

A Study on Predicting Credit Ratings of Korean Companies using TabNet

Hyeokjin Choi*, Gyeongho Jung*, Hyunchul Ahn*

*Master's Candidate, Graduate School of Business IT, Kookmin University, Seoul, Korea

*Master's Candidate, Graduate School of Business IT, Kookmin University, Seoul, Korea

*Professor, Graduate School of Business IT, Kookmin University, Seoul, Korea

[Abstract]

This study presents TabNet, a novel deep learning method, to enhance corporate credit rating accuracy amidst growing financial market uncertainties due to technological advancements. By analyzing data from major Korean stock markets, the research constructs a credit rating prediction model using TabNet. Comparing it with traditional machine learning, TabNet proves superior, achieving a Precision of 0.884 and an F1 score of 0.895. It notably reduces misclassification of high-risk companies as low-risk, emphasizing its potential as a vital tool for financial institutions in credit risk management and decision-making.

▶ **Key words:** Corporate Credit Rating Prediction, TabNet, Machine Learning, Deep Learning, Credit Risk Management

[요 약]

최근 IT 기술의 발전과 더불어 금융 시장에서의 불확실성이 증대되는 상황에서 기업 신용등급 평가의 중요성을 인식하고, 이를 개선하기 위한 새로운 접근 방식으로 딥러닝 모델인 TabNet을 제안한다. 이에 본 연구에서는 TabNet을 활용하여 기업 신용등급을 예측하고, 이의 예측 성능을 기존 머신러닝 방법론과 상세하게 비교한다. 한국의 주요 증권시장에 상장된 기업들의 재무 데이터를 기반으로 TabNet 알고리즘을 적용하여 신용등급 예측 모델을 구축하고, 다양한 머신러닝 모델과의 성능을 비교 분석하였다. 실험 결과, TabNet 모델은 Precision 0.884, F1이 0.895로 기존의 머신러닝 모델들보다 우수한 성능을 보였으며, 고위험 기업을 저위험 기업으로 잘못 분류하는 경우가 다른 머신러닝 모델보다 적어 TabNet의 우수성을 확인하였다. 이는 TabNet이 기업 신용등급 예측에 있어 효과적인 도구로 활용될 수 있으며, 금융기관의 신용 위험 관리 및 의사 결정 과정을 지원할 수 있을 것으로 기대한다.

▶ **주제어:** 기업신용등급 예측, TabNet, 머신러닝, 딥러닝, 신용 위험 관리

- First Author: Hyeokjin Choi, Corresponding Author: Gyeongho Jung
- *Hyeokjin Choi (hyeokjin0350@naver.com), Graduate School of Business IT, Kookmin University
- *Gyeongho Jung (gyeongho9878@kookmin.ac.kr), Graduate School of Business IT, Kookmin University
- *Hyunchul Ahn (hcahn@kookmin.ac.kr), Graduate School of Business IT, Kookmin University
- Received: 2024. 03. 15, Revised: 2024. 04. 30, Accepted: 2024. 05. 08.

I. Introduction

금융기관과 금융투자회사들에게 기업의 신용등급을 평가하고 측정하는 것은 매우 중요한 일이다. 특히 지난 1997년 IMF 외환 위기와 2007년 서브프라임 모기지 사태(subprime mortgage crisis), 2020년 COVID-19 발생에 따른 경제 위기 등과 같은 금융 위기가 발생하면서 기업의 부도예측과 신용위험관리에 대한 중요성이 더욱 부각되고 있다. 또한 최근 갈수록 심화되고 있는 글로벌 경제의 불확실성 확대와 더불어 국내 경기 침체에 따른 금융시장의 불안정성이 대두되면서 보다 정확한 모델을 통한 기업 신용평가 시스템이 필요해졌다. 학계에서도 이와 같은 이유로 전통적 통계 방법론인 로지스틱 회귀분석과 머신러닝 기법인 의사결정나무 기반의 모델을 적용한 기업 신용평가 모델 연구를 지난 수년간 활발하게 수행하여 왔다[1].

기업채권에 대한 신용등급 예측모형 연구가 세계적으로 활발히 진행되고 있는 가운데, 최근 IT기술의 발달로 인간의 지성과 지적능력을 컴퓨터에 도입하고자 하는 인공지능 기술이 각계각층 다양한 분야에서 뛰어난 성능을 보여 주고 있다. 특히 심층신경망(deep neural network)을 학습하여 결과를 도출하는 딥러닝 기반 예측 알고리즘이 통계적인 분석기법과 머신러닝 기반 예측 알고리즘보다 회귀와 분류 문제 있어 우수한 성능을 보여주고 있다. 딥러닝 알고리즘은 복잡한 데이터 형태를 최소한의 전처리 과정과 변수 선택만으로도 효과적으로 학습할 수 있다는 장점이 있어 금융산업을 포함한 다양한 사회 과학 분야에서 부도 예측 모델과 신용평가 모형 등에 딥러닝 알고리즘을 도입하는 추세이다[2]. 따라서, 본 연구는 금융 분야에서 딥러닝 알고리즘의 적용이 확산되는 현상을 반영하여, 특히 TabNet과 같은 딥러닝 모델이 기업 신용등급 예측에 있어 어떠한 성능을 보이는지 평가함으로써, 이 분야의 발전 가능성을 탐색하고자 한다.

이러한 배경 하에, 본 연구는 TabNet이라는 새로운 딥러닝 기술의 특성과 이를 통한 예측의 효율성에 대해 깊이 있게 분석하고자 한다. Arik and Pfister(2021)이 고안한 새로운 딥러닝 기술인 Tabnet 알고리즘은 순차적인 어텐션(Sequential Attention)을 적용하여 각각의 의사결정 단계에서 순차적으로 Feature를 선택하여 보다 효율적인 학습과 예측 결과에 대한 해석이 가능한 딥러닝 모형이다[3]. 본 연구는 정교한 기업신용등급 평가를 위해 TabNet 기반의 다중 분류 모델을 적용함으로써, TabNet이 신용등급 예측에 있어 정확성과 안정성을 양립할 수 있는 유효한 방법임을 검증하고자 한다. 더불어, 본 연구는 기존의

머신러닝 모델들과의 성능 비교를 통해 TabNet의 상대적인 우수성을 평가할 것이다.

