Araştırma Makalesi/Research Article *DOU Dergi*, 2023, *24*(2), 589-604

DOĞUŞ ÜNİVERSİTESİ DERGİSİ

DOGUS UNIVERSITY JOURNAL

e-ISSN: 1308-6979

https://dergipark.org.tr/tr/pub/doujournal

FINANCIAL DISTRESS PREDICTION FROM TIME SERIES DATA USING XGBOOST: BIST100 OF BORSA ISTANBUL

XGBOOST İLE ZAMAN SERİSİ VERİLERİNDEN FİNANSAL BAŞARISIZLIK TAHMİNİ: BORSA ISTANBUL BİST100

Umut ENGİN[1], Salih DURER[2]

Abstract: This study utilized financial and non-financial data from 233 companies listed in the Borsa Istanbul BIST SINAI Index from 2010 to 2020. The XGBOOST machine learning algorithm was employed to predict whether these companies would encounter financial distress. The machine was trained using supervised learning, with 80% of the data used for training and 20% for testing purposes. Financial ratios were utilized as independent variables in predicting financial distress. The 25 financial ratios can be categorized into four main headings: Liquidity, Financial Structure, Activity, and Profitability Ratios. Furthermore, the model allowed for individual analysis of each company. In predicting whether companies would experience financial distress, the maximum F1 score (85.1%), recall (84.5%), precision (85.7%), and accuracy (91.6%) were achieved.

Keywords: XGBoost, BIST100, Financial Distress, Prediction, Stock, BIST SINAI

JEL: C53, G17, E44

Öz: *Bu çalışmada, Borsa İstanbul BIST SINAI Endeksi'nde yer alan 233 şirketin 2010'dan 2020'ye kadar olan finansal ve finansal olmayan verileri kullanılmıştır. Bu firmaların finansal sıkıntıya girip girmeyeceklerini tahmin etmek için bir makine öğrenmesi algoritması olan XGBOOST kullanıldı. Denetimli öğrenme şeklinde makine eğitildi, verinin %80' i eğitim, %20' si ise test için kullanıldı. Finansal sıkıntıyı tahmin ederken finansal oranlar bağımsız değişkenler olarak kullandı. 25 adet finansal oranı 4 ana başlık altında toplayabiliriz: Likidite, Finansal Yapı, Faaliyet ve Karlılık Oranları. Ayrıca model, firmaları tek tek analiz etmeyi sağladı. Şirketlerin finansal sıkıntıya girip girmeyeceklerini tahminlemede maksimum F1 puanı (%85.1), hatırlama (%84.5), kesinlik (%85.7) ve doğruluk (%91.6) elde edildi.*

Anahtar Kelimeler: *XGBoost, BIST100, Finansal Sıkıntı, Tahmin, Hisse Senedi, BIST SINAI*

1. Introduction

According to the literature, the term "financial distress" is commonly used interchangeably with concepts such as bankruptcy, failure, default, and insolvency. These concepts refer to the loss of a business's ability and capacity to sustain its operations (Altman & Hotchkiss, 2010). In financial markets, various factors such as unemployment data, economic indicators like inflation rate, political decisions, evolving global conditions, financial choices, and expectations of their impact on these variables, contribute to the periodic changes in the stock market climate (Abu-Mostafa & Atiya, 1996). The dynamic and non-linear nature of financial success and

[1] Yıldız Teknik Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü umutengin22@gmail.com, ORCID: 0000- 0003-4665-8169

[2] Yıldız Teknik Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü; durer@yildiz.edu.tr, ORCID: 0000-0003-2575-2842

Geliş/Received: 18-01-2023; Kabul/Accepted: 27-06-2023

Atıf bilgisi: Engin, U. ve Durer, S. (2023). Financial distress prediction from time series data using xgboost: BIST100 of Borsa Istanbul. *Doğuş Üniversitesi Dergisi*, *24*(2), 589-604. DOI: 10.31671/doujournal.1238432

failure also influences stock prices. Moreover, accurate predictions of financial distress are crucial for companies involved in financial relationships with each other and for investors to make informed investment decisions. Numerous methods exist today for determining multivariate financial distress indicators in companies. Most of these methods rely on structured data and technical indicators (Marcek, 2004). Generally, technical and fundamental analyses are employed in the literature to forecast future stock prices. Technical analysis utilizes historical stock prices, including daily, weekly, and monthly data, while fundamental analysis requires economic indicators such as exchange rates, inflation rates, interest rates, and unemployment rates. For portfolio management, obtaining company-specific financial distress data and making precise predictions of future stock prices can potentially lead to higher profits or the mitigation/elimination of potential losses.

