

자연어 처리 및 기계학습을 활용한 손익계산서 자동화: 비용 계정 분류를 중심으로

전선함, 유현창
고려대학교 컴퓨터정보통신대학원
jsh8637@korea.ac.kr, yuhc@korea.ac.kr

A Technique for Automating Income Statement Using Natural Language Processing and Machine Learning: Focusing on Cost Account Classification

Seonham Jeon, Heonchang Yu
Dept. of Computer & Information Technology, Korea University

요 약

본 연구의 목적은 발생한 거래에 대해 적합한 회계 계정을 선택하는 예측 모델을 생성하는 것이다. 이를 통해 단기적으로 보조 수단으로 활용하여 회계 전표 승인에 대한 소요 시간을 단축하거나, 장기적으로 손익계산서가 일정 오차율 이내에서 자동으로 빠르게 작성됨으로써 재무 정보의 적시성을 올려주고, 기업의 실적을 나타냄에 있어 개별 담당자의 주관이 개입될 여지가 줄어든다는 면에서 재무 정보의 객관성을 올려줄 것으로 기대한다. 제안하는 모델은 비용 전표 입력의 적요를 자연어 처리하고 학습한 모델을 통해 1차(공시용) 비용 계정을 분류한다. 분류 결과를 범위로 활용하는 기계 학습을 활용하여 좀 더 세밀한 범위의 2차(관리용) 비용 계정을 분류하였다.

1. 서론

최근 빅데이터와 인공지능의 발전으로 많은 기업들이 다방면에서 프로세스 효율화, 자동화 등을 통한 비용 절감 방안을 모색하고 있다. 특히 회계 관련 직무는 효율화, 자동화에 대한 요구가 잦은 편인데, 일정 주기마다 명확한 기준을 갖고 동일한 방식으로 작업하는 업무 특성 때문에 기인한 것으로 보인다.

이는 여러 연구 결과를 통해서도 확인할 수 있는데, Frey and Osborne[1]은 회계 관련 직업을 자동화 가능성이 높은 직업군으로 분류하였으며, 전영승[2]은 인간이 행해온 단순 계정과목의 결정 혹은 분개의 작성 등 정형화된 업무부터 결산 및 마감까지 모두 인공지능에 의한 자동화 가능성이 높다고 언급하였다.

<표 1>은 일반적인 회계 업무의 흐름을 나타낸다. 주요 담당을 통해 작성으로부터 감사로 갈수록 점점 더 회계에 관한 전문적인 지식 혹은 자격을 요구하고 있음을 확인할 수 있다. 회계사들에 의해 수행되는 회계감사보다는 회계 전표 작성이나 승인 등 업무 흐름상 먼저 수행되는 작업들부터 자동화의 대상이 될 것이라 추정해볼 수 있다.

<표 1> 회계 업무 흐름

| 연번 | 업무 | 주요 담당 |
|----|----------|--------------|
| 1 | 회계 전표 작성 | 영업팀 |
| 2 | 회계 전표 승인 | 회계팀 |
| 3 | 마감 및 결산 | |
| 4 | 재무제표 작성 | |
| 5 | 회계 감사 | 외부감사인(공인회계사) |

시스템의 발전으로 회계 업무의 첫 단계인 회계 전표 작성 영역은 많은 부분 자동화되었다. 최근 개발된 회계 프로그램을 살펴보면 전표의 자동 작성 기능 등이 이미 구현되어 있는 경우도 있다. 다만, 이는 처음부터 해당 프로그램으로 장부를 작성하는 신규 기업에게는 유용할 수 있으나, 기존 기업들은 과거 이력을 바탕으로 일관성 있는 장부를 작성해야 과거 자료와 비교 가능성이 성립할 것이다. 따라서 자동화를 위해 단순히 해당 프로그램을 사용하는 것과는 다른 접근 방법이 필요할 수 있다.

