

기업 신용평가 모형에 설명 가능한 AI 적용을 위한 TabNet 기반 프레임워크: 기업 신용평가 모형 연구를 중심으로*

정 승 호** · 김 창 호*** · 이 군 희****

국문 초록(Korean Abstract): 늘어나는 기업 대출에 맞춰 은행권에서는 기업 신용평가 모형의 고도화를 통한 여신 리스크 관리의 필요하다. 기업 신용평가 고도화를 위하여 기업 재무정보를 활용한 통계, 머신러닝, 딥러닝 기법들의 활용 연구가 활발히 진행되었다. 해당 연구들에서 사용된 대표적인 모델들은 전통적인 통계 기법인 로지스틱 회귀, 머신러닝 기법인 의사결정나무 기반의 모델들이 존재하며, 최근 여러 분야에서 강조되는 딥러닝 기반의 알고리즘을 활용한 연구 또한 높은 예측 성능을 보인다. 하지만 기업 신용 평가 연구에서 성능과 함께 집중해야하는 내용은 결과에 대한 해석력과 불균형 데이터의 처리이다. 본 연구에서는 성능을 위하여 새로운 변수를 생성하는 전처리법과 불균형 데이터를 처리하기 위한 샘플링 기법, 그리고 딥러닝 기반의 설명가능한 인공지능(XAI)인 TabNet을 활용한 새로운 프레임 워크를 제안한다. 연구결과 로지스틱회귀, 다층 퍼셉트론 대비 새로운 프레임워크의 높은 성능을 확인하였다. 또한 설명력을 분석한 결과 안정성, 수익성 지표가 예측에 크게 영향을 미쳤고, 성장성과 활동성 지표는 적게 영향을 미침을 확인하였다.

[주제어] 금융, 신용평가, 설명가능한 인공지능(XAI), TabNet, 불균형데이터, 이상치, 딥러닝

* 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 글로벌핵심인재양성지원사업의 연구결과로 수행되었음(2021-0-02126)

** 서강대학교 일반대학원 경영학과 석사과정 (seunghoch@naver.com, 주저자)

*** 서강대학교 일반대학원 경영학과 박사과정 (charley@sogang.ac.kr, 공동저자)

**** 서강대학교 경영학과 교수(ghlee@sogang.ac.kr, 교신저자)

TabNet-Based Framework for Application of XAI to Corporate Credit Rating Models: On the Credit Rating Model for Enterprise*

Seungho Chung** · Changhyo Kim*** · Gunhee Lee****

Abstract: Credit risk management has recently been emphasized through the advancement of the corporate credit evaluation model. Representative models used in studies related to corporate credit evaluation have various statistics and machine learning techniques, and recently, deep learning-based algorithms that show good performance in various fields have been studied to show high predictive performance. However, in corporate credit evaluation research, not only performance but also the content to focus on together is an interpretation of results and the processing of imbalanced data. In this work, we propose a preprocessing method that generates new variables for performance, sampling techniques for handling unbalanced data, and a new framework utilizing TabNet, an explainable artificial intelligence (XAI) based on deep learning. As a result of the study, the high performance of the new framework was confirmed compared to logistic regression and multilayer perceptron. In addition, as a result of analyzing explanatory power, it was confirmed that stability and profitability indicators had a significant effect on prediction, and growth and activity indicators had a negligible effect.

[Keywords] *Credit Rating, Explainable Artificial Intelligence(XAI), TabNet, Unbalanced Data, Outlier, Deep Learning*

* This research was conducted as part of the research outcomes of the Ministry of Science and ICT and the National IT Industry Promotion Agency's Global Core Human Resources Development Support Project (2021-0-02126).

** M.S., Department of Business Administration, Sogang University(seunghoch@naver.com, First Author)

*** Ph.D. Candidate, Department of Business Administration, Sogang University (charley@sogang.ac.kr, Co-Author)

**** Professor, College of Business Administration, Sogang University (ghlee@sogang.ac.kr, Corresponding Author)

I. 서 론

‘고용 없는 성장’이라는 국내 경제의 흐름 속에서 기업들의 역할과 중요성은 과거 대비 더욱이 강조되고 있다. 대기업과 중견기업은 전문적인 역량과 기술 경쟁력으로 국가 경제성장과 고용창출 이끌고 중소기업은 탄탄한 산업구조를 형성하고 신기술과 제품 개발을 통한 산업발전을 주도한다. 하지만 이러한 중견기업과 중소기업은 대기업 대비 자금 조달의 한계가 존재한다(김소정, 이근희, 2022). 이에 기업들은 금융권을 통하여 기업 대출을 실시한다. 하지만 은행권의 신용평가모형은 재정 안정성이 높은 기업들에 맞춰 고도화되어 있다. 은행권은 늘어나는 기업대출에 걸맞은 신용평가모형 개발이 필요하며 이를 통한 여신 리스크 관리의 중요성이 강조되고 있다. 이러한 필요성에 따라본 연구에서는 기업 재무 데이터를 활용하여 금융권에서 활용 가능한 기업의 부도 여부를 예측하고자 한다.

최근 인간의 지적 능력을 컴퓨터로 구현하고자 하는 인공지능(artificial intelligence)의 활발한 연구로 인공지능 기술이 다양한 분야에서 우수한 성능을 보이고 있다. 특히 심층신경망(deep neural network)을 학습하여 결과를 예측하는 딥러닝 알고리즘은 회귀와 분류 문제에 있어 탁월한 성능을 보이고 있다. 딥러닝의 강력한 성능과 가능성에 금융 시장을 포함한 여러 사회 과학 분야에서 부도 예측 모델에 딥러닝 기술을 도입하는 등 다양한 인공지능 기술 적용을 시도하는 추세이다(Zhen, 2019).

하지만, 이러한 딥러닝의 강력한 성능에 반하여 한계가 되는 것은 딥러닝의 결과 도출 과정이 해석할 수 없는 블랙박스(black-

box)라는 점이다. 즉, 입력값에 대하여 원하는 출력값을 얻을 수 있지만 해당 결과가 도출되는 과정을 직관적으로 알지 못한다. 이러한 이유로 금전적인 피해를 직접적으로 야기할 수 있는 금융 산업에서는 인공지능 기술의 적용에 신중해야하며, 금융 데이터를 이용한 분석을 위해서는 모델 결과 도출 과정에 대한 설명력이 요구된다. 이러한 문제를 해결하기 위해, 설명 가능한 인공지능 XAI(explainable AI)와 관련된 많은 연구가 이루어지고 있다. 또한 금융 데이터는 정해진 형식에 따라 저장되는 정형 데이터(tabular data)로 이루어진 경우가 많다. 자연어 처리나 영상 처리와 같은 비정형 데이터를 이용한 연구는 많이 진행된 반면에, 정형 데이터에 관한 연구는 부족한 실정이다. 그러나 최근 정형화된 데이터에 특화된 XAI 모델인 TabNet(Arik and Pfister, 2021)이 제안되었다. 해당 모델은 의사결정나무의 특징과 XAI의 특징을 지니고 있다. 의사결정나무의 결정경계 분류법을 심층신경망으로 모방하여 목적함수를 경사하강법으로 계산하는 특징이 있다. 동시에 전역 설명력과 지역 설명력을 제공할 수 있어 정형화 데이터에서 우수한 성능을 보이는 설명 가능한 인공지능 모델로 손꼽힌다.