II. Preliminaries

1. Corporate credit rating assessment

기업신용등급 평가는 기업의 채무 이행능력과 신용 위험을 평가하는 프로세스로, 주로 재무적인 관점에서 이루어진다. 이는 기업의 채무 수준을 이해하고 부채 발행이나 대출 시 신뢰성을 확보하는 데 중요한 역할을 한다. 이러한 평가는 주로 재무상태 분석, 신용평가 모형, 그리고 시장 정보를 통해 이루어진다. 재무상태 분석은 기업의 재무 건전성과 지급 능력을 평가하는 과정으로, 주요 지표로는 부채비율, 이익률, 유동성 등이 사용된다. 신용평가 모형은 통계적 기법이나 머신러닝 알고리즘을 활용하여 기업의 신용 위험을 평가하며, 시장 정보는 기업체의 시장가격이나 신용스프레드 등을 분석하여 신용평가에 반영된다[1-4].

국내외 주요 신용평가 기관으로는 한국신용평가, 한국기업평가, NICE 신용평가 등이 있으며, 국제적으로는 무디스, S&P, 피치 등이 있다. 이들 기관은 각자의 평가 기준과 방법을 가지고 있으며, 글로벌 금융시장에서 큰 영향력을 가지고 있다. 이들 신용평가 기관은 각자의 등급 체계를 가지고 있는데, 무디스는 Aaa부터 Baa3까지의 등급을 사용하고 한국신용평가는 AAA부터 BBB-까지의 등급을 사용한다. 이러한 등급은 기업의 신용 위험 수준을 표시하는 데 사용된다. 기업신용등급평가는 기업의 신용 위험을 평가하는 중요한 프로세스로, 재무적인 관점에서 기업의 재무 건전성과 지급 능력을 평가한다. 이를 통해 부채 발행이나 대출 시 신뢰성을 보장하고 금융 시장에서의 신용등급을 확보하는 데 중요한 역할을 한다[5].

한편 IT 기술의 발전으로 빅데이터와 인공지능 기술을 결합한 기업신용평가 시스템이 등장하였다. 기존의 수작업 방식으로 이루어지던 신용평가는 시간이 많이 소요되고 오류 발생 가능성도 높았으나, AI의 도입으로 신용평가 모델의 안정성을 유지하면서 동시에 정밀한 예측이 가능해졌다. 또한, 기업의 SNS, 뉴스 기사, 고객 피드백 등 다양한 비정형 데이터를 분석하여 보다 광범위한 관점에서 신용도를 평가할 수 있다. 이로 인해 기업의 숨겨진 가치나 위험 요소를 보다 정확하게 파악하는 것이 가능해졌다[6]. 이에 따라 AI 기반 기업 신용평가는 기업의 신용도 판단에 있어 중요한 역할을 하고 있으며, 앞으로도 계속될 것으로 보인다.

2. Corporate credit rating using AI/ML

기계학습(Machine Learning)은 AI의 한 분야로서 데이터를 바탕으로 내재된 패턴을 탐색하고 분석하여 결과를 예측하는 방법론을 의미한다. 기계학습을 활용한 기업 신용등급 평가 분야의 초창기 연구들은 OLS(Ordinary Least Squares), 다중판별분석 등과 같은 통계적 방법론에 기반한 신용등급평가모형을 연구해 왔으나, 낮은 예측 정확도 문제로 인해 1980년 이후부터 인공신경망, 귀납적 학습방법, 사례기반추론, 유전자 알고리즘, 다분류 서포트 벡터 머신(support vector machine, SVM) 등 인공지능 기법이 주로 적용되고 있다. 이 중에서 최근까지도 많이 연구되고 있고, 현장에서 활용되고 있는 방법으로 인공신경망과 다분류 SVM이 있다[7].

초기 연구인 Shin and Han (1999)에서는 유전자 알고리즘(GA)을 기반으로 한 사례 기반 추론(CBR)을 통해 회사채 신용등급 예측 모델을 개발, 75.5%의 예측 정확도를 달성했다고 보고했다. 이는 회사채 신용등급 예측을 위한 머신러닝 기법의 적용 가능성을 시사한다[8]. 서포트 벡터 머신(SVM)과 인공 신경망(ANN)의 성능을 비교한 Huang et al. (2004)의 연구에서는 SVM이 일반적으로 ANN보다 우수한 성능을 보였음을 발견했다. 이 결과는 SVM이 회사채 신용등급 예측에 있어 더 효과적인 기법일 수 있음을 시사한다[9]. 또한, 인공 신경망과 선형 판별 분석(LDA)을 비교한 Kumar and Bhattacharya (2006)의 연구에서는 인공 신경망이 복잡한 데이터와 비선형 문제 해결에서 LDA보다 우수한 성능을 보였다. 이는 인공 신경망이 더 복잡한 신용등급 예측 모델에 적합할 수 있음을 제시한다[10]. 다중 클래스 분류 문제에 대한 연구에서 Ye et al. (2008)은 PSVM(proximal SVM)이 예측 정확도 84%로 우수한 성능을 보였으며, 연산 시간도 줄일 수 있음을 보고했다. 이는 기업 신용등급 다중 클래스 분류에 PSVM의 적용 가능성을 강조한다[11]. 마지막으로, Huang (2009)의 연구에서는 차원 축소 기반 통합 SVM 알고리즘을 통해 기업 신용평점 예측의 성능을 향상시킬 수 있었다. KGE(kernel graph embedding) 방식을 활용한 SVM은 다른 멀티 클래스 SVM이나 기존 분류 방식에 비해 우수한 성능을 나타냈다. 이러한 연구 결과들은 회사채 신용등급 예측을 위한 머신러닝 기법의 발전 방향성을 제시한다[12].