In recent years, various machine learning methods, including decision tree modeling, support vector machines, artificial neural networks, ARIMA, time series analysis, linear regression, and Markov chain/process, have been employed in financial data processing. However, machine learning algorithms have proven to yield the highest precision and accuracy in predicting financial distress (Alkhatib et al., 2013; Altan & Demirci, 2022). Among these algorithms, the XGBoost algorithm, based on decision trees, has shown promise. It is an expanded and enhanced version of the gradient-boosted decision trees algorithm, capable of producing reliable solutions to many estimation problems (Chen et al., 2015; Chen & Guestrin, 2016; Ustalı et al., 2021).

In other words, financial distress refers to the inability of businesses to meet their financial obligations, leading to potential bankruptcy. It is crucial to develop models that can predict and prevent such situations, as they have significant negative effects on businesses and the economy as a whole. Financial failure not only affects the concerned business but also impacts stakeholders such as banks, investors, and suppliers. The increase in failed businesses results in unemployment and economic decline. Therefore, financial failure is both a micro-level concept and a sociological phenomenon. Predicting and mitigating financial failure is vital for businesses, and various models, including XGBOOST machine learning, have been developed for this purpose.

This study aimed to predict the financial distress of companies listed in the BIST100 SINAI index of Borsa Istanbul by conducting a historical analysis of data from the past ten years (2010-2020) using the Gradient Boosting Trees Algorithm. The following sections provide a detailed description and preparation of the dataset, the proposed methodology, and the results obtained.

2. Material and Method

The dataset used in the study is collected from the financial data of 233 companies in BIST100 SINAI of Borsa Istanbul, each described by 25 independent financial variables (from X1 to X25) over the period from 2010 to 2020 labeled based on the financial distress of each company per year where zero (0) represents a non-financially distressed company and one (1) is considered financially distressed. . In order to understand whether companies are in financial distress, data based on financial statements and data based on non-financial statements were used. For data not based on financial statements, stock market statements of companies were examined (KAP, 2022a). Financial distress criterias are summarized in Table 1. This is supervised learning methodology in machine learning. Supervised learning aims to predict expected effects using dataset labeled by humans (Goecks, Jalili et al., 2020). In a supervised neural network, the output of the input is already known, the estimated output of the neural network is compared with the actual output. Based on the error, the parameters are changed and then fed back to the neural network, the supervised neural network is used in the feedforward neural network (Coelho and Richert, 2015).

According to Civan and Dayı (2014), financial failure refers to the situation where the operating income fails to cover all costs, including capital costs. Baş and Çakmak (2012) define financial failure as the inability of a business to meet its debt obligations, resulting from a deteriorating financial structure and ongoing distress. Aktaş, Doğanay, and Yıldız (2003) take a comprehensive perspective on financial failure, considering factors such as production cessation, a loss of 10% of business assets, debts exceeding total assets, three consecutive years of losses, difficulties in debt repayment, and loss of capital. Kılıç and Seyrek (2012) define financial failure as the failure of businesses due to their policies and decisions, leading to the inability to achieve their goals. Aktaş (1993) considers financial failure as three consecutive years of losses and the cessation of production due to a financial crisis. Yıldız (1999) defines financial failure as bankruptcy, the loss of half of the capital and 10% of the total assets, three consecutive years of losses, difficulties in debt repayment, cessation of production, and debts surpassing assets. Keskin (2002) defines financial failure as bankruptcy, while Torun (2007) includes factors such as the closure of the stock market, suspension of operations, and two or more consecutive years of losses. Özdemir (2011) considers financial failure as incurring losses in the last two years and experiencing depreciation compared to the general stock exchange index where the stock is traded.