금융권에서 다루지는 신용 사기 분류 모델과 부도 예측 모델 등과 같은 분류문제에서는 데이터 불균형 문제가 존재한다. 부도 예측 문제에서 분류하고자 하는 부도 기업의 데이터는 전체 기업의 데이터에서 극히 일부이다. 이는 정상 범주의 데이터들에 과적합(overfitting)이 되어 분류기가 제대로 학습이 되지 않는 문제를 야기한다. 이에 본 연구에서는 데이터 샘플링 기법 중 하나인 SMOTE+ENN(Batista, Prati and Monard,

2004)를 적용하여 데이터 불균형의 문제를 보완하였다.

본 연구에서는 이러한 배경을 바탕으로 부도예측에 사용되는 데이터의 불균형 문제 해소를 위한 복합샘플링 방법론과 함께, 정형 데이터(tabular data)를 기반으로 심층 신경망의 우수한 성능과 설명가능한 인공지능(XAI)의 장점을 지닌 TabNet을 이용하여 부도 예측 및 예측 결과에 대한 설명이 가능한 프레임워크를 제안하고자 한다.

성능을 검증하기 위해 연구에 사용된 데이터는 기업 외감 데이터로 기업의 재무 상태와 현금흐름을 바탕으로 생성된 변수값과 기업의 부도 여부를 포함한다. 이러한 전처리 과정은 기존 정형화된 데이터를 더욱이 정형화된 데이터의 특성을 띄게 만든다. 실험 과정에서 학습과 추론은 k-fold (k=4)를 통한 교차검증을 실시하였다. 데이터 전처리과정과 TabNet 인코더를 활용한 프레임워크의 결과는 전통적인 분류 모델인 로지스틱 회귀와 다중 퍼셉트론(multi-layer perceptron, MLP) 모델을 활용한 프레임워크 대비 AUC, F1 스코어, FNR, 균형정확도 결과 값에서의 증가를 보였다. 모델의 결과 예측에 영향을 미치는 변수들을 분석해본 결과 재무제표의 안정성과 수익성에 해당하는 변수들이 부도 예측에 주요한 영향을 끼침을 확인하였다. 또한, 데이터 이상치 및 결측치의 전처리를 통해 생성한 변수들도 모델의 결과 예측에 영향을 끼침을 확인하였다. 모델의 성능이 증가하였으므로 해당 전처리 기법이 모델 학습에 추가 정보를 제공하였다고 해석할 수 있다. 추가 실험으로 데이터 불균형 문제를 다루기 위하여 해당 프레임워크의 데이터 전처리과정에 SMOTE+ENN의 복합샘플링 기법을 추가하여 기존 결과

와 비교하였다. 이를 통해 SMOTE+ENN 복합샘플링법이 추가로 이루어진 프레임워크가 모든 평가척도에서 우수한 분류 성능 보임을 확인하였다.

II. 선행 연구 및 방법

2.1 선행 연구

기업 재무정보는 기업의 성과와 경쟁력 등 많은 정보를 담고 있는 정형화된 데이터이다(안태일, 김종민, 2002). 이러한 기업 재무정보를 활용한 신용정보 예측 혹은 분류 모델 연구는 과거부터 활발히 연구되어 왔다(김진선, 최영문, 2006; Ohlson, 1980). 연구들의 기법으로 전통적인 통계 방법인 로지스틱 회귀부터 최근에는 각종 머신러닝 기법들과 인공지능 기법 등이 활용되었다(오우석, 김진화, 2017). 머신러닝 기법 중 Support Vector Machine(SVM)을 활용한 기업부도예측 방법론은 다변량판별분석 및 로지스틱회귀분석과 같은 통계분석모형 보다 우수한 예측력을 보였다(민재형, 이영찬, 2005). 또 다른 머신러닝 기법인 트리기반의 모델은 해석력이 뛰어나고 정형화된 데이터에 특화된 좋은 성능을 보였으며(조준희, 강부식, 2007), 인공지능 기법으로는 심층신경망과 같은 딥러닝을 적용하는 방법들을 이용하여 신용평가 모형에 활용하고 있다.

하지만 인공지능 모델은 결과 도출 과정을 알 수 없다는 한계점이 존재한다. 이에 인공지능 모델의 입력에 따른 결과의 변화를 분석하는 설명가능한 인공지능(XAI) 기술이 등장하였다. XAI에는 대표적으로 샵(SHAP)(Lundberg and Lee, 2017)과 라임

(LIME)(Gehring, Auli, Grangier, Yarats and Dauphin, 2017) 모델이 있다. 두 모델은 대리 모델(surrogate model)로 입력 값의 변화에 따른 예측 값의 변화를 분석하여 변수 중요도를 출력한다. 즉, 결과에 대한 귀납적 추론을 통해서 설명력을 확보한다. 샵 모델은 전역 설명력(global interpretability), 라임 모델은 지역 설명력(local interpretability)을 제공한다(Adadi and Berrada, 2018).

2.2 선행 연구 방법론

해당 장에서는 본 연구에 적용된 TabNet 과의 모델 성능 비교분석을 위해 사용된, 로지스틱 회귀 및 다층 퍼셉트론 기법에 대해 설명한다.

2.2.1 로지스틱 회귀

로지스틱 회귀란 독립 변수의 선형 결합을 이용하여 종속 변수를 예측하는 확률 추론 통계 기법이다. 로지스틱 회귀 p 의 수식은 다음과 같다.

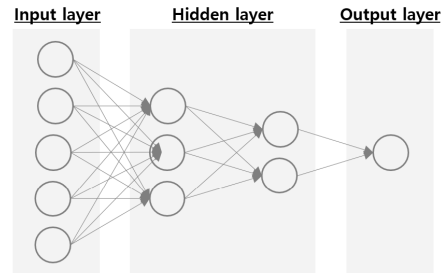
$$Y = p(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}$$

Y 는 종속 변수, X 는 독립 변수이며 β 는 회귀계수 값으로 각 변수를 k 로 나타낸다. 종속 변수가 범주형 데이터일 때를 대상으로 사용되며 분류 문제를 풀기 위하여 활용된다. 이러한 로지스틱 회귀는 의료, 경제학 등의 분야에서 예측을 위하여 활용된다. 해당 연구에서는 이항형 로지스틱 회귀를 사용할 것이다.

2.2.2 다층 퍼셉트론

다층 퍼셉트론(multi-layer perceptron, MLP)

이란 퍼셉트론 뉴런으로 이루어진 층을 여러 개 쌓은 인공 신경망이다. 입력층(input layer)과 출력층(output layer) 사이에 하나 이상의 은닉 층(hidden layer)을 가지고 있으며, <그림 1>와 같이 표현된다.



<그림 1> 다층퍼셉트론(2개 은닉층)

또한 입력에 대해 비선형적인 연산을 통하여 신경망을 학습시키는데, 이 과정에서 각 층에서는 활성화함수(activation function)를 통해 입력을 처리한다. 그림1을 수식으로 표현하면 다음과 같다(Goodfellow, Yoshua and Aaron, 2016).

$$Y = g^{(2)}(W^{(2)T}(g^{(1)}(W^{(1)T}x + b^{(1)})) + b^{(2)})$$

여기서 x 는 입력층, W 와 b 는 은닉층의 가중치와 편향(bias)이며, 학습을 통해 추정되는 계수(parameter)가 된다. 끝으로 g 는 활성화함수를 나타내며, 괄호안의 숫자는 은닉층수를 의미한다.

신경망의 출력값인 예측 값과 실제 값의 차이(error)를 손실 함수(loss function)를 통해 계산하며, 이 차이를 바탕으로 역전파(back propagation)를 이용하여 인공 신경망의 가중치(weight)를 계산한다. 인공신경망을 학습시키기 위해 경사하강법(gradient descent) 최적화 알고리즘을 주로 사용한다.