2010년 이후 기업 신용등급 예측을 위한 머신러닝 기술의 발전과 하이브리드 방식의 적용에 초점을 맞춘다. Guo et al. (2012)의 연구에서는 퍼지 군집화와 SVM을 결합한 신용 예측 모델이 제안되어, 다중 분류 문제에 있어서 효율적인 데이터 전처리와 경계 데이터 포인트의 선택을 통

해 연산 과정을 줄이면서도 약 73%의 예측 정확도를 달성했다고 보고되었다[13]. 또한, 기업 신용등급의 다단계 예측 모델 구현에 대한 연구도 진행되었다. 여기서는 데이터 전처리 과정에서의 특성 선정, 데이터 군집화 및 리샘플링을 통해 첫 단계를 완성하고, 다음 단계에서 다양한 분류 기법 및 앙상블 기법을 적용하여 최종 예측 모델을 구현한다. Wu et al. (2014)에 의하면, 데이터 리샘플링을 적용한 Bagging-DT 방식이 82.96%의 예측 정확도로 가장 우수한 결과를 보였다[14].

한편, 김성진과 안현철 (2016)의 연구에서는 랜덤 포레스트 분석 기법이 기존 머신러닝 분석기법에 비해 과적합에 덜 취약하고, 데이터의 잡음이나 이상치를 효과적으로 처리할 수 있다고 주장한다. 이 연구는 1,295개의 국내 상장 기업 데이터를 사용하여 최대 75%의 예측 정확도를 달성한 랜덤 포레스트 기반 예측모델을 개발함으로써, 랜덤 포레스트의 유효성을 입증한다[15]. 이외에도 박형권 et al (2018)의 연구에 따르면, 신용등급 예측은 다중 클래스 분류 문제로서, 기업 신용등급의 다양성과 불균형 분포를 고려할 때 비교적 높은 성능의 학습 모델이 요구된다. 이에 따라, 일반적으로 5개 내외의 클래스로 등급을 통합하는 방식이 적용되며, 이는 분석의 복잡성을 줄이고 예측 모델의 성능을 최적화하는 데 기여한다[16]. 이후 2020년대 들어, 기업 신용등급 예측 연구에서는 비구조화된 데이터 집합의 활용이 증가하였으며, 이미지 분야에서는 CNN(Convolutional Neural Network) 알고리즘이 좋은 성과를 보이고, 시계열 자료의 예측이나 텍스트 자료의 문맥 인식에는 RNN(Recurrent Neural Network) 알고리즘이 사용되고 있다[17].

III. Research Model

본 연구에서는 기업신용등급평가 예측을 위하여 TabNet을 제안한다[3]. TabNet은 정형 데이터 분석을 위한 딥러닝 기반의 엔드 투 엔드(end to end) 모델로, 기존 부스팅 기반 머신러닝 모델에 비교하여 동등하거나 우수한 성능을 제공함으로써 정형 데이터 처리에 있어 딥러닝의 적용 가능성을 확장하고 있다. TabNet은 예측력을 강화하기 위해 핵심적인 특징(Feature)들을 동적으로 선택하는 구조를 갖추고 있다. 이 과정에서, 모델은 입력 데이터를 초기 단계에서 배치 정규화(Batch Normalization, BN)를 통해 처리함으로써 중요한 역할을 한다. 배치 정규화는 입력 데이터의 평균을 0으로, 분산을 1로 조정하여, 레이어 간 입력 분포가

변화하는 문제, 즉 내부 공변량 변화(Internal Covariate Shift)를 줄인다. 이 초기 단계는 TabNet이 중요한 특성을 효과적으로 식별하고 활용할 수 있는 기반을 마련해 주며, 결과적으로 예측 정확성을 높이는 데 기여한다[18].

모델의 핵심은 GLU(Gated Covariant Unit)블록을 사용하여 데이터의 복잡한 특성 간의 상호작용을 모델링하는 것이다. GLU블록은 입력에 대해 두 개의 경로를 사용하는 활성화 함수이다[19]. 하나는 선형 변환을 사용하고 하나는 게이트의 역할을 하는 시그모이드 함수를 사용하여 중요한 정보를 선택적으로 강조하고, 덜 중요한 정보는 억제할 수 있도록 한다. 이러한 변환 과정을 통해, GLU 블록들은 데이터의 다층적인 표현을 학습하고, 모델이 입력 데이터의 복잡한 패턴을 효과적으로 이해할 수 있도록 한다. TabNet은 이러한 GLU 블록을 독립적으로 또는 공유되게 사용하여, 모델의 일반화 능력을 향상시키고 다양한 데이터 구조에 대한 적응성을 증진한다.

그림 1과 같이 TabNet의 인코더는 다양한 의사결정 단계(Decision Steps)로 구성되어 있으며, 이는 주로 인코딩을 담당하는 변수 트랜스포머(Feature Transformer)와 마스크 생성을 위한 어텐티브 트랜스포머(Attentive Transformer)로 구성된다.

변수 트랜스포머는 그림 2와 같이 FC(Fully Connected Layer), BN(Batch Normalize), GLU의 조합으로 구성된다. 이 구조는 데이터를 고차원으로 매핑하는 FC, 학습 과정을 안정화시키는 BN, 그리고 중요한 정보를 선택적으로 강조하는 GLU로 이루어진다. 변수 트랜스포머 내에서, FC-BN-GLU의 조합은 총 네 번 반복되며, 이 중 앞의 두 번은 의사결정 단계 간에 공유되어 가중치를 공유함으로써 일관된 특성 추출을 가능하게 한다. 이 공유 메커니즘은 데이터의 공통적인 및 기본적인 정보를 효율적으로 파악하게 하며, 뒤의 두 번은 보다 세밀한 특성을 추출하여 패턴 인식 능력을 강화한다. 이러한 구조는 모델이 필요로 하는 파라미터 수를 줄이면서도, 효과적인 특성 조합 및 표현 학습을 가능하게 하여, TabNet의 경량화 및 효율성을 도모한다.

끝으로 그림 3은 어텐티브 트랜스포머를 보여준다. 어텐티브 트랜스포머(Attentive Transformer)는 변수 트랜스포머와 연결되어 FC-GLU-Sparsemax를 거쳐 마스크(Mask)를 생성한다[20]. 이 과정에서 변수 트랜스포머의 출력은 희소(Sparse)한 분포로 변환되어 중요한 특징만을 선택하도록 한다. 이때 이전 변수 척도(Prior Scales)를 활용하여 데이터의 다양한 측면을 고려하고, 중복되는 특징 선택을 방지한다. 생성된 마스크는 변수 트랜스포머에 전

달되어 어텐티브 트랜스포머에서 가중치가 할당된 특성은 강조되고 그렇지 않은 특성은 억제된다. 이 과정은 각 의사결정 단계에서 반복되며, TabNet은 이러한 메커니즘을 통해 입력 데이터를 바탕으로 기업신용등급을 예측한다.