Table 1. Financial Distress Criterias

BASED ON FINANCIAL STATEMENTS	NOT BASED ON FINANCIAL STATEMENTS
Loss for 3 Years in a Row	Independent Audit Report Containing Adverse Opinion or Avoiding Expressing Opinion
Losing 10% of its assets	Partial or Total Suspension of Activities or Become Impossible
Losing 50% of Capital	Financial Fixed Asset Sale
Total Debt Exceeds Total Assets	Enforcement Follow-ups
	Bankruptcy/Bankruptcy Postponement
	Delist/Trading Ineligibility
	Starting to Trade Structured Debt Instruments

If a company meets at least one of the aforementioned criteria, it is marked as financially distressed in the respective year. Table 2 presents a partial list generated as a result of these markings.

Table 2. Sample of Financial Distress Companies

.	Years																						
Company Codes	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**Grand Total**
ABANA									1									1					2
ADEL																	1			1			2
ADESE																					1		1
AEFES															1								1
AFYON												1	1						1				3
AKCNS						1																	1
AKENR						1													1	1	1		4
AKPAZ																			1				1
AKSA						1		1															2
AKSEN																		1					1
AKSUE												1											1
ALKA															1								1
ALMAD																							
ANELE												1											1
APEKS	1																						1
ARAT								1															1
ARBUL																	1	1					2
ARCLK						1																	1
ARTOG																1	2						3
ASCEL																			1				1
ASUZU										1													1
ATAC															1								1
ATEKS										1													1
ATPET																				1			1
AVOD																			1				1
AVTUR								1			1												2
AYDEM																					1		1
AYEN							1	1										1	1				4
AYGAZ								1															1
BAGFS				1						1													2
BALAT																	1						1
BERDN														1									1
BEYAZ										1				1	1	1			1				5
BFREN									2				1										3
BIZIM										1													1
BJKAS							1		1		1	1	1								1		6
BLCYT										2													2
BMEKS																				1			1
BMELK																		1	1				2
BOSSA															1		2						3
BRKO																	1				1	1	3
BRMEN																				1	2		3
BSOKE																				1	1		2
BTCIM																				1	1		2
BURCE								1															1
BURVA								1						1									2
BYSAN									1			1											2
CASA																						1	1
CBSBO														1									1
CCOLA					1																		1
CELHA													1										1
CEMAS															1			1		1			3
CLEBI				2				1				1	1										5
COMDO																	1						1
CRFSA							2										1	2	1	1	1		8
DARDL		1			1				1	1	2		1	2		1					1		11

Consequently, the dataset comprises a total of 2563 company-year observations, with 359 observations classified as financially distressed and 2204 as non-financially distressed, indicating a class imbalance that requires attention.

Figure 1 displays the distribution of companies' states, whether distressed or non-distressed, spanning the period from 2010 to 2020. It is noteworthy that, despite being the minority class, the year 2016 witnessed the highest number of financially distressed companies, whereas in 2020, only 15 companies experienced financial distress. The subsequent subsections elucidate the steps taken for data cleaning, data preparation, addressing class imbalance, and feature engineering, accompanied by exploratory data analysis (EDA) to gain further insights into the dataset prior to its utilization in the XGBoost model.