인공신경망 방법론은 다층 퍼셉트론 외에도 이미지 인식(image recognition) 분야에 서 좋은 성과를 나타내는 CNN(Convolutional Neural Network) 알고리즘과 시계열 자료의 예측이나 텍스트(text) 자료의 문맥 인식 등에 사용되는 RNN(Recurrent Neural Network) 알고리즘 등이 있다(김형준, 류두진, 조훈, 2019).

이러한 인공지능 모델은 다른 알고리즘에 비해서 복잡한 형태의 모델에 대해서 보다 높은 분석능력을 갖고 있는 반면, 분석시간이 상대적으로 매우 길고, 블랙박스 모형으로 수치의 해석에 한계점이 있다.

III. 실험내용 및 방법

3.1 실험 Framework 및 구성도

본 연구에서는 재무정보에서 추가적인 정보를 얻기 위해 시행한 예외값 및 결측치 전처리 기법과 불균형 데이터의 처리를 위한 복합샘플링 기법 그리고 분류기를 하나의 프레임워크로 제안하고 로지스틱 회귀, 다층퍼셉트론, TabNet 인코더를 분류기로 사용하여 각 분류기에 따른 프레임워크의 결과 비교를 통한 성능 확인을 하였다. 이후 TabNet을 활용하여 변수설명력을 해석

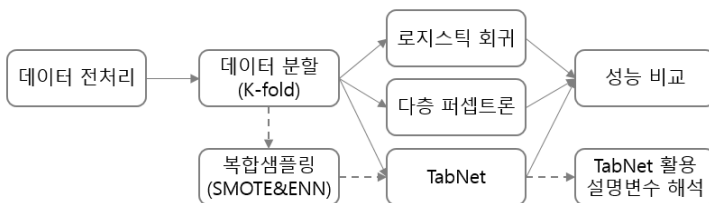
하였다. 전반적인 실험 프레임워크는 <그림 2>와 같다.

3.2 실험 방법

3.2.1 데이터 전처리

본 연구에서 기업의 부도 예측 모델의 성능 검증을 위해 중소기업통합시스템(SIMS)에 포함된 비금융산업 기업의 2008년부터 2021년 기간 동안 수집한 외감 대상 기업의 재무 데이터를 사용하였다. 해당 데이터는 158,336건의 기업 정보를 포함한다. 부도를 예측하기 위한 연구 모형에서의 종속 변수는 기업의 부도 여부로 정상 기업을 0, 부도인 기업을 1로 나타내는 범주형 데이터이다. 이는 코리아크레딧뷰로(Korea Credit Bureau)가 제공한 정보를 활용하였다. 부도를 설명하기 위한 독립 변수는 크게 재무비율 수치형 변수 27개와 이 수치형 변수의 전처리를 통해 생성한 명목 변수 52개로 이루어져 있다. 해당 변수들은 부록 A.2에 나타내었다. 기업의 재무상태표와 현금 흐름표의 수치를 기반으로 계산된 재무비율 수치형 변수는 안정성 지표, 활동성 지표, 수익성 지표, 성장성 지표로 분류한다. 해당 변수들은 <표 1>에 제시하였다.

명목 변수는 업종을 구분하는 변수와 재무변수에 대한 이상치를 제거하는 전처리



<그림 2> TabNet을 활용한 실험 프레임워크 구성도

〈표 1〉 재무비율 수치형 변수

지표종류	특성 항목명
안정성 지표	유동비율
	당좌비율
	부채비율
	자기자본비율
	차입금/자기자본
	차입금의존도
	단기차입금/총자산
	영업활동후현금흐름/매출액
	이자보상배율
	부채상환능력계수
	차입금/EBITDA
	EBITDA대금융비용
활동성 지표	총자산회전율
	순운전자본회전율
	총자산회전율
	매출채권회전율
수익성 지표	매출액순이익률
	총자산순이익률
	자기자본순이익률
	매출액영업이익률
	매출액총이익률
	EBIT대매출액 (법인세비용차감전순이익)
성장성 지표	매출액증가율
	영업이익률
	순이익증가율
	총자산증가율
	법인세비용차감전순이익증가율

과정을 통해 만들어진 더미(dummy) 변수이다. 생성된 데이터인 결측치 및 예외값, 이상치 상한(upper), 이상치 하한(lower)으로 더미 변수에 속한다. 결측치 및 예외값 더미는 유효하지 않은 수치이다. 0으로 나눌 수 없는 경우거나 부채 금액이 음수인 경우가 이

에 속한다. 이러한 수치는 중앙값으로 대체 후 더미변수를 생성하여 마킹(marking)하였다. 이상치의 경우 사분위수 범위(Inter Quantile Range, IQR) 절대값 5배를 초과하는 관측치를 극단적 이상치로 정의한다. 이상치는 경계값으로 대체 후 더미 변수를 생성하여 마킹하였다. 상한의 경우 +5배 상한 사분위수 범위를 초과한 경우이고 하한의 경우 -5배 하한 사분위수 범위를 초과한 경우이다.

이러한 전처리 과정을 통해 생성된 명목 변수 중 모든 데이터가 단일 값(0 또는 1)만 존재하는 경우를 제외하였다.

3.2.2 샘플링 기법

데이터불균형 문제를 해결하기 위하여 데이터를 샘플링하는 방법, 학습 시 클래스 별로 가중치를 주는 방법, 데이터를 추가로 생성하는 방법 등, 다양한 방법들이 제시되고 있다. 데이터 샘플링은 크게 두 가지로, 데이터 수가 적은 클래스를 데이터 수가 많은 클래스에 맞추기 위한 오버샘플링(over-sampling) 방법과 반대로 데이터 수가 많은 클래스를 데이터 수가 적은 클래스에 맞추는 언더샘플링(under-sampling)이 존재한다(Barandela, Baldovinos, Sanchez and Ferri, 2004).

SMOTE(Synthetic Minority Over-sampling Technique)는 합성 데이터를 생성하는 기법으로 기존 소수 샘플을 보간하여 새로운 소수 데이터를 합성한다. 임의의 소수 데이터를 선택한 후 KNN(K-Nearest Neighbors)을 활용하여 K개의 가장 가까운 이웃 사이의 차이를 계산한 후 해당 차이에 0과 1 사이의 난수를 곱하여 기존 선택된 데이터에 해당 벡터를 더하여 데이터를 합성한다. 이와 같

이 계산되는 SMOTE를 하기 위하여 범주형 데이터 또한 처리 가능한 SMOTENC 알고리즘을 활용하여 소수데이터 합성을 하였다. ENN(Edited Nearest Neighbors)은 언더샘플링 기법으로 KNN을 사용하여 이웃한 데이터 중 자신과 같은 클래스보다 다른 클래스의 데이터가 많은 경우 해당 데이터를 제외하는 기법이다. 본 연구에서는 오버샘플링 기법인 SMOTE와 언더샘플링 기법인 ENN을 합친 복합 샘플링(hybrid sampling) 방법을 활용하였다.