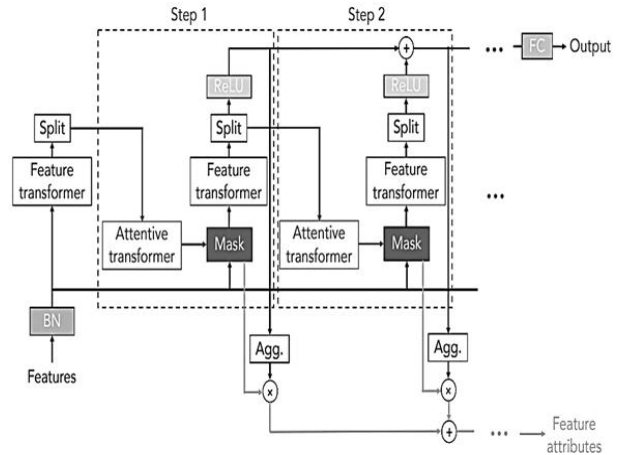


Fig. 1. TabNet encoder architecture

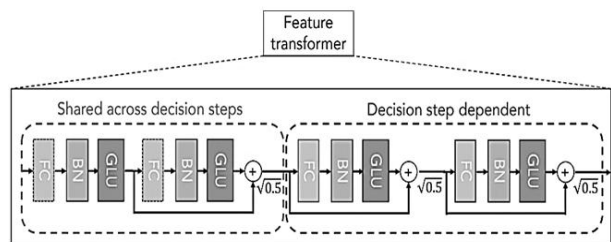


Fig. 2. Feature transformer block

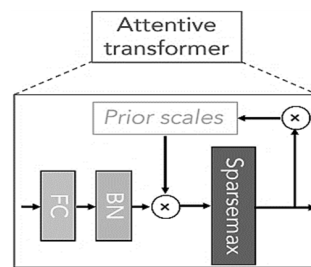


Fig. 3. Attentive transformer block

IV. Empirical Test

1. Proposed credit rating model

본 연구에서 제안하는 전체적인 프로세스는 그림 4와 같다. 신용등급평가와 분석을 위한 데이터 수집을 위해 풍부한 데이터를 수집하였다. 이후 모델링 단계에서는 데이터를 훈련 세트와 테스트 세트로 8:2 비율로 나누고 학습 모델을 구축하였다. 훈련 데이터셋에 대해서는 오버샘플링

과 계층화 5-폴드 교차 검증(Stratified 5-Fold Cross Validation)을 적용하였다. 이후 다수의 모델과 비교실험을 진행하고, 결과를 분석하여 학습모델의 우수성을 검증하였다.

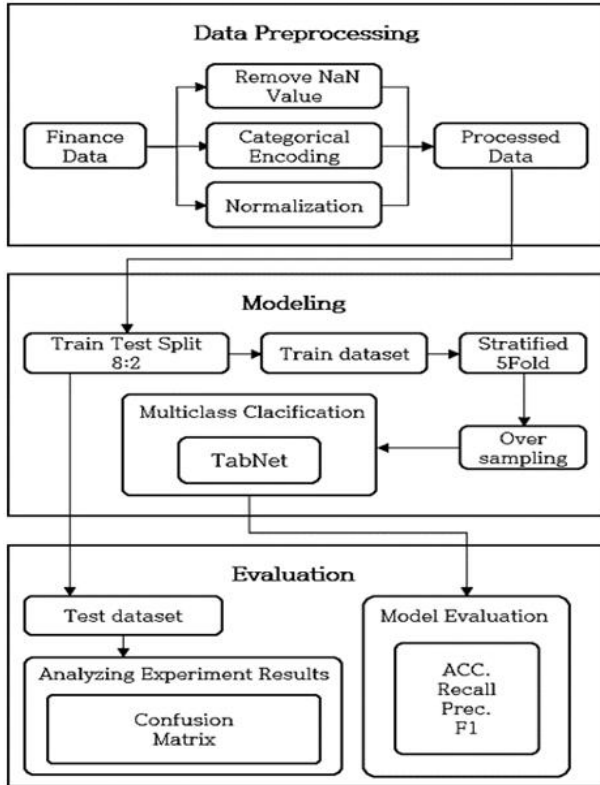


Fig. 4. Proposed Model

2. Data Collection and Preprocessing

본 연구에서 활용된 데이터는 한국의 주요 증권시장인 KOSPI, KOSDAQ, KONEX에 상장된 제조업, 은행업, 보험업을 포함하는 총 1,517개 기업에서 추출하였다. 이 데이터는 2002년부터 2022년까지 20년간 공시된 정보를 기반으로 하며, 11,890건의 기업 신용등급 자료를 포함한다. 이는 S&P, Moody's, Fitch, 한국기업평가, 한국신용평가, NICE신용평가, 서울신용평가 등 국내외 주요 신용평가기관으로부터 수집된 것이다.

표1을 참조하면, 본 연구에서 수집된 데이터는 AAA부터 D까지 총 22개의 신용등급으로 구분되어 있으며, 상대적으로 수집 건수가 적은 D등급 데이터는 C등급과 통합하여 분석에 활용하였다. 이 연구는 다양한 업종 및 신용등급을 포함하는 대규모 데이터셋을 기반으로 하여 TabNet 알고리즘의 성능을 종합적으로 평가하고자 한다.

연구에 사용된 입력 변수는 산업분류, 성장성, 수익성, 안정성, 활동성, 생산성, 부가가치, 투자지표, EBITDA 지

표를 포함하는 194개의 변수이다. 이는 TabNet 알고리즘의 특징 선택의 장점을 최대한 활용하여, 알고리즘의 예측 성능 및 신용등급 분류 능력을 검증하기 위함이다.

데이터 전처리 과정은 알고리즘의 광범위한 예측 능력을 확인하기 위해 최소한으로 제한하였으며, 이는 산업 분류의 인코딩, 결측값(NaN)의 제거, 정규화 과정이 포함된다.