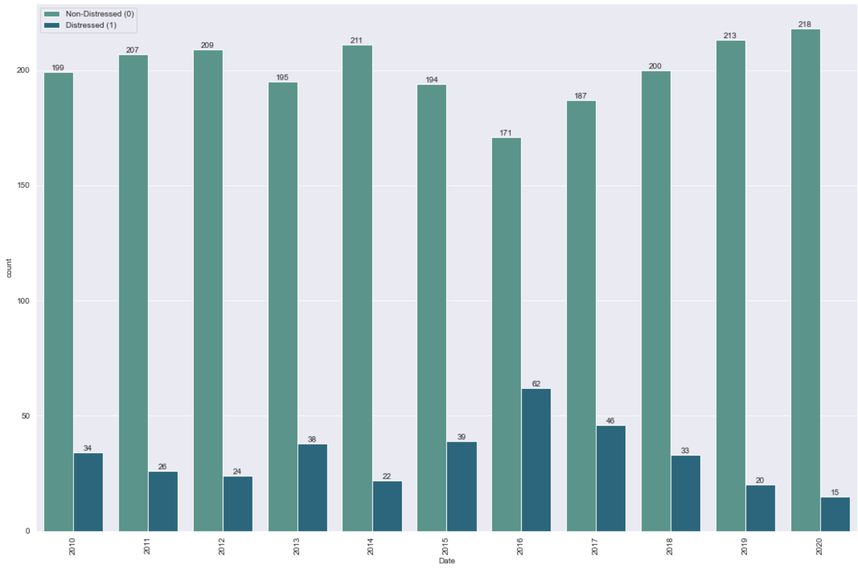

Figure 1. Distribution of the companies' state over the given time period

During the process of data cleaning and preparation, the initial dataset was obtained from the financial statements of companies, including the relevant performance variables used to calculate the independent variables (X1-X25). Nevertheless, a significant step in data preparation was undertaken, involving the transformation of the dataset from a wide format to a long format. This involved representing the different time values (years) in a single column, which was repeated for each company, as depicted in Figure 2.

Current Assets	2000/12	2001/12	2002/12	2003/12	2004/12
.V.O.D Gıda Ve Tarım	0.0	0.0	0.000000e+00	0.0	0.0
Acıpayam Selüloz	0.0	0.0	0.000000e+00	0.0	0.0
Adel Kalemcilik	8824746.0	13144768.0	1.843350e+07	19770782.0	22830241.0
Afyon Çimento	3669289.0	3631102.0	9.142103e+06	10473196.0	14885721.0
Akçansa	51592192.0	72209117.0	1.112380e+08	143235957.0	205628123.0

(a)

	Company	Current Assets	Date	Sheet1
0	ABANA	Abana Elektromekanik Sanayii Ve Ticaret A.Ş.	2000/12	2.201067e+06
1	ABANA	Abana Elektromekanik Sanayii Ve Ticaret A.Ş.	2001/12	1.694979e+06
2	ABANA	Abana Elektromekanik Sanayii Ve Ticaret A.Ş.	2002/12	2.292213e+06
3	ABANA	Abana Elektromekanik Sanayii Ve Ticaret A.Ş.	2003/12	1.690793e+06
4	ABANA	Abana Elektromekanik Sanayii Ve Ticaret A.Ş.	2004/12	7.565670e+05
5	ABANA	Abana Elektromekanik Sanayii Ve Ticaret A.Ş.	2005/12	6.882530e+05

(b)

Figure 2. An example snapshot of the data frame in (a) wide format, (b) long format

One important observation is the presence of certain features that exhibit only zero values across different classes. This characteristic poses a challenge to the model's ability to accurately classify or distinguish such instances. Addressing such records may require domain expertise or additional data collection efforts. However, for the purpose of the current study, these observations were removed from the "non-distressed" class, retaining only those belonging to the "distressed" class.

Following this data preprocessing step, the distribution of the data classes was as follows: 1725 observations belonged to the non-distressed (0) class, while 359 observations belonged to the distressed (1) class, as depicted in Figure 3.