3.2.3 TabNet

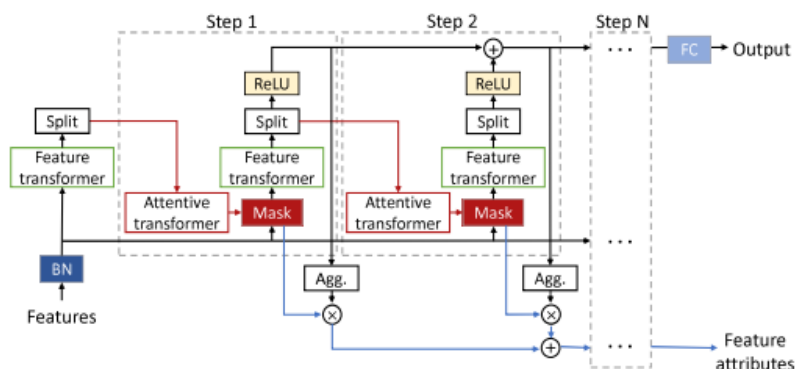
본 연구에서는 정형화된 중소기업 외감 데이터를 이용한 부도 예측 프레임워크를 구축하기 위해 TabNet(Arik and Pfister, 2021)의 인코더(encoder)를 적용한다. TabNet은 구글에서 2019년에 개발한 모델로 의사결정나무 모델의 설명력과 변수 선택 기능을 가진, 심층신경망(deep neural network) 기반의 종단 간 학습(end-to-end learning) 모델이다. 현재 정형화 데이터 분석에서 많이 쓰이는 기존의 의사결정나무(decision tree) 기반 모델보다 좋은 성능을 보여 정형화 데이터

딥러닝(deep learning) 분야에서 주목 받는 최신 기법 중 하나이다.

기존 의사결정나무 알고리즘의 특징인 결정 경계(decision boundary) 나눠 결과를 도출하는 특징을 모방하고 경사하강법(gradient descent) 기반 최적화를 통하여 학습한다는 특징이 있다.

<그림 3>은 TabNet의 인코더 구조를 나타낸다. 인코더의 각 의사결정단계(step)는 크게 특징변환기(feature transformer)와 입력 변수변환기(attentional transformer)로 이루어져 있다. 인코더의 입력 값은 해당 변환기들을 통과하여 마스크(mask)를 출력하고, 이 마스크를 통해 입력변수선택(feature masking)이 이루어진다. 이는 인코더의 각 의사결정단계마다 순차적으로 사용될 변수선택(feature selection)이 된다는 것을 의미한다. 의사결정 단계의 횟수는 하이퍼파라미터(hyper-parameter)로 지정할 수 있으며, 이를 거쳐 최종적으로 결과 값과 변수기여도(feature attributes)를 구할 수 있게 된다.

해당 장에서는 인코더의 구성요소인 특징변환기와 입력변수변환기에 대해 설명하고 이러한 구조를 통해 TabNet이 갖는 설명가

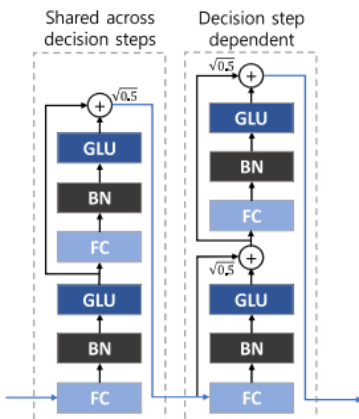


〈그림 3〉 TabNet Encoder 구조(Arik and Pfister, Figure 4(a), 2021)

능한 AI(explainable artificial intelligence, XAI)의 특징인 결과 해석력을 다루고자 한다.

(1) 특징변환기

특징변환기의 구조는 <그림 4>와 같다. 특징변환기는 변수(feature)를 모델이 학습하기 위한 형태로 변환한다. 완전연결계층(fully connected layer, FC layer)과 배치정규화(batch normalization, BN), 게이트선형유닛(gated linear unit, GLU)을 차례로 연결한 FC layer-BN-GLU가 하나의 블록을 구성한다. BN은 배치(batch) 간의 내부 공변량 변화(internal covariant shift)를 방지하여, 학습 과정에서 기울기 소실(gradient vanishing) 또는 폭주(gradient exploding) 문제를 방지할 수 있다.



<그림 4> 특징변환기 구조(Arik and Pfister, Figure 4(c), 2021)

게이트선형유닛은 입력된 값의 절반을 정보로 이용하고, 나머지 절반을 시그모이드 함수를 거친 게이트로 사용함으로써 정보의 흐름을 제어한다. 이의 정의는 아래 수식 (1)과 같다. σ 는 시그모이드 함수이며,

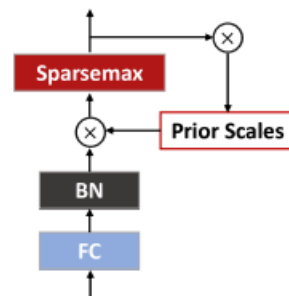
A와 B는 완전연결계층과 배치 정규화를 거친 출력 값의 앞의 절반과 뒤의 절반을 의미하며 같은 차원을 가진다. 이러한 FC layer-BN-GLU인 하나의 블록은 학습 가능한 활성화 함수의 역할을 하게 된다. 이 블록은 의사결정단계에 따라 가중치(weight)를 공유하는지의 여부에 따라, 공유블록과 개별 블록으로 나뉘어 특징변환기를 구성한다. 즉, 공유블록은 인코더의 모든 의사결정 단계에서 학습이 되며, 개별블록은 의사결정 단계별로 학습된다. 이러한 형태는 모델용량을 줄이고 계산 효율성을 향상시킨다.

$$GLU(A, B) = A \otimes \sigma(B). \quad (1)$$

마지막으로 잔차정규화연결(Normalized residual connection)을 통하여 학습 과정에서 분산이 극적으로 변하지 않고 안정적으로 학습된다(Gehring et al, 2017). 이러한 과정을 거친 특징변환기의 결과 값은 변수 선택을 위해 입력변수변환기로 전달된다.

(2) 입력변수변환기

<그림 5>은 입력변수변환기의 구조이다. 입력변수변환기는 중요한 변수를 마스킹(masking)하기 위해 사용한다.



<그림 5> 입력변수변환기 구조(Arik and Pfister, Figure 4(d), 2021)

먼저, 사전변수사용척도(prior scale)는 해당 의사결정단계에서 어떠한 변수를 사용할 지 결정하고, 이는 아래 수식(2)과 같다.

$$P[i] = \prod_{j=1}^i (\gamma - M[j]). \quad (2)$$

의사결정단계(step)를 나타내는 j와 변수 사용조절변수 γ 는 모델 구축 시 설정 하는 하이퍼파라미터이다. γ 는 의사결정단계에서 기존 사용되었던 변수의 재사용 여부에 영향을 끼친다. 즉, $\gamma = 1$ 로 설정한다면 각 변수들은 하나의 의사결정단계에서만 선택 될 수 있으며, γ 를 증가시키면 여러 의사결정단계에서 사용될 확률이 증가한다. M은 마스크(mask) 값으로, 이전 의사결정단계들의 스펬스맥스(sparsemax)(Martins and Astudillo, 2016)의 출력 값이다.

FClayer와 BN을 거친 입력변수변환기의 입력 값은 사전변수사용척도와 요소별 곱(element wise product)을 사용하여 스펬스맥스로 입력된다. 스펬스맥스는 소프트맥스(softmax)의 극단적인 형태의 정규화 기법으로 수식 (3)으로 나타낸다.

$$\text{sparsemax}(t) = \begin{cases} 1, & \text{if } t > 1 \\ (t+1)/2 & \text{if } -1 \leq t \leq 1 \\ 0 & \text{if } t < -1. \end{cases} \quad (3)$$

이러한 스펬스맥스를 활용하여 입력된 값은 0과 1로 이동된다. 이를 통해 각 변수의 중요성을 극명하게 구분지어 해당 의사결정단계에서 어떠한 변수를 사용할지를 결정하는 마스크로 출력된다.