산업 분류에 대해서는 One-Hot Encoding 방법을 적용하여, 각 산업 분류를 해당하는 이진 변수로 전환함으로써 모델이 각 산업을 명확하게 구별할 수 있도록 하였다. 결측값은 재무 정보의 부재를 나타내므로, 이를 0으로 대체하여 처리하였다. 정규화의 경우, Z-score 정규화 방식을 적용하여 모든 수치형 변수가 동일한 척도로 비교될 수 있도록 하였다.

Table 1. Credit ratings of Korea data

Label	Grade	Count	Ratio(%)
0	AAA	93	0.782
1	AA+	107	0.899
2	AA	368	3.095
3	AA-	516	4.340
4	A+	1477	12.42
5	A	1523	12.81
6	A-	1642	13.81
7	BBB+	1009	8.486
8	BBB	1033	8.688
9	BBB-	874	7.351
10	BB+	801	6.737
11	BB	645	5.425
12	BB-	557	4.685
13	B+	430	3.616
14	B	383	3.221
15	B-	240	2.019
16	CCC+	31	0.261
17	CCC	70	0.589
18	CCC-	5	0.042
19	CC	14	0.118
20	C, D	72	0.606
Total		11,890	100.0

3. Experimental Model

본 연구의 모형 구축 과정에서, 전체 데이터 세트의 약 80%를 훈련용으로 할당하고 나머지 20%를 검증용으로 설정하였다. 신용등급 데이터의 균일한 분포를 보장하기 위하여, 계층화 5-폴드 교차 검증을 적용하였다. 이 방식은 각 교차 검증 폴드에 신용등급이 균등하게 배치되도록 하여, 모델이 신용등급의 전 범위에 걸쳐 일관된 학습 및 검증 과정을 거치도록 구성하였다.

데이터 세트에 존재하는 신용등급의 불균형 분포는 특정 등급에 대한 모델의 학습이 과도하게 집중될 위험을 내포하

고 있으며, 이는 모델의 종합적인 성능 저하를 초래할 수 있다. 이러한 위험을 완화하기 위해, 본 연구에서는 소수 클래스의 샘플을 인위적으로 증가시키는 SMOTE(Synthetic Minority Oversampling Technique) 방법을 도입하였다. SMOTE 기법을 통해, 모델이 다양한 신용등급에 대해 보다 균형 잡힌 학습을 수행할 수 있도록 하여, 모든 신용등급에 대한 예측 능력을 향상시키도록 한다[21].

Table 2. TabNet Hyperparameters Description

n_d	14
n_a	14
n_steps	1
tabnet_gamma	2.0
n_shared	5
n_independent	1
lambda_sparse	0.000035
optimizer	Adagrad
lr	0.11
scheduler	StepLR
scheduler_step_size	15
scheduler_gamma	0.99
mask_type	entmax
loss	FocalLoss
loss_gamma	0.4

본 연구에서 사용된 TabNet의 학습에 사용된 하이퍼파라미터(Hyperparameter)는 표 2와 같으며 TabNet의 기본 설정과 다르게 설정된 항목에 대해서만 명시하였다. 주요한 특징으로는 스케줄러(scheduler)의 경우 세밀한 조정이 가능한 StepLR 방식을 채택하였다. 이는 학습 과정에서의 학습률을 주기적으로 조정하여, 모델의 수렴 속도와 최종 성능을 최적화한다. 손실 함수로는 클래스 불균형 문제에 효과적으로 대응하기 위해 FocalLoss를 사용하였다. 이는 오분류된 클래스에 더 큰 가중치를 부여하여, 특히 소수 클래스의 예측 정확도를 향상시키는 데 중점을 두었다. 이러한 접근 방식은 기업신용등급 평가와 같이 클래스 불균형이 심한 문제에 있어서 모델의 성능을 개선하는데 중요한 역할을 한다.

4. Experimental Results

본 연구에서는 기업 신용 등급 예측을 위한 TabNet 모델의 성능을 평가하고자 하였다. 이를 위해 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), 및 F1 점수를 포함한 다양한 평가 지표를 사용하여 TabNet을 포함한 5가지 머신러닝 알고리즘의 성능을 비교 분석하였다. 기업 신용 등급 데이터의 불균형한 특성을 고려하여, 모델의 종합적인 성능 평가를 위해 이러한 지표들을 채택하였다. 또한 모든 모델의 성능을 보다 공정하고 일관된 방식

으로 비교하기 위해, 각 폴드가 전체 데이터셋의 클래스 비율을 반영하도록 보장하는 계층화 5-폴드 교차 검증을 적용하였으며, 실험의 공정성을 확보하기 위해 모든 모델에 동일한 변수를 가지는 데이터로 동일한 전처리 과정을 적용함으로써 알고리즘 자체의 특성을 평가하고자 하였다.

표 3에서 TabNet은 모든 지표에서 뛰어난 일관성을 보였으며, 평균 정확도는 0.983, 재현율은 0.911, 정밀도는 0.884, 그리고 F1 점수는 0.895로 나타났다. 이는 TabNet의 특징 선택 능력과 순차적인 의사 결정 과정이 데이터의 중요한 패턴을 효과적으로 포착하는 동시에 Focal Loss가 소수의 클래스에 대하여 예측 성능을 높여 좋은 성과를 낼 수 있음을 확인하였다. 반면 XGB의 경우, 각 폴드에서의 성능 지표는 눈에 띄게 변동을 보인다. 정확도의 경우 최저 0.823이며, 최대 0.944를 보여주는 것을 보아, 데이터 분할에 따라 상이한 성능을 나타낼 수 있음을 확인하였다. 또한, 평균 정확도는 0.884로, 모델의 전반적인 성능이 안정적이라고 할 수 있으나, 폴드 간 성능의 변동성은 모델의 일관성 있는 예측력에 영향을 줄 수 있는 중요한 요소로 작용할 수 있음을 확인하였다. 이로 인해 XGB가 특정 데이터 세트에 민감하게 반응할 수 있음을 확인하였다.