재무제표 중 하나인 손익계산서는 크게 매출과 비용으로 구성되며, 보다 세부적인 내역을 파악할 수 있도록 매출과 비용 아래 계정이라는 일종의 구분자를 사용한다. 이 때 매출 및 매출원가의 경우 물류

시스템 등을 활용하여 회계 데이터로 연계되는 경우가 많으며, 비용을 입력할 때도 국세청과 연동을 통해 세금계산서의 거래업체, 거래일자, 금액 등을 자동으로 불러올 수 있다는 점을 감안하면 결국 담당자는 해당 비용의 성격을 거래 내용 등을 근거로 판단하여 비용 전표의 회계 계정을 선택하는 작업만을 수행한다고 볼 수 있다. 단, 이 때 거래의 실질에 대한 판단에 있어서 담당자의 주관이 개입될 수밖에 없는 구조이다. 이에 대한 위험과 관련하여 권대현[3] 등은 일부 비용의 부적절한 수동 분류로 인한 기업의 이익 과장 관련 사례를 소개한 바 있다.

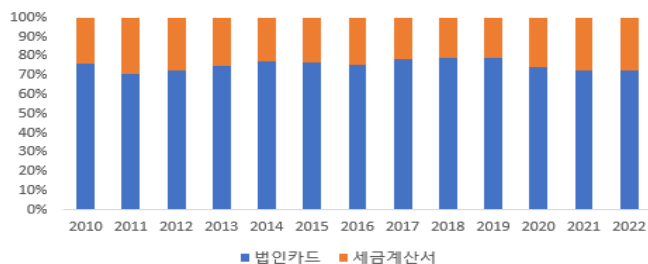
이에 본 연구에서는 나형종[4]의 연구 등을 통해 회계 분야에 대한 텍스트 분석의 효과가 강조된 이력 등을 참고하여 기존의 비용 전표가 가진 적요 및 기타 금액 등의 특징을 활용하여 비용 계정을 예측할 수 있는 모델을 제안한다. 제안하는 모델은 (그림 4)와 같이 비용 계정을 분할하고 분할한 기준마다 각기 다른 모델을 활용하여 분류 결과를 도출한다.

2. 관련 연구

CNN을 활용한 회계 계정과목 분류[5], 머신러닝을 활용한 세금 계정과목 분류[6] 등의 연구를 통해 자연어 처리를 활용한 회계 전표 관련 분류 모델 연구들이 진행되었으나 두 연구 모두 연구 대상을 설정함에 있어 법인카드 사용내역에 한정하거나 혹은 법인카드 사용내역을 포함하고 있다.

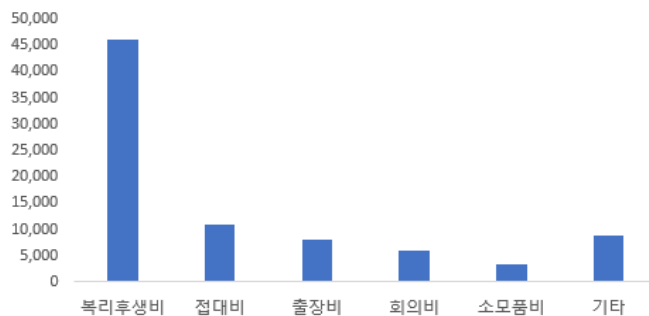
법인카드는 임직원 개인이 소지하며 식대 혹은 음료 등에 빈번히 사용되는 결제 수단이기에 비용 전표 중 많은 비중을 차지한다. 세금계산서의 경우 1억월 기준으로 거래내역을 합산하여 1건으로 거래하는 경우가 일반적이라는 특징을 감안하면 기업의 업종과 무관하게 세금계산서 대비 법인카드의 비중이 높을 수밖에 없는 구조임을 알 수 있다.

본 연구는 소비재의 생산, 유통 및 고객에 대한 판매를 주요 사업으로 하는 기업의 데이터를 활용하였다. 각 연도별 결제수단에 따른 비용 전표의 비율을 (그림 1)을 통해 확인할 수 있으며, 법인카드의 비중이 압도적으로 높다는 사실을 확인할 수 있다.



(그림 1) 결제 수단별 회계전표 비율

이에 반해 법인카드를 처리하는 계정은 복리후생비, 접대비 등으로 한정되는 경우가 많다. (그림 2)를 통해 최근 연도인 2022년의 법인카드 구성 내역을 살펴보면 대표적인 5개 정도의 항목이 대부분을 차지한다. 기타는 약 70여개의 항목이 합쳐져 있음에도 불구하고 그 비중이 크지 않음을 확인할 수 있다.