이상과 같이 TabNet은 각 의사결정단계(step)별 입력변수변환기(attentive transformer)와 특징변환기(feature transformer)를 통해 변수선택(feature selection) 및 변수 가중치

(weight)를 학습하게 된다. 특히 사전변수사용척도(prior scale) 및 스펬스맥스(sparse-max)를 활용한 변수 마스크(mask)를 통해 입력단위로 변수(feature)를 선택할 수 있는 장점이 있다(Arik and Pfister, 2021). 때문에 본 연구에서도 데이터 탐색을 통한 사전 변수 제거 없이, 모든 변수를 학습에 사용하였다.

(3) 설명가능한 TabNet

TabNet은 설명 가능한 설명가능한AI로 지역설명력(local interpretability)과 전역설명력(global interpretability)을 지닌다. 지역 설명력은 각 예측별 (instance wise) 대한 변수별 중요도(feature importance)를 통해 확보할 수 있다. 이를 위해 각 의사결정단계별 변수 기여도(Feature attributes)를 계산해야한다. 변수기여도는 해당 의사결정단계에서 쓰이는 변수를 선택하는 마스크 값인 M과 렐루 함수(ReLU)를 통해 나온 결과 값의 요소별 곱을 통해 구할 수 있다. 이러한 각 의사결정단계의 변수 기여도들을 취합하여 해당 예측 값에 대한 변수별 중요도를 계산한다. 전역설명력은 모든 예측치에 대한 각 변수별 중요도를 계산하는 것으로 확보된다. 즉, 각 예측치에 대한 변수별 중요도를 모두 계산한 후 그 값들을 평균내는 것으로 전체 예측 변수 중요도를 구할 수 있다.

3.2.4 평가 지표

본 연구에서는 예측 모델을 평가하기 위해 AUC(area under the ROC curve), F1-Score, FNR(False Negative Rate), 균형 정확도(balanced accuracy) 총 4가지 평가 지표를 사용한다. 모델 예측에 따라 표 2 같이 혼동 행렬(confusion matrix)을 구할 수 있다.

〈표 2〉 혼동 행렬

Actual Class			
Positive (1)		Negative (0)	
Predicted Class			
Positive (1)	Negative (0)	Positive (1)	Negative (0)
True Positive (TP)	False Negative (FN)	False Positive (FP)	True Negative (TN)

그림에서 TP는 실제 양성(positive)인 경우를 모델이 양성으로 옳게 예측하는 것을 의미하며 FN은 실제 양성인 경우를 음성(negative)으로 잘못 예측하는 것을 의미한다. TN은 실제 음성인 경우를 음성으로 옳게 예측하는 것을 의미하며 FP는 실제 음성인 경우를 양성으로 잘못 예측하는 것을 의미한다. 이 값들을 이용하여 모델의 성능을 평가하는 아래의 평가 지표들을 계산할 수 있다.

(1) Area Under the ROC curve (AUC)

AUC는 ROC 곡선과 x축이 이루고 있는 면적의 넓이를 나타낸다. ROC 곡선의 x축은 FPR (False Positive Rate), y축은 TPR (True Positive Rate)를 이용하여 계산되며, 해당 값들은 다음 수식으로 계산할 수 있다.

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

FPR은 실제로 음성인 전체 경우 중에 모델이 양성으로 잘못 예측한 경우의 비율을 의미하며 TPR은 실제로 양성인 전체 경우 중에 모델이 양성으로 옳게 예측한 비율을 나타낸다. 따라서 경계값(threshold)에 따른 FPR과 TPR을 표시하여 ROC 곡선을 그릴

수 있다. AUC 값이 1에 가까울수록 이상적인 분류 성능을 보이는 것을 의미한다.

(2) F1-Score

F1-Score은 모델의 정밀도(precision)와 재현율(recall)의 조화 평균으로 다음과 같이 구할 수 있다.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

정밀도는 모델이 양성으로 예측한 경우가 실제로 양성인 비율을 의미하고 재현율은 실제 양성인 경우 중 모델이 양성으로 맞춘 비율을 의미하며 다음과 같은 수식으로 계산된다.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

정밀도와 재현율 값은 트레이드오프(trade-off)로 이들의 조화 평균인 F1-Score는 데이터 클래스 간의 불균형이 심한 경우에 평가 척도로 활용되며 이 값이 증가할수록 좋은 모델 성능을 의미한다.

(3) False Negative Rate(FNR)

FNR은 실제 양성인 경우를 모델이 음성으로 검출한 비율을 의미한다.

$$FNR = \frac{FN}{TP + FN}$$

본 연구에서는 실제 부도인 기업의 경우를 모델이 정상 기업으로 잘못 검출한 비율을 의미한다. 따라서 이 값이 감소할수록 좋은 모델 성능을 의미한다.

(4) 균형정확도

균형 정확도는 불균형한 데이터에서 정확도를 측정할 때 주로 사용되며 각 클래스

별 정확도를 평균을 낸 것으로 다음과 같이 계산된다.

$$\text{균형정확도} = \frac{1}{2} \times \left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP} \right)$$

해당 지표는 증가할수록 좋은 모델 성능을 의미한다.

IV 실증분석 결과

4.1 분류기 성능 비교

기업 부도예측 모델의 성능 비교를 위해 3.2장에서 설명한 데이터셋을 이용하여 로

지스틱 회귀, MLP (multi-layer perceptron)와 TabNet으로 학습 및 평가를 진행하였다. MLP는 입력층과 3개의 은닉층 그리고 출력층으로 구성되어 있다. 활성화 함수는 시그모이드(sigmoid) 함수를 사용했으며 손실 함수로 binary cross-entropy를 사용하였다. 학습의 최적화는 Adam을 통해 이루어졌다. TabNet의 의사결정단계는 5번으로 설정하고 변수사용조절변수인 γ 는 1로 설정하였다. TabNet 또한 최적화를 위하여 Adam을 사용하였다.

모델에서 정의되는 초매개변수(hyper parameter)는 MLP 및 TabNet 두 모델 모두 격자탐색(grid search) 방법을 통해 모델성능을

〈표 3〉 분류기 별 성능 비교

k-fold #	AUC ↑	F1-Score ↑	FNR ↓	Balanced Accuracy ↑
Logistic Regression				
1	0.846	0.864	0.222	0.760
2	0.838	0.860	0.227	0.764
3	0.850	0.860	0.227	0.759
4	0.831	0.814	0.300	0.731
Mean	0.841	0.850	0.244	0.754
MLP				
1	0.846	0.838	0.261	0.759
2	0.834	0.789	0.333	0.743
3	0.846	0.909	0.150	0.747
4	0.828	0.778	0.350	0.730
Mean	0.838	0.829	0.273	0.745
TabNet				
1	0.866	0.895	0.166	0.803
2	0.883	0.891	0.172	0.805
3	0.857	0.883	0.189	0.780
4	0.869	0.849	0.238	0.791
Mean	0.868*	0.880**	0.191**	0.795*

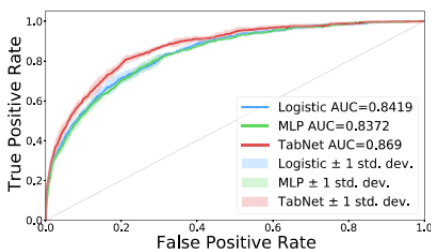
TabNet vs. MLP 비교

* 유의수준 5%에서 통계적으로 유의함, ** 유의수준 10%에서 통계적으로 유의함

최적화시키는 값으로 정의하였다.