KNN의 경우 각 폴드에서의 성능 차이가 나타나는 것을 확인할 수 있다. 평균 정확도는 0.915, 재현율은 0.915, 정밀도는 0.912, 그리고 F1 점수는 0.912로 좋은 성능을 보여주는 것을 확인하였다. CatBoost는 성능 지표에서 각 폴드 별 성능의 차이가 크다는 것을 확인하였다. 평균 정확도는 0.864, 재현율은 0.864, 정밀도는 0.861, 그리고 F1 점수는 0.861을 보여준다. 그리고 LGBM의 경우 본 연구 결과에서 정확도 0.872, 재현율 0.872, 정밀도 0.871, 그리고 F1 점수 0.870을 기록함으로써, 각 폴드에서 상이한 성능을 보일 수 있음을 확인하였다. 마지막으로, RF 모델은 다른 모델들에 비해 상대적으로 낮은 성능을 보였으며 정확도 0.809, 재현율 0.809, 정밀도 0.804, 그리고 F1 0.803을 달성하였다.

본 연구에서는 모델의 예측 성능 뿐만 아니라, 계산 효율성을 평가하기 위해 각 모델의 학습 및 추론 시간을 측정하였다. 실험에 사용된 하드웨어 환경은 AMD Ryzen9 7950x3D CPU, 64GB RAM, NVIDIA GeForce RTX 4090 GPU이며, 소프트웨어 환경은 Python 3.10, Pytorch 2.2.1, CUDA 12.1이다. 각 모델의 학습 시간은 XGB, KNN, RF, LGBM, TabNet, CatBoost순으로 빠른 것을 확인하였다.

추가적으로, 2,377개의 테스트 데이터셋을 활용하여 TabNet과 XGBoost 모델을 통한 기업 신용 등급 예측의

정확성을 평가하였다. 기업 신용 평가에서 등급을 과대 평가하는 것은 과소 평가하는 것에 비해 더 심각한 부정적 결과를 초래한다. 이는 과대 평가된 경우, 부도 위험이 높은 기업이 과도하게 낮은 위험으로 평가될 수 있으며, 이로 인해 발생하는 기회비용이 더 크기 때문이다.

이러한 관점에서, 모델의 우수성은 예측된 등급이 실제 등급보다 높은 오류를 적게 발생시키는 능력으로 평가될 수 있다. 즉, 고위험 기업을 저위험 기업으로 잘못 분류하는 경우가 저위험 기업을 고위험 기업으로 잘못 분류하는 경우보다 더 적어야 한다는 것이다. 분석 결과, XGBoost 모델은 실제 등급보다 높게 예측한 경우가 653건 발생한 반면, TabNet 모델은 이러한 유형의 오류가 616건으로 나타나, TabNet이 상대적으로 더 우수한 성능을 보임을 확인할 수 있다.

더욱이, 등급 오류의 범위에 따른 분석에서도 XGBoost 모델은 한 등급 과대 평가 시 301건, 두 등급 과대 평가 시 163건, 세 등급 과대 평가 시 55건의 오류를 보였다. 반면, TabNet 모델은 각각 290건, 150건, 58건의 오류를 기록하며, 상대적으로 더 낮은 오분류 빈도를 보였다. 이러한 결과는 TabNet 모델이 기업 신용 등급 평가에 있어 XGBoost 모델에 비해 일관되게 더 정확한 예측을 제공함을 확인하였다.

Table 3. Comparison of Machine Learning Models' Performance Using 5-Fold Cross-Validation

TabNet				
Fold	ACC.	Recall	Prec.	F1
1	0.983	0.912	0.883	0.894
2	0.983	0.912	0.885	0.896
3	0.982	0.907	0.884	0.894
4	0.983	0.912	0.882	0.894
5	0.984	0.912	0.886	0.896
Avg	0.983	0.911	0.884	0.895
Std	0.000	0.002	0.001	0.001

XGB				
Fold	ACC.	Recall	Prec.	F1
1	0.823	0.823	0.827	0.824
2	0.850	0.850	0.849	0.849
3	0.881	0.881	0.878	0.879
4	0.923	0.923	0.919	0.919
5	0.944	0.944	0.943	0.941
Avg	0.884	0.884	0.883	0.882
Std	0.045	0.045	0.043	0.043

KNN				
Fold	ACC.	Recall	Prec.	F1
1	0.875	0.875	0.871	0.872
2	0.893	0.893	0.890	0.890
3	0.907	0.907	0.904	0.905
4	0.940	0.940	0.936	0.935
5	0.961	0.961	0.961	0.959
Avg	0.915	0.915	0.912	0.912
Std	0.031	0.031	0.032	0.031

CatBoost				
Fold	ACC.	Recall	Prec.	F1
1	0.807	0.807	0.809	0.808
2	0.832	0.832	0.829	0.830
3	0.861	0.861	0.856	0.858
4	0.898	0.898	0.893	0.894
5	0.922	0.922	0.920	0.918
Avg	0.864	0.864	0.861	0.861
Std	0.042	0.042	0.041	0.041

LGBM				
Fold	ACC.	Recall	Prec.	F1
1	0.808	0.808	0.812	0.809
2	0.833	0.833	0.833	0.832
3	0.867	0.867	0.863	0.864
4	0.915	0.915	0.911	0.911
5	0.935	0.935	0.933	0.932
Avg	0.872	0.872	0.871	0.870
Std	0.048	0.048	0.045	0.046

RF				
Fold	ACC.	Recall	Prec.	F1
1	0.767	0.767	0.761	0.761
2	0.782	0.782	0.776	0.775
3	0.803	0.803	0.795	0.795
4	0.836	0.836	0.832	0.830
5	0.857	0.857	0.857	0.852
Avg	0.809	0.809	0.804	0.803
Std	0.033	0.033	0.036	0.034