(그림 2) 법인카드 사용 건 별 처리 계정 수 (2022년 기준)

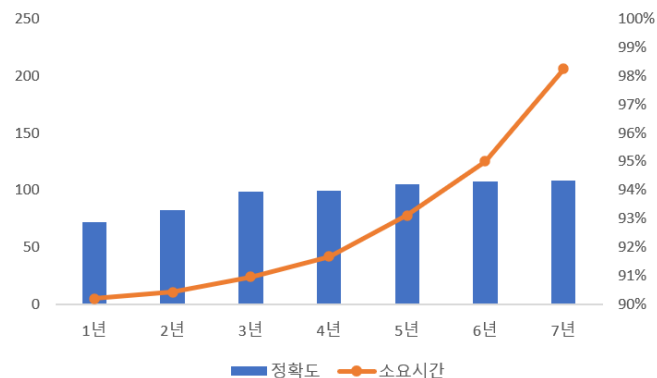
결국 연구 대상에 법인카드 사용내역이 많이 포함될수록 분류 모델의 정확도는 과대 측정되었을 가능성이 높다. 따라서 앞선 연구들의 경우 연구 대상의 설정에 대한 부분에서 한계점을 지닌다.

본 연구에서는 법인카드 사용내역을 모두 배제하고 실제 기업 운영을 위해 기업 간 거래 시 주로 사용하는 세금계산서 거래만으로 비용 계정 예측 모델을 만들었다는데 의의가 있다.

3. 연구 방법

3.1 학습 데이터의 범위 선정

공시용 비용 계정 예측을 위한 학습 데이터의 범위는 예측 데이터 기준 최근 과거 3개년으로 선정하였다. (그림 3)은 비용 전표의 적요를 형태소 분석하여 서포트 벡터 머신(Support Vector Machine)을 통해 비용 계정을 분류한 결과 값이다. 이를 통해 3년 이상의 데이터 학습 시 소요시간 증가 대비 효율이 떨어진다고 판단하였다.



(그림 3) 학습데이터 추가에 따른 학습시간 및 정확도
(좌측 축: 소요시간(분), 우측 축: 정확도)

단, 학습 대상이 되는 비용 전표 중에서 <표 2>의 내역과 사유에 해당하는 건들을 배제하였다.

<표 2> 배제 내역의 구분 및 배제 사유

| 구분 | 배제 사유 |
|-----------|---|
| 법인카드 | 건수 대비 복리후생비 등 한정된 용도로 모델의 정확도 과다 측정 |
| 예외 데이터 | 법인카드 사용이 일반적이나, 결제 문제 등으로 기타 수단으로 처리한 건 → 법인카드와 동일한 건으로 취급 |
| 자산 | 계약서에 대한 자연어 처리 및 분석이 필요, 이를 위해서 ERP 시스템 자체의 변경이 필요 → 지정된 위치에 항상 계약서가 포함되어야 함 |
| 극소 데이터 | 발생 건수가 기준 이하인 경우 자동화의 의미가 감소하며, 모델의 정확도에 악영향을 줌 (1년 간 발생 30건 이하인 경우 배제) |
| 잡이익 / 잡손실 | 사용하고 있는 계정 내 적합한 내역이 없는 기타의 개념이므로 예측이 어려움 |

3.2 형태소 분석기의 선정

연구 대상과 모델의 특성에 맞는 형태소 분석기를 선정하기 위해 <표 3>과 같이 학습연도 추가에 따른 형태소 분석기별 서포트 벡터 머신을 통한 모델에서의 분류 정확도와 소요시간을 확인하였다. 일반적으로 많이 사용되는 Mecab 대비 정확도나 속도 측면에서 크게 불리하지 않으면서 불용어 사전 구성이 용이한 Kiwi 를 사용했다. 불용어 사전은 형태소 분석을 통해 나뉜 결과를 바탕으로 잘못 인식된 단어 및 브랜드명 등 도메인 지식을 일부 활용하여 분류와 관계없는 항목들을 제외했다.