모델의 성능 평가는 AUC, FI-Score, FNR, 균형 정확도로 진행되었으며 각 모델에 대한 결과는 표 3에 나타내었다. 해당 데이터셋은 불균형한 데이터셋이기 때문에 데이터 분할에 따라 성능에 편차가 존재한다. 따라서 k-fold (k=4)를 통하여 모든 데이터셋을 훈련하여 모델의 성능을 검증하고자 한다. 데이터셋 분할은 훈련, 검증, 테스트를 6:1:1의 비율로 진행하였다. 또한 데이터를 분할할 때 종속 변수의 클래스 비율을 맞춰주기 위해 계층적 샘플링(stratified sampling)을 적용하였다.

<표 3>에서 각 모델의 마지막 행은 k-fold 별 결과의 평균값을 나타낸다. TabNet 모델의 AUC, FI-Score, 균형 정확도는 로지스틱 회귀 모델 대비 각각 0.0261, 0.03, 0.0413, MLP 모델 대비 0.0291, 0.0511, 0.05 증가하였다. FNR의 경우 로지스틱 회귀 모델 대비 0.526, MLP 모델 대비 0.0291 감소하였다. 즉, TabNet 모델은 로지스틱 회귀 모델과 MLP 모델과 비교하여 가장 좋은 분류 성능을 보인다. 분석 모델로 사용된 각 모델의 k-fold 결과와 이의 평균에 대한 ROC 곡선은 부록 A.1에 첨부하였다. 세 모델의 비교 ROC 곡선은 <그림 6>과 같다.



<그림 6> 세 분류기의 k-fold 평균 결과값 ROC 커브 비교

4.2 설명가능한 인공지능 결과 해석

본 장에서는 TabNet 모델 중 가장 높은 성능을 보인 모델의 결과를 바탕으로 전역 설명력과 지역 설명력을 해석하고자 한다.



<그림 7> 전역 설명력

(1) 전역 설명력

<그림 7>는 TabNet 인코더가 제공하는 전역설명력의 출력 값으로 모든 테스트 데이터를 대상으로 각 변수들이 예측에 끼친 영향력을 시각화한 것이다. 우측 범례는 변수의 영향력을 나타낸다. 영향력이 가장 높은 변수는 파란색, 가장 낮은 변수는 붉은색으로 표시되며, 수치 값은 변수별로 중앙에 출력되어 있다. 영향력의 수치는 모든 테스트 데이터의 지역 설명력의 평균값을 통해서 구할 수 있다. 0.78까지의 변수명은 부록 A.2에 제시하였다. 전역설명력에서 영향력이 높은 변수는 0.01 이상의 영향력 수치를 가진 값으로 정의한다. 수치형 변수에서는 ‘자기자본비율’, ‘차입금 의존도’, ‘부채상환 능력’, ‘총자산순이익률’, ‘EBIT대매출액’,

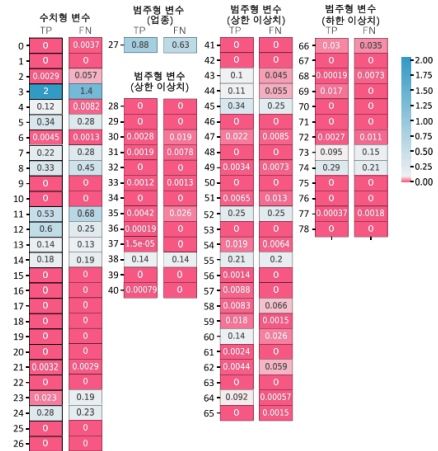
‘매출액영업이익율’, ‘차입금/자기자본’, ‘부채상환능력계수’, ‘영업활동후현금흐름/매출액’ 변수가 예측 값에 주요하게 영향을 끼쳤으며, 이 중 ‘자기자본비율’은 모든 변수를 통틀어 가장 큰 영향을 끼친다. 범주형 변수 중에서는 ‘업종’이 영향력이 컸으며 예외값의 경우 ‘이자보상배율’, ‘법인세비용차감전순이익증가율’이 영향력이 큰 변수이다. 또한 범주형 변수 중 이상치 값들은 상한 이상치의 경우 ‘매출액영업이익율’, ‘순운전자본회전율’, ‘영업활동후현금흐름’, ‘단기차입금’이 영향력이 컸으며, 하한 이상치의 경우 ‘총자산 순이익률’이 중요한 변수로 관찰되었다. 영향력이 높은 변수를 집계해보니 해당 모델의 결과는 안정성, 수익성, 성장성, 활동성의 순서로 높은 것으로 영향력이 해석되었다. 즉, 안정성 54%, 수익성 35%, 성장성 5%, 활동성 4%만큼 모델의 예측 값에 영향을 끼친다고 해석할 수 있다.

(2) 지역 설명력

TabNet 인코더의 결과로 각 케이스의 결과 예측에 변수가 얼마나 영향을 미쳤는지를 수치화한 값을 얻을 수 있다. 이 값은 모든 케이스에 대하여 구할 수 있으며 그 중 3가지 케이스만 시각화하여 <그림 8>에 나타내었다. 케이스 별로 경향성이 비슷한 것을 확인할 수 있다. 영향력의 수치는 케이스 별로 차이가 존재하였지만, 결과에 영향을 끼친 변수가 유사하였다.

(3) TN, FN 케이스 데이터 전역 설명력

TabNet의 지역설명력을 활용하여 TP와 FN으로 분류된 예측값을 따로 집계하여 TP, FN에 해당하는 케이스의 설명력을 계산하였다. 해당 TP, FN 설명력은 <그림 9>와 같다. <그림 9>의 수치값의 범위는 0~2



<그림 8> 3개의 케이스 지역 설명력

이다. TP, FN 데이터에 해당되는 전역설명력 수치값에 대하여 설명력이 큰 기준값을 0.1로 설정하였다. TP와 FN에 해당되는 데이터 중 전역설명력이 0.1 이상의 설명력을 갖는 변수와 해당 변수에 대한 TP case - FN case 값은 <표 4>와 같다. TP로 분류된 케이스와 FN으로 분류된 케이스는 각각 실제 부도인 케이스를 예측에 성공한 것과 실패한 것을 의미한다. 따라서 TP로 분류된 케이스와 FN으로 분류된 케이스의 차이를 통하여 어떤 변수가 부도 예측 실패와



<그림 9> TP, FN 케이스의 설명력 비교

〈표 4〉 TP, FN 케이스 설명력 비교

수치형 변수			
변수	TP	FN	TP-FN
자기자본비율	2.00	1.40	0.60
자기자본순이익율	0.60	0.25	0.35
단기차입금/총자산	0.28	0.23	0.05
총자산순이익률	0.53	0.68	-0.15
이자보상배율	0.34	0.28	0.06
총자본회전율	0.33	0.45	-0.12
순운전자본회전율	0.22	0.28	-0.06
매출액영업이익율	0.14	0.13	0.01
차입금/자기자본	0.12	0.01	0.11
범주형 변수 예외값 변수			
업종	0.88	0.63	0.25
매출액총이익율	0.18	0.19	-0.01
단기차입금/총자산	0.14	0.14	0.00
차입금의존도	0.02	0.19	-0.17
범주형 변수 이상치 상황			
이자보상 배율	0.34	0.25	0.09
차입금/자기자본	0.11	0.05	0.06
부채비율	0.10	0.04	0.05
자기자본순이익율	0.25	0.25	0.00
범주형 변수 이상치 하한			
EBIT대매출액(법인세비용차감전순이익)	0.29	0.21	0.08
법인세비용차감전순이익증가율	0.14	0.02	0.11
매출액총이익율	0.09	0.15	-0.05
EBIT대매출액(법인세비용차감전순이익)	0.21	0.20	0.01

성공에 영향을 미치는지 확인하였다.