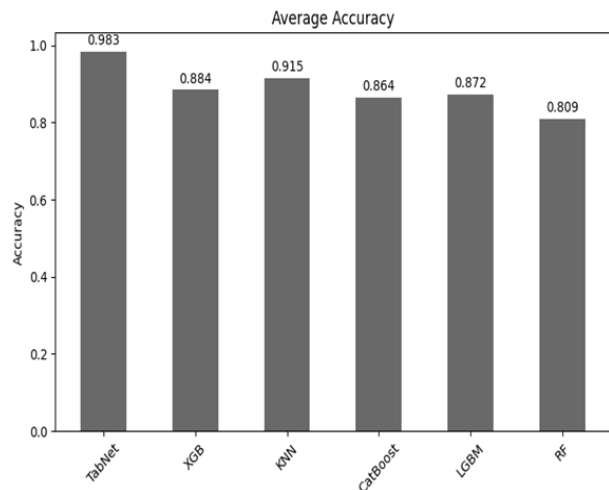


Fig. 5. Average Accuracy

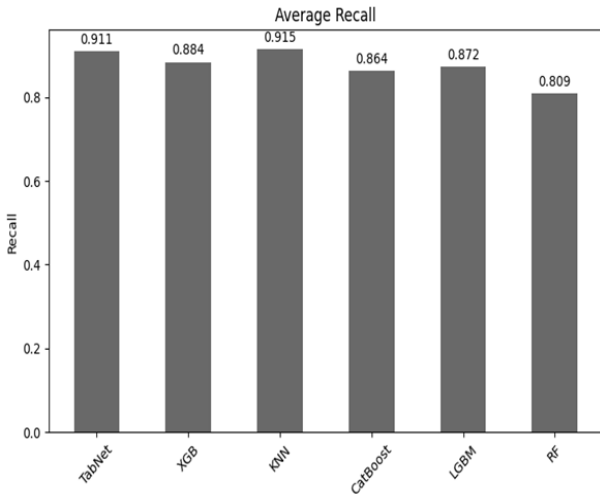


Fig. 6. Average Recall

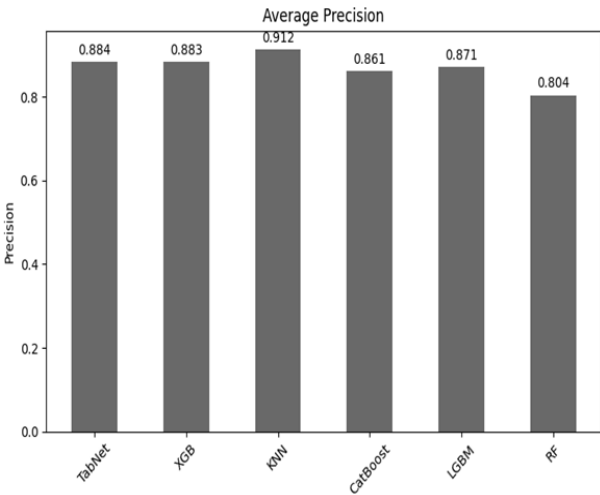


Fig. 7. Average Precision

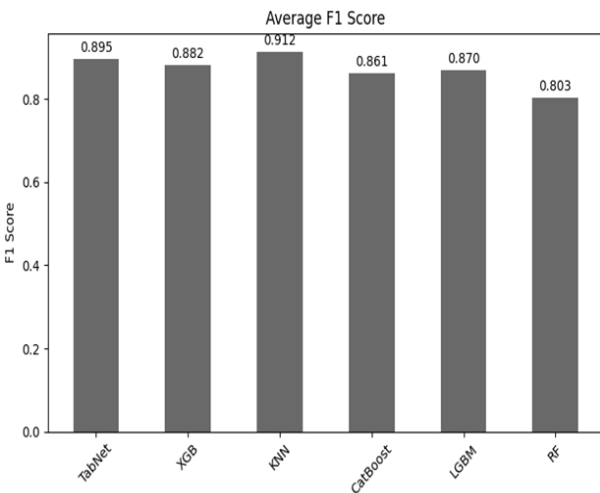


Fig. 8. Average F1 Score

V. Conclusions

본 연구에서는 정형 데이터 분석을 위해 TabNet 모델을 활용하였으며, 이를 통해 기업 신용 등급 평가의 정확성을 향상시킬 수 있음을 확인했다. TabNet 모델은 동적 특성 선택, 배치 정규화, 그리고 GLU 블록을 통합함으로써 내부 공변량 변화를 최소화하고 중요 정보를 강조하는 기법을 채택하였다. 본 연구의 결과는 한국 주요 증권시장에 상장된 1,517개 기업에서 수집된 데이터에 기반한 실증분석을 통해 검증되었으며, 모델의 예측 정확도 및 실용성이 실제 신용 등급 적용을 통해 TabNet이 더 정확한 예측을 한다는 것을 확인하였다.

본 연구는 이론과 실제 사이의 간극을 효과적으로 좁히는 데에 학술적으로 의미 있는 기여를 하고 있다. 실무적 관점에서 볼 때, 본 연구는 실제 기업 사례 분석을 통해 얻은 결과를 기반으로 하고 있어 TabNet의 적용가능성을 충분히 검증했다는 점에서 의의가 있다. 특히, 재무 정보를 활용한 딥러닝의 적용이 기존 머신러닝 기법보다 더 정확한 신용 등급 예측을 가능하게 함으로써, 비재무 정보가 부족한 소규모 기업들에게 실질적인 이점을 제공한다. 또한, 본 연구의 성과는 정확한 신용 등급 예측을 통해 여신 및 대출 기관의 위험 관리 개선에 기여할 수 있다는 점을 시사한다. 이는 금융 기관이 잠재적 손실을 예방하고 효율적인 자산 관리를 수행하는 데 있어 중요한 발전을 의미한다.

본 연구는 정형 데이터 분석에 TabNet 모델을 적용하여 기업 신용 등급 평가의 정확성을 향상시키는 방안을 탐구하였으나, 비재무 정보의 활용이 이루어지지 않았다는 점과 모델의 하이퍼파라미터 튜닝에 따라서 성능의 변화가 크다는 점에서 한계를 가진다. 특히, 비재무 정보의 포함 여부가 모델의 예측 능력에 미치는 영향에 대한 추가적인 검증이 필요하며, TabNet의 계산 효율성 및 하이퍼파라미터 튜닝 과정의 최적화를 통한 리소스 사용의 개선이 요구된다. 이에 따라, 향후 연구는 비재무 정보를 통합하여 모델의 성능 향상 가능성을 탐구하고, TabNet 모델의 동적 특성 선택 기능을 활용한 해석 가능한 모델 구축하여 모델의 투명성 및 신뢰성을 증진할 필요가 있다. 이러한 접근 방식은 기업 신용 등급 평가의 정확성 및 실용성을 더욱 강화하고, 금융 분야에서의 리스크 관리 및 의사 결정 과정을 효과적으로 지원할 수 있을 것이다.