<표 3> 형태소분석기에 따른 차이내역 (시간단위: 분)

| Morphs Analyzer | 1 year | | 3 years | |
|-----------------|----------|-------|----------|-------|
| | Accuracy | Time | Accuracy | Time |
| Kiwi | 0.9255 | 4.77 | 0.9376 | 24.93 |
| Meacb | 0.9294 | 4.21 | 0.9387 | 23.60 |
| Okt | 0.9285 | 5.42 | 0.9395 | 26.07 |
| Kkma | 0.9297 | 14.18 | 0.9395 | 44.38 |
| Komoran | 0.9142 | 4.30 | 0.9207 | 24.54 |
| Hannanum | 0.8105 | 7.67 | 0.8345 | 47.62 |

3.3 예측 라벨의 정의 및 분할

비용 계정은 모델을 통해 예측하고자 하는 라벨에 해당한다. 본 연구에서는 모델의 분류 정확도를 높이기 위해 (그림 4)와 같이 공시용, 관리용 비용 계정으로 임의의 기준을 정의하였다. 공시용 비용 계정은 기업의 실적을 외부에 공시하는데 필요한 수준의 비용 계정이며 관리용 비용 계정은 예산, 실적 등의 관리를 위해 공시용을 세분화한 것으로 이해할 수 있다.

공시용 비용 계정

광고선전비-모델비

관리용 비용 계정

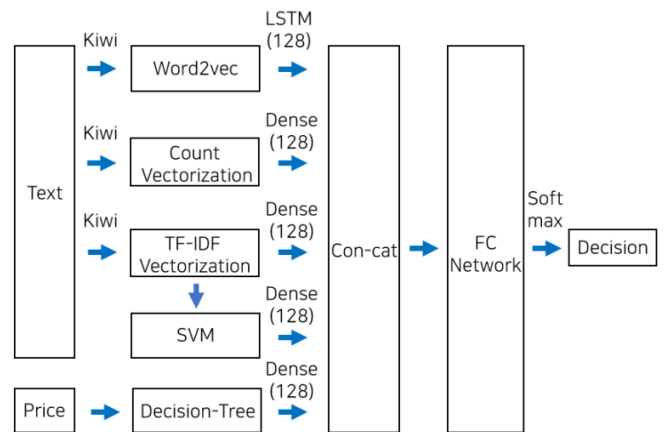
(그림 4) 공시용과 관리용 비용 계정의 구분

3.4 분류 모델의 선정

공시용 비용 계정을 예측하기 위한 모델은 서포트 벡터 머신(Support Vector Machine)을 선정하였다. 엑스트라 트리(Extra Tree) 등이 근소한 차이를 보였으나 서포트 벡터 머신의 분류 정확도가 가장 높았다.

관리용 비용 계정을 예측하기 위해 신경망 모델(Neural Network)을 사용하였으며 각 벡터화(Vectorization)의 특징을 반영하고자 대표적인 벡터화 방법인 Word2vec, Count, TF-IDF(Term-Frequency-Inverse Document Frequency)를 모두 입력 값으로 활용하였다. Word2vec의 경우 한국어 위키피디아 등을 학습한 공개 모델[7]을 활용했으며, Word2vec의 결과 값을 다른 결과 값과 합치는 과정에서 텍스트 분류 연구에 많이 사용되는 장기 단기 기억(Long Short-Term Memory) 모델을 사용했다.

공시용 비용계정이 관리용 비용계정의 상위 레벨이라는 측면에서 서포트 벡터 머신을 통한 출력 값을, 금액 대를 통한 추정도 가능하다는 측면에서 결정나무모형(Decision-Tree)의 출력 값을 신경망 모델의 입력 값으로 함께 활용하였다. 이를 간단하게 표현하면 (그림 5)와 같다.



(그림 5) 단계별 계정 분류 Model 구조도

4. 실험결과

4.1 실험 환경

본 연구의 실험은 Google Colaboratory 서비스를 통해 Intel(R) Xeon(R) CPU @ 2.20GHz, RAM 12GB를 사용하여 진행했다. 각 모델은 전처리 이전 기준으로 예측연도의 라벨 20,376 건의 예측을 위해 과거 3개년의 라벨 56,615 건을 학습하였다.

4.2 공시용 비용 계정(Level 1)의 예측 결과

예측연도의 공시용 비용 계정은 총 14 개이다. 관리용보다 분류 선택지가 적은 만큼 상대적으로 정확도가 높다.