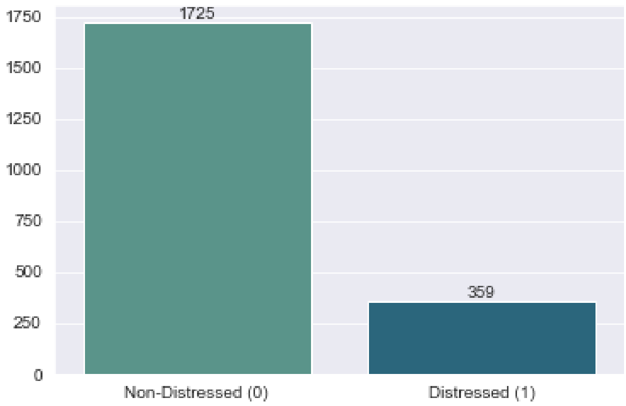

Figure 3. Classes distribution after data cleaning

Next, in the majority of models created to predict the financial failure of businesses, commonly used and deemed significant ratios have been employed as independent variables to forecast failure. In this study, financial ratios were utilized as independent variables in the constructed models. There exist numerous financial ratios for use in financial analysis. However, from among these ratios, those frequently employed in studies related to predicting financial failure and acknowledged as significant for businesses were selected. A total of 25 different financial ratios from four distinct categories were utilized in the analyses. It is observed that in most of the financial failure prediction studies conducted thus far, financial ratios derived from financial statements are employed as independent variables. Although a large number of financial ratios can be calculated theoretically to assess the financial conditions of businesses, the ratios regarded as important and commonly used in the literature are selected as the independent variables in the study. Therefore, in this study, financial ratios obtained from the financial statements of businesses, which have reached a consensus and are considered important in the literature, were chosen as independent variables and utilized. The analyses conducted in this study incorporated 25 financial ratios utilized in the analyses of Akkaya et al. (2009), Torun (2007), Kılıç and Seyrek (2012), Shirata (1998), Aktaş et al. (2003), Albayrak and Yılmaz (2009), Li and Sun (2011), and Yakut and Elmas (2013). The ratios used in the analyses for predicting financial failure were categorized and presented in Table 3.

Considered as a reliable diagnostic method for data analysis, Figure 4 illustrates the correlation matrix, which displays the degree of linear association (correlation) between the data variables. Positive correlation is represented by red colors, while negative correlation is indicated by blue colors, with the intensity of the colors reflecting the strength of the correlation.

One important finding from the correlation matrix is the presence of collinearity among certain variables or features. For instance, X1 and X2 exhibit a high positive correlation. Conversely, X20 and X14 show a negative correlation, suggesting potential redundancy in the feature space, where these variables may provide similar information or lead to model overfitting.

While the correlation matrix is an important data analysis tool that warrants attention, it solely captures linear correlations between variables. Therefore, during model training experiments, different approaches will be attempted to address this collinearity issue. However, it is important to note that collinearity is not always a problem, as there may exist nonlinear relationships between variables.

Before addressing the issue of class imbalance, a random portion of the data was set aside for model evaluation. The dataset was divided into training and testing sets, with an 80%-20% ratio allocation, respectively.

Table 3. Independent variables formulas

Independent Variable	Formula
Liquidity Ratios	
X1	Current Assets / Short Term Liabilities
X2	(Current Assets -Stocks) / Short Term Liabilities
X3	Cash and Cash Equivalents / Short Term Liabilities
X4	Stocks / Total Assets
X5	Commercial Debts / Total Assets
Financial Structure Ratios	
X6	Short Term Liabilities / Equities
X7	Total Debt / Equities
X8	Fixed Assets / Equities
X9	Current Assets / Total Assets
X10	Short Term Liabilities / Total Assets
X11	Long Term Liabilities / Total Assets
X12	Total Debts / Total Assets
Activity Rates	
X13	Revenue / Trade Receivables
X14	365 / Receivable Turnover Rate
X15	Revenue / Cash and Cash Equivalents
X16	Cost of sales / Average Stocks
X17	Revenue / Current Assets
X18	Net Sales / Total Assets
X19	Revenue / Equities
Profitability Ratios	

X20	Gross Profit / Revenue
X21	Core Operating Profit / Revenue
X22	Net Profit / Revenue
X23	Net Profit / Equities
X24	Main Operating Profit Loss / Total Resources
X25	Net profit / Total Assets

Addressing data imbalance is a prevalent challenge encountered in many real-world classification datasets. This issue arises when the dataset exhibits skewed or biased class proportions, wherein classes with a substantial representation are referred to as majority classes, while those with a comparatively smaller representation are referred to as minority classes.