REFERENCES

- [1] J. H. Park, G. Y. Kim, J. H. Ju, H. Lee, and H. J. Choi, "Factors Affecting Corporate Insolvency Prediction Based on Explainable Artificial Intelligence," *Journal of Digital Contents Society*, vol. 24, no. 9, pp. 2093-2105, Sep 2023, DOI: 10.9728/dcs.2023.24.9.2093
- [2] X. L. Zheng, M. Y. Zhu, Q. B. Li, C. C. Chen and Y. C. Tan, "FinBrain: when finance meets AI 2.0," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 7, pp. 914-924, Aug 2019. DOI: 10.1631/FITEE.1700822
- [3] S. Ö. Arik and T. Pfister, "TabNet: Attentive Interpretable Tabular Learning", *AAAI*, vol. 35, no. 8, pp. 6679-6687, May 2021. DOI: 10.1609/aaai.v35i8.16826
- [4] Y. j. Kim and M. J. Chung, "Initial Credit Ratings and Asymmetric Cost Behavior," *KOREAN JOURNAL OF MANAGEMENT ACCOUNTING RESEARCH*, vol. 21, no. 2, pp. 23-46, Aug 2021. DOI: 10.31507/KJMAR.2021.8.21.2.23
- [5] Y. W. Jeong and H. C. Chung, "The impact of bond ratings by domestic and foreign credit rating agencies on stock prices," *Korean Journal of Business Administration*, Vol. 27, No. 3, pp. 371-391, Mar 2014.
- [6] Y. J. Chen and Y. M. Chen, "Forecasting corporate credit ratings using big data from social media," *Expert Systems with Applications*, Volume 207, Nov 2022. DOI: 10.1016/j.eswa.2022.118042
- [7] H. C. Ahn, K. J. Kim and I. G. Han, "Intelligent Credit Rating Model for Korean Companies using Multiclass Support Vector Machines," *korean management review*, Vol. 35, No. 5, pp. 1479-1496, Oct 2006.
- [8] I. G. Han and K. S. Shin, "Development and Application of Intelligent Credit Scoring System: Boram Bank Case," *Information Systems Review*, Vol. 1, No. 1, pp. 51-61, 1999.
- [9] Z. Huang, 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*, Vol. 37, No. 4, pp. 543-558, Sep 2004. DOI: 10.1016/S0167-9236(03)00086-1
- [10] K. Kumar and S. Bhattacharya, "Artificial neural network vs linear discriminant analysis in credit ratings forecast: A comparative study of prediction performances," *Review of Accounting and Finance*, Vol. 5 No. 3, pp. 216-227, Jul 2006. DOI: 10.1108/14757700610686426
- [11] Y. Ye, S. Liu and J. Li, "A Multiclass Machine Learning Approach to Credit Rating Prediction," *2008 International Symposiums on Information Processing*, pp. 57-61, Moscow, Russia, May 2008. DOI: 10.1109/ISIP.2008.37
- [12] S. C. Huang, "Integrating nonlinear graph based dimensionality reduction schemes with SVMs for credit rating forecasting," *Expert Systems with Applications*, Vol. 36, No. 4, pp. 7515-7518, May 2009. DOI: 10.1016/j.eswa.2008.09.047
- [13] X. Guo, Z. Zhu and J. Shi, "A corporate credit rating model using support vector domain combined with fuzzy clustering algorithm," *Mathematical Problems in Engineering*, vol. 2012, Jul 2012. DOI: 10.1155/2012/302624
- [14] H. C. Wu, Y. H. Hu and Y. H. Huang, "Two-stage credit rating prediction using machine learning techniques," *Kybernetes*, Vol. 43, No. 7, pp. 1098-1113, July 2014. DOI: 10.1108/K-10-2013-0218
- [15] S. j. Kim and H. C. Ahn, "Application of Random Forests to Corporate Credit Rating Prediction," *The Journal of Industrial Innovation*, Vol. 32, No. 1, pp. 187-211, Mar 2016.
- [16] H. k. Park, J. Y. Kang, S. W. Heo and D. H. Yu, "Comparative study of prediction models for corporate bond rating," *The Korean Journal of Applied Statistics*, Vol. 31, No. 3, pp. 367-382, Jun 2018. DOI: 10.5351/KJAS.2018.31.3.367
- [17] S. H. Chung, C. H. Kim and G. H. Lee, "TabNet-Based Framework for Application of XAI to Corporate Credit Rating Models: On the Credit Rating Model for Enterprise," *Asia Pacific Journal of Small Business*, vol. 45, no. 3, pp. 45-65, Sep 2023. DOI: 10.36491/APJSB.45.3.3
- [18] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," In *International conference on machine learning*, pp. 448-456, pmlr, Jun 2015. DOI: 10.48550/arXiv.1502.03167
- [19] Y. N. Dauphin, A. Fan, M. Auli and D. Grangier, "Language modeling with gated convolutional networks," In *International conference on machine learning*, pp. 933-941, pmlr, Jul 2017. DOI: 10.48550/arXiv.1612.08083
- [20] A. Martins and R. Astudillo, "From softmax to sparsemax: A sparse model of attention and multi-label classification," In *International conference on machine learning*, pp. 1614-1623, pmlr, Jun 2016. DOI: 10.48550/arXiv.1602.02068
- [21] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, Jun 2002. DOI: 10.1613/jair.953

Authors



Hyeokjin Choi received his B.S. in Political Diplomacy from Kookmin University, Korea in 2022. He is currently enrolled in the master's program at the Graduate School of Business IT at Kookmin University.

His main research areas of interest are business analytics, deep learning and machine learning.



Gyeongho Jung received his B.S. in Smart Automotive Engineering from Halla University, Korea in 2023. He is currently enrolled in the master's program at the Graduate School of Business IT at Kookmin

University. His main research areas of interest are business analytics, deep learning and machine learning.



Hyunchul Ahn received a BS in Industrial Management from KAIST, and a ME and PhD from KAIST Graduate School of Management. He is currently working as a professor of the Graduate School of Business

IT at Kookmin University. His main research areas include AI applications in finance and customer relationship management as well as behavioral models related to information system acceptance.