<표 4>는 서포트 벡터 머신을 통해 예측한 공시용 회계 계정 별 확률을 내림차순 했을 때, 몇 번째 순위 안에 정답이 포함되는지에 대한 확률을 누적 값으로 표현한 것이다.

<표 4> 정답 포함 범위에 따른 Accuracy

| 정답 포함 범위 | Accuracy |
|----------|----------|
| 1 순위만 | 0.9438 |
| 2 순위까지 | 0.9806 |
| 3 순위까지 | 0.9908 |

14 개의 계정 중 3 번째로 높은 확률로 예상되는 비용 계정 안에 99% 이상의 회계 전표의 계정에 대한 정답이 포함됨을 확인할 수 있다.

4.3 관리용 비용 계정(Level 2)의 예측 결과

<표 5>를 통해 모델 별 조합에 따른 예측 정확도를 비교해볼 수 있다.

<표 5> 모델 별 관리용 회계 계정 예측 Accuracy

| Model | Accuracy |
|--|----------|
| SVM (Level 2) | 0.7692 |
| CNN (Level 2) | 0.7420 |
| SVM (Level 1) + LSTM, CNN (Level 2) | 0.7954 |
| SVM (Level 1) + LSTM (Level 2) | 0.8033 |

최종 모델에 관한 실험 세부 내역은 <표 6>을 통해 확인할 수 있다. Con-cat 을 위한 과정에는 ReLU 를 활성화 함수로, 최종 분류 결과 값을 도출하는 과정에서는 Softmax 를 활성화 함수로 사용했다.

<표 6> Parameter Configuration

| Parameter name | Size |
|-----------------|-------|
| Num of Classes | 111 |
| Embedding dim | 100 |
| Sequence Length | 30 |
| Hidden dim | 128 |
| Learning Rate | 0.001 |
| Batch Size | 32 |
| Num of Epochs | 10 |

5. 결론 및 향후연구

본 연구에서는 자연어 처리 및 기계학습을 통하여 과거 3 년 간의 데이터를 학습한 비용 계정 예측모델을 생성하였고, 이를 통해 2022 년도 1 년 간의 데이터를 예측하여 약 80.3%의 정확도를 달성하였다.

해당 연구는 담당자들이 입력하는 비용 전표가 자산인지 비용인지 구분할 수 있는 기준을 제시하고 있지 못하다는 점에서 한계를 지닌다.

다만 모델을 통해 담당자의 주관이 포함되지 않은 일정 수준 이상의 재무제표를 만들어낼 수 있다는 점에서 의의를 가진다. 이는 곧 손익계산서 등의 자동화에 대한 가능성이 충분히 높아졌다는 의미도 될 수 있다.

추후 연구에서는 딥러닝 모델 구조 개선 및 트랜스포머(Transformer) 등 고비용 모델을 통한 실험을 진행할 예정이다. 또한 다양한 연차의 회계 담당자 및 비 회계 담당자 등을 대상으로 모델의 입력 값과 동일한 정보를 기반으로 실험한 예측 정확도를 수집하여 약 80%의 예측 결과가 사람과 비교했을 때 어느 정도 수준에 해당하는지 비교해볼 계획이다.

참고문헌

- [1] Frey and Osborn, The future of employment: How susceptible are jobs to computerization?, In Technological Forecasting & Social Change, pp. 254-280, 2013
- [2] 전영승, 인공지능이 회계실무에 미치는 영향, 글로벌경영학회지 14(2), 47, 2017, pp. 311-333
- [3] 권대현, 안태식, 황인이, 박진하, 결산 자동화 시스템 사례, 기업가정신과 벤처연구 20(1), 2017 pp. 47-57
- [4] 나형중, The Impact of Text Analysis on Accounting and Finance Research, 지역산업연구 44(2), pp.53-74, 2021
- [5] 김병수, CNN(Convolution Neural Network)을 활용한 회계 계정과목 분류, 국내석사학위논문 고려대학교 대학원, 2020
- [6] 최동빈, 조인수, 박용범, 머신러닝을 이용한 세금 계정과목 분류, 반도체디스플레이기술학회지 제 18 권 제 2 호, 2019, pp. 73-77
- [7] <https://ratsgo.github.io/embedding>