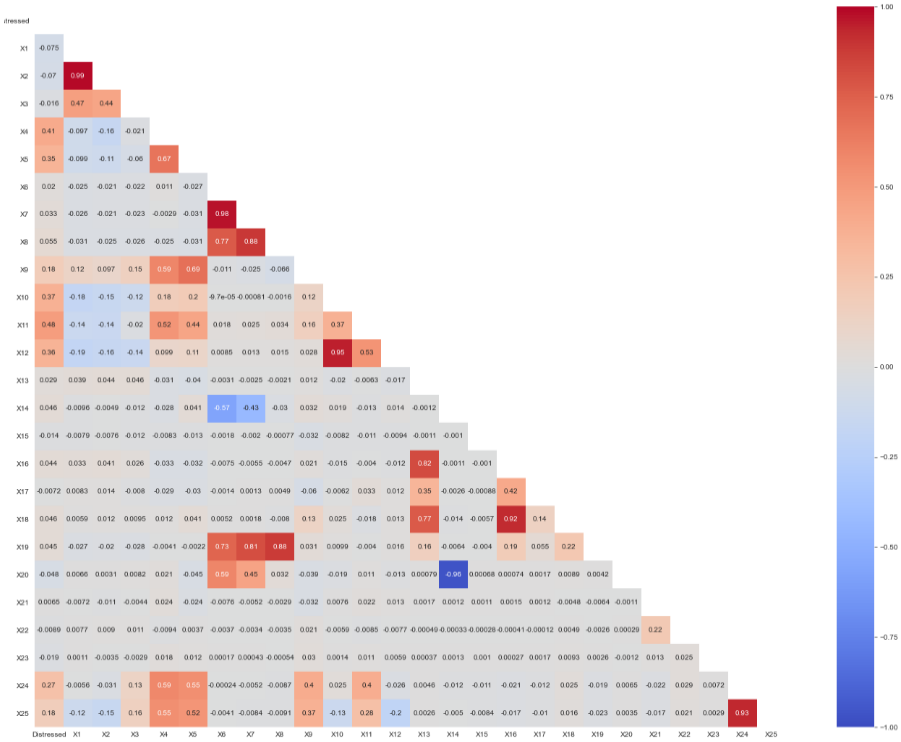

Figure 4. The correlation matrix between data variables

In the dataset, as previously explained, there is a class imbalance issue where the majority class of "non-distressed" companies accounts for 86% of the original dataset. This class imbalance poses a challenge because it can be misleading when evaluating the model's accuracy. The model may predict the majority class accurately but struggle to predict the minority class observations, yet still achieve a high accuracy score.

Given that the primary concern with the dataset is to prioritize the classification of financially distressed companies, it is crucial to address the class imbalance before training the model. Therefore, two common techniques for handling class imbalance were employed:

Oversampling the minority class: This technique involves resampling the minority class, which means creating additional samples/observations by duplicating random records from the minority class. Although this approach may lead to overfitting, it is still a viable option.

Synthetic Minority Oversampling Technique (SMOTE): SMOTE is a type of data augmentation that generates synthetic data points based on the original data. Unlike duplicating records, SMOTE creates new data points that are slightly different from the existing ones. One significant advantage of SMOTE is that it avoids generating exact duplicates.

It's worth noting that class imbalance is only addressed in the training set and not the test set to prevent information leakage. Both techniques aim to increase the proportions of the minority class to match those of the majority class, as depicted in Figure 5.

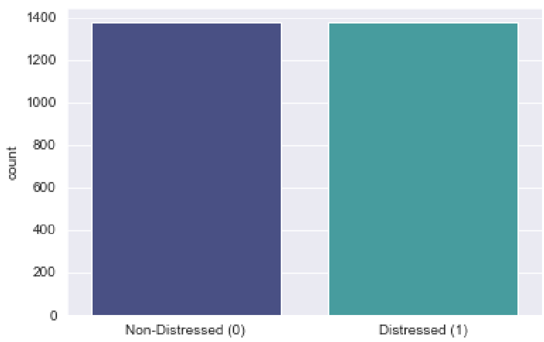

Figure 5. Class distribution after handling imbalance class problem

In the feature engineering phase, special attention is required for the time factor due to the nature of the data, which can be regarded as a time series analysis problem. Past variables may contain important information about the future or have an impact on it. Our dataset includes a series of financial variables spanning 10 years for each company. To address this, lag and rolling features were added for each company group, as explained below.

