효과에 관한 연구,” *국제회계연구* (12), pp 149-163.
14. 모예린 2018. “주석 내용의 변동과 주식시장,” *이화여자대학교 대학원*
15. 최형규, 이상용 2019. “재무제표 주석의 텍스트 분석 통한 재무비율 예측연구 : 텍스트 마이닝을 활용한 예측력 향상,” *한국지능정보시스템학회 학술대회논문집*, pp 4-5.
16. 황선필, 윤재원, 김경호 2017. “K-IFRS 도입 전후의 주석공시사례분석과 전문가 평가.” *商業教育研究* (31:2), pp. 179-205.
17. 황선희 2006. “현금흐름표를 이용한 EPS 예측,” *인하대학교 일반대학원*

[국외 문헌]

1. Maeda, A. C. A., Carvalho, R. S., and Carvalho, R. N. 2017. “Evaluating the Use of Brazilian Companies’ Financial Footnotes Texts for Debt Variation Prediction,” *IEEE International Conference on Machine Learning and Applications* (16), pp. 1030-1034.
2. Bergmeir, C., and Benítez, J. M. 2012. “On the Use of Cross-Validation for Time Series Predictor Evaluation,” *Information Sciences* (191), pp 192-213.
3. Bishop, C. M. 1995. “Training with Noise is Equivalent to Tikhonov Regularization,”

- Neural Computation (7:1), pp 108-116.
4. Bushman, R. M., and A. J. Smith. 2001. "Financial Accounting Information and Corporate Governance," *Journal of Accounting and Economics*, (32:1-3), pp 237-333.
 5. Cao, Q., and Parry, M. E. 2009. "Neural Network Earnings Per Share Forecasting Models: A Comparison of Backward Propagation and the Genetic Algorithm," *Decision Support Systems* (47:1), pp 32-41.
 6. Chen, C. L., Liu, C. L., Chang, Y. C., and Tsai, H. P. 2013. "Opinion Mining for Relating Subjective Expressions and Annual Earnings in Us Financial Statements," *Journal of Information Science and Engineering*(29:4), pp 743-764.
 7. Chen, Y. J., Lin, J. A., Chen, Y. M., and Wu, J. H. 2019. "Financial Forecasting with Multivariate Adaptive Regression Splines and Queen Genetic Algorithm-Support Vector Regression," *IEEE Access* (7), pp 112931-112938.
 8. Hjorth, U., and Hjort, U. 1982. "Model Selection and Forward Validation," *Scandinavian Journal of Statistics* (9:2), 95-105.
 9. Jones, J. J. 1991. "Earnings Management During Import Relief Investigations," *Journal of Accounting Research*(29:2), pp 193-228.
 10. Karapandza, R. 2016. "Stock Returns and Future Tense Language in 10-K Reports," *Journal of Banking and Finance* (71), pp 50-61.
 11. Krizhevsky, A., Sutskever, I., and Hinton, G. E. 2012. "Imagenet Classification with Deep Convolutional Neural Networks," *Paper presented at the Proceedings of the 25th International Conference on Neural Information Processing Systems*(1), pp 1097-1105.
 12. Lai, S., and Li, H. "The Predictive Power of Quarterly Earnings Per Share Based on Time Series and Artificial Intelligence Model," *Applied Financial Economics*, Vol.16, No.18(2006), 1375-88.
 13. Lang, M. H., and Lundholm, R. J. 1996. "Corporate Disclosure Policy and Analyst Behavior," *Accounting Review*, (71:4), pp 467-492.
 14. Lee, B., Park, J. H., Kwon, L., Moon, Y. H., Shin, Y., Kim, G., and Kim, H. J. 2018. "About Relationship between Business Text Patterns and Financial Performance in Corporate Data," *Journal of Open Innovation: Technology, Market, and Complexity*(4:1).
 15. Lee, H., Surdeanu, M., MacCartney, B., and Jurafsky, D. 2014. "On the Importance of Text Analysis for Stock Price Prediction," *Proceedings of the 9th International Conference on Language Resources and Evaluation*, pp 1170-1175.
 16. Lin, M., Lucas Jr, H. C., and Shmueli, G. 2013. "Research Commentary—Too Big to Fail: Large Samples and the P-Value Problem." *Information Systems Research* (24:4), pp

- 906-917.
17. Meier, J. H., Esmatyar, W., and Frost, R. 2018. "The Predictive Power of the Sentiment of Financial Reports," *CEUR Workshop Proceedings* (2104), pp 30-44.
 18. Qiu, X. Y., Srinivasan, P., and Hu, Y. 2014. "Supervised Learning Models to Predict Firm Performance with Annual Reports: An Empirical Study," *Journal of the American Society for Information Science and Technology* (65:2), pp 400-413.
 19. Ramya, M. R. J. 2017. "A Comparison of Intelligent Soft Computing Techniques for Forecasting Earnings Per Share," *International Journal of Pure and Applied Mathematics* (114:9), pp 167-77.
 20. Rawte, V., Gupta, A., and Zaki, M. J. 2018. "Analysis of Year-over-Year Changes in Risk Factors Disclosure in 10-K Filings," *Proceedings of the 4th International Workshop on Data Science for Macro-Modeling*(a8)
 21. Trusov, R., Natekin, A., Kalaidin, P., Ovcharenko, S., Knoll, A., and Fazylova, A. 2016. "Multi-Representation Approach to Text Regression of Financial Risks," *Proceedings of Artificial Intelligence and Natural Language and Information Extraction*, pp 110-117.
 22. Verrecchia, R. E. "Discretionary Disclosure," *Journal of Accounting and Economics* (5:C), pp 179-194.
- [URL]**
1. 박선후 2016. <http://research.ibk.co.kr/research/board/finance/download/51315/1>
 2. 임근호 2018. <https://www.hankyung.com/finance/article/2018091138771>
 3. FnGuide 2012. <https://fnguidesns.tistory.com/217>
 4. Mikolov, Tomas, Kai C., Greg C., and Jeffrey D. 2013. <https://ui.adsabs.harvard.edu/abs/2013arXiv1301.3781M>.
 5. Dai, Andrew M., Christopher O., and Quoc V. L. 2015. <https://ui.adsabs.harvard.edu/abs/2015arXiv150707998D>
 6. Lakshmipathi N. 2019. <https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>

● 저 자 소 개 ●



최형규 (Hyoung-Gyu Choe)

한양대학교 에리카 캠퍼스에서 경제학과를 졸업한 후, 한양대학교 대학원에서 비즈니스 인포매틱스 석사학위를 취득하였다. 주요 관심 분야는 자연어처리, 인공지능, 기계학습 등이다.



이상용 (Sang-Yong Tom Lee)

현재 한양대학교 경영대학 교수로 재직 중이다. 서울대학교 경제학과를 졸업하고, Texas A&M University에서 박사학위를 취득하였다. 주요 관심 분야는 정보경제, 개인정보보호 (privacy) 및 보안, 소셜미디어, 정보통신정책, 기술경영 등이다. 관련 연구들을 MIS Quarterly, Management Science, Journal of Management Information Systems를 비롯한 다수의 저널에 관련 논문을 게재하고 있다.