차이값의 절대값이 클수록 예측의 오류에 영향을 크게 미치는 변수로 해석한다. <표 4>에서 확인할 수 있듯이, ‘자기자본비율’은 0.6으로 가장 큰 차이를 보였으며 부도를 예측하는데 가장 중요하게 다뤄야할 변수로 해석된다. ‘자기자본순이익율’은 0.350의 차이를 보였으며 이 값은 전체 차이 값들 중 두 번째로 큰 차이를 보인다. ‘업종’은

0.250으로 범주형 변수 중 가장 큰 차이를 보였으며, 전체 차이 중에서는 3번째 차이를 보였다. ‘자기자본비율’, ‘자기자본순이익율’, ‘업종’ 변수 모두 TP로 분류되는데 큰 영향을 끼치는 것으로 파악된다. 반대로 FN으로 오분류될 때 TP로 분류된 케이스보다 영향력 이 큰 수치형 변수는 ‘총자산순이익율’이 있다. 이상치 상황과 하한 중에서는 대체로 유의미한 차이를 보이지 않았으나

〈표 5〉복합샘플링 효과 비교

k-fold #	AUC ↑	F1-Score ↑	FNR ↓	Balanced Accuracy ↑
TabNet				
1	0.8662	0.8953	0.1667	0.8036
2	0.8831	0.8916	0.1722	0.8057
3	0.8577	0.8834	0.1899	0.7800
4	0.8690	0.8499	0.2389	0.7917
Mean	0.8680	0.8801	0.1919	0.7954
TabNet + Hybrid Sampling				
1	0.9045	0.9147	0.1278	0.8362
2	0.8689	0.9073	0.1500	0.7854
3	0.9105	0.9219	0.1117	0.8512
4	0.9186	0.9080	0.1389	0.8350
Mean	0.9006**	0.9130*	0.1321*	0.8270**

* 유의수준 5%에서 통계적으로 유의함

** 유의수준 10%에서 통계적으로 유의함

이상치 하한에 속하는 ‘법인세비용차감전 순이익증가율’이 소폭의 차이를 보였다. TP로 분류된 케이스와 FN으로 분류된 케이스의 유의미한 차이를 보인 변수를 확인해 본 결과 부도인 기업을 부도로 옳게 분류하는데 있어서 안정성과 수익성 그리고 어떠한 업종인지가 예측 결과에 주요하게 영향을 끼친 것으로 파악된다.

4.3 단순 TabNet 대비 샘플링 기법을 포함한 성능비교

TabNet의 효과를 확인한 후 불균형 데이터 문제를 해결하기 위해 복합 샘플링을 적용하였다. 적용 결과는 표 5에서 확인할 수 있다. 복합샘플링과 TabNet 모형을 함께 활용한 모형의 결과가 k-fold 1,3,4에서 모든 평가척도에서 성능 향상을 보였고, k-fold 2에서는 AUC와 balanced Accuracy에서 감소, F1-Score와 FNR에서는 향상을 보였다.

결과적으로 기업의 부도예측에 사용되 는 데이터처럼 정상/부도의 데이터 불균형 문제에 있어, 본 연구에서 제안한 프레임워크 즉, 오버샘플링 기법인 SMOTE와 언더샘플링 기법인 ENN을 합친 복합 샘플링 (hybrid sampling) 방법을 활용할 경우, 단순 TabNet 만을 활용한 결과보다 모든 평가척도에 대해서 평균값은 성능 향상을 보였다. 이를 통하여 불균형 데이터 처리에 복합샘플링 이 효과를 보임을 모델의 성능 향상을 통하여 확인하였다.

IV. 결 론

본 연구에서는 기존의 전처리 방법은 데이터 손실이 존재한다고 판단하여 기업 재무 데이터의 이상치와 예외값을 새로운 방식으로 전처리하였다. 예외값과 이상치 상향과 하한의 변환을 통하여 새로운 범주형

변수를 생성하여 학습에 활용하였다. 이러한 전처리 과정을 통한 학습 데이터는 정형화 데이터의 성질을 강하게 보인다.

또한 정형화 데이터의 강점이 있는 딥러닝 기반의 모델인 TabNet의 인코더를 사용하여 기존 연구들에 많이 활용된 로지스틱 회귀와 다층퍼셉트론 대비 우수한 분류 성능을 확인하였다. 하지만 데이터의 종속변수 값인 부도여부가 불균형한 데이터로 같은 비율로 테스트(test)와 학습(train), 검증(validation) 데이터 셋을 나누었지만 k-fold에 따른 성능 차이를 보였다. 이러한 데이터 불균형 문제를 다루고자 복합샘플링 기법을 사용하였고 분류 성능 향상의 효과를 확인하였다.

TabNet의 설명가능력을 통하여 결과 분석을 한 결과 부도여부 예측에 안정성 지표와 수익성 지표의 중요성을 확인하였고 각 케이스별로 설명력을 제공할 수 있다는 장점 또한 확인하였다.

본 연구에서 제안한 프레임워크에 대해 기업의 부도를 예측해야하는 실무적 관점에서 고려해 볼 때 다음과 같은 기여도가 기대된다. 첫째, 우수한 예측성능과 함께 결과 도출 과정에 대한 설명을 제공할 수 있고 둘째, 부도예측에 사용되는 데이터의 불균형 문제를 해소할 수 있다. 끝으로 TabNet은 데이터 전처리 없이 end-to-end 학습이 가능하여, 개발이후 모델의 정기적 개선(tuning) 작업을 용이하게 한다.

본 연구에서는 모델의 예측성능 입증을 위해 로지스틱 회귀 및 다층 퍼셉트론을 중심으로 비교분석 하였으나, 향후 연구에서는 GBM(Gradient Boost Machine) 및 XGBoost(Extreme Gradient Boosting) 등 예측결과에 대한 설명이 가능한 트리기반 모델들과의

예측성능 및 변수 설명력을 비교분석하고자 한다. 또한 예측결과에 따라 TabNet이 제시하는 설명결과 대비 샵(SHAP)과 라임(LIME) 같은 대리모델(surrogate model)이 제시하는 변수 중요도와의 비교 연구를 통해 TabNet의 XAI 성능을 추가 검증하고자 한다.

논문접수일: 2022년 08월 03일

논문수정일: 2023년 08월 11일

게재확정일: 2023년 09월 01일

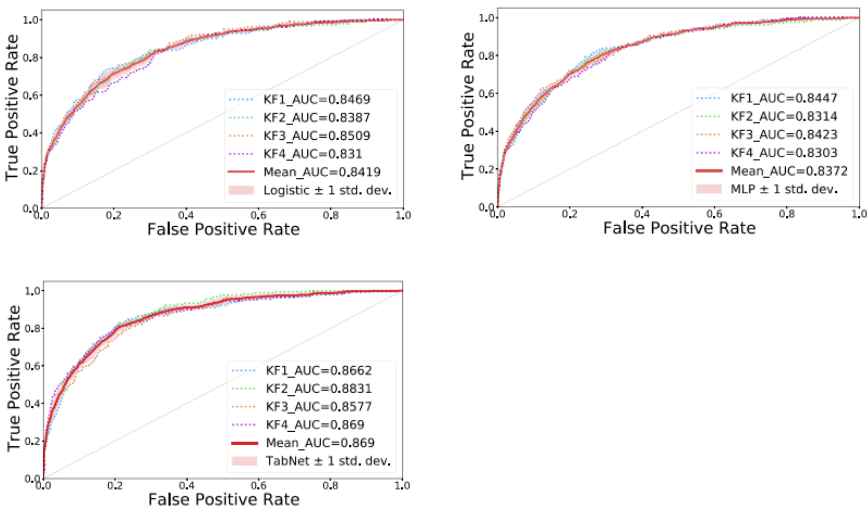
참 고 문 헌

- 김소정, 이균희. (2022). 심층신경망의 설명 가능성과 하이퍼파라미터 특성에 관한 연구-중소기업 신용평가를 중심으로-. *중소기업금융연구*, 42(1), 3-37.
- 김진선, 최영문. (2006). 재무변수 및 주식가격 변수를 이용한 회사채 신용등급 예측모형의 개발. *회계, 세무와 감사 연구*, 43, 185-217.
- 김형준, 류두진, 조훈. (2019). 기업부도 예측과 기계학습. *금융공학연구*, 18(3), 131-152.
- 민재형, 이영찬. (2005). Support Vector Machine을 이용한 부도예측모형의 개발. *한국경영과학회지*, 30(1), 55-74.
- 안태일, 김종민. (2002). 주식투자자들의 재무정보 중요성 인식에 관한 연구. *세무회계연구*, 10, 209-233.
- 오우석, 김진화. (2017). 인공지능기법을 이용한 기업부도 예측. *산업융합연구*, 15, 17-32.
- 조준희, 강부식. (2007). 코스닥기업의 도산 예측모형에 관한 연구. *산업경제연구*, 20(1), 141-160.

- Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138-52160.
- Arik, S. O., & Pfister, T. (2021). TabNet: Attentive interpretable tabular learning. *AAAI*, 35, 6679-6687.
- Barandela, R., Valdivinos, R. M., Sanchez, J. S., & Ferri, F. J. (2004). The imbalanced training sample problem: Under or over sampling? In *Joint IAPR International Workshops on Structural, Syntactic, and Statistical Pattern Recognition (SSPR/SPR'04), Lecture Notes in Computer Science*, 3138, 806-814.
- Batista, G. EAPA., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter*, 6(1), 20-29.
- Gehring, J., Auli, M., Grangier, D., Yarats, D., & Dauphin, Y. N. (2017). Convolutional sequence to sequence learning. *International Conference on Machine Learning, PMLR*, 1243-1252.
- Goodfellow, I., Yoshua, B., & Aaron, C. (2016). *Deep learning*, Massachusetts: The MIT Press.
- Lundberg, S. M., & Lee, S. J. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
- Martins, A., & Astudillo, R. (2016). From softmax to sparsemax: A sparse model of attention and multi-label classification. *International Conference on Machine Learning, PMLR*, 1614-1623.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18, 109-131.
- Zhen, W. (2019). Machine learning applications in finance: Some case studies. *Ph. D.'s thesis, Department of Computing, Imperial College London*.

별첨

A.1. 모델 성능 비교 ROC 곡선



A.2. 데이터셋 독립 변수명

〈표 A.2.1〉 데이터셋 독립변수명: 수치형 변수와 업종

수치형 변수			
변수	카테고리	항목명	산식
0	유동성	유동비율	유동자산 / 유동부채
1	유동성	당좌비율	당좌자산 / 유동부채
2	안정성	부채비율	부채총액 / 자기자본
3	안정성	자기자본비율	자기자본 / 위험가중자산
4	안정성	차입금/자기자본	
5	부채상환능력	이자보상배율	영업이익 / 금융비용(이자비용)
6	활동성	총자산회전율	매출액 / 총자산
7	활동성	순운전자본회전율	매출액 / 순운전자본
8	활동성	총자본회전율	매출액 / 총자본
9	활동성	매출채권회전율	순외상매출액 / 연평균 매출채권 잔액
10	수익성	매출액순이익률	순이익 / 매출액
11	수익성	총자산이익률	당기순이익 / 총자산
12	수익성	자기자본순이익률	당기순이익 / 자기자본
13	수익성	매출액영업이익률	영업이익(영업외손익 제외) / 매출액
14	수익성	매출액총이익률	총이익(매출원가 제외) / 매출액
15	수익성	EBIT대매출액	법인세비용차감전순이익 / 매출액

16	성장성	매출액증가율	당기 - 전기 / 전기 매출액
17	성장성	영업이익증가율	현 - 전 / 전 회계연도 영업이익
18	성장성	순이익증가율	당기 - 전기 / 전기의 당기순이익
19	성장성	총자산증가율	당기 - 전기 / 전기말 총자산
20	성장성	법인세비용차감전순이익증가율	현 - 전 / 전 회계연도 법인세비용차감전순이익
21	부채상환능력	부채상환계수	EBITDA / (단기차입금 + 금융비용)
22	현금흐름	영업활동후현금흐름/매출액	-
23	안정성	차입금의존도	차입금 / 총자본
24	안정성	단기차입금/총자산	-
25	부채상환능력	차입금/EBITDA	-
26	부채상환능력	EBITDA대금융비용	EBITDA / 금융비용
범주형 변수			
27	업종	업종	

〈표 A.2.2〉 데이터셋 독립변수명: 범주형 변수 예외값

범주형 예외값		
28	안정성	자기자본비율
29	부채상환능력	이자보상배율
30	성장성	매출액증가율
31	성장성	영업이익증가율
32	성장성	순이익증가율
33	성장성	총자산증가율
34	성장성	법인세비용차감전순이익증가율
35	부채상환능력	부채상환계수
36	현금흐름	영업활동후현금흐름/매출액
37	안정성	차입금의존도
38	안정성	단기차입금/총자산
39	부채상환능력	차입금/EBITDA
40	부채상환능력	EBITDA2대금융비용

〈표 A.2.3〉 데이터셋 독립변수명: 범주형 변수 상한 이상치

범주형 예외값		
41	유동성	유도비율
42	유동성	당좌비율
43	안정성	부채비율
44	안정성	차입금/자기자본
45	부채상환능력	이자보상배율
46	활동성	총자산회전율
47	활동성	순운전자본회전율
48	활동성	총자본회전율
49	활동성	매출채권회전율
50	수익성	매출액순이익율
51	수익성	총자산순이익율
52	수익성	자기자본순이익율

53	수익성	매출액영업이익율
54	수익성	EBIT대매출액(법인세비용차감전순이익)
55	성장성	매출액증가율
56	성장성	영업이익증가율
57	성장성	순이익증가율
58	성장성	총자산증가율
59	성장성	법인세비용차감전순이익증가율
60	부채상환능력	부채상환계수
61	현금흐름	영업활동후현금흐름/매출액
62	안정성	차입금의존도
63	안정성	단기차입금/총자산
64	부채상환능력	차입금/EBITDA
65	부채상환능력	EBITDA대금융비용

〈표 A.2.4〉 데이터셋 독립변수명: 범주형 변수 하한 이상치

범주형 예외값		
66	안정성	자기자본비율
67	부채상환능력	이자보상배율
68	수익성	매출액순이익율
69	수익성	총자산순이익율
70	수익성	자기자본순이익율
71	수익성	매출액영업이익율
72	수익성	매출액총이익율
73	수익성	EBIT대매출액(법인세비용차감전순이익)
74	수익성	영업이익증가율
75	성장성	순이익증가율
76	성장성	법인세비용차감전순이익증가율
77	부채상환능력	부채상환계수
78	현금흐름	영업활동후현금흐름/매출액