Lag features involve incorporating the value of a previous timestamp as a new feature at the current timestamp. In Python, this is achieved using the shift() function from the Pandas library. It shifts the index by a specified number of steps, either forward or backward. Lag shifting was applied with periods of 1 and 2 for each feature.

Rolling windows are employed to smooth the data by computing statistics using a moving window. The rolling() function in Python with a window size of 3 was utilized to calculate the average per company over time. However, only the interaction terms between feature variables were computed using the PolynomialFeatures() function from the scikit-learn library, aiming to address collinearity issues identified through the correlation matrix. Redundant features were dropped accordingly.

Many machine learning models have hyperparameters that need to be specified by the model builder to determine the training behavior of the model. In the XGBoost modeling phase of the study, the XGBoost machine learning algorithm was implemented using the XGBoost library in Python's scikit-learn. Six experiments were conducted to train the XGBoost model, examining the impact of each preprocessing step on the model's performance. Eventually, the best model was selected based on the achieved metrics.

The following parameters are the model parameters determined to find the most suitable parameters for our data.

Alpha is a regularization term. It is used to prevent the model from overfitting to the training data (e.g., to compensate for overfitting to the training data and not performing well with the test data). It can be said that as alpha increases, there is less "overfitting" occurring.

Gamma is a regularization parameter. It is a value used to control when we stop splitting the tree leaves in the tree model (decision tree). As the tree depth increases, its value also increases Raschka, S. (2015).

Table 4. The best XGBoost model hyper-parameters for each experiment

Experiment no.	Model Hyper-parameters					
	max_ depth	learning_ rate	n_ estimators	gamma	reg_ alpha	colsample _bytree
1) Raw data without any preprocessing	10	0.06	800	2	0.8	0.06
2) Applying data cleaning	8	0.06	800	2	0.8	0.06
3) Data cleaned+ Oversampling	12	0.06	800	1	0.8	0.09
4) Data cleaned + SMOTE	12	0.09	1000	2	2	0.06
5) Data cleaned + SMOTE + Lagging & Rolling features	15	0.04	1000	0.3	2	0.04
6) Data	13	0.04	1000	0.3	2	0.04

cleaned + SMOTE + Lagging & Rolling + Interaction features	

As previously mentioned, relying solely on accuracy is not suitable when working with a dataset that exhibits class imbalance. Therefore, the model's performance will be evaluated using additional metrics such as the confusion matrix, F1-score, precision, and recall. This comprehensive evaluation aims to ensure that the chosen model performs optimally. In the following section, we will present the results obtained from each experiment and engage in a discussion regarding these outcomes.

3. Results and Discussion

The results of the XGBoost models trained for each experiment and tested on the holdout test set are summarized in Table 5. Additionally, Figure 6 presents the corresponding confusion matrices to provide visual representation and better understanding of the results.

Table 5. Results summary of the five conducted training experiments

Experiment no.	Evaluation metric			
	F1-score (%)	Recall (%)	Precision (%)	Accuracy (%)
(1)	63.01	59.8	81.7	87.9
(2)	79.7	74.6	90.4	90.6
(3)	80.5	76.6	87.3	90.4
(4)	77.02	79.8	75.1	85.6
(5)	85.1	84.5	85.7	91.6

Analyzing these results, it is evident that the model's performance improves progressively with each step, particularly in terms of the recall metric. However, upon closer examination, it becomes apparent that the fourth experiment outperforms the third experiment solely in predicting the majority class ("non-distressed"), while failing to effectively learn the minority class ("distressed"). In this problem, the key challenge lies in training the model using various preprocessing techniques until it achieves satisfactory performance in predicting both classes.

The best overall performance was observed in the fifth training experiment, where the model achieved a maximum F1 score of 85.1%, recall of 84.5%, precision of 85.7%, and accuracy of 91.6%. This experiment successfully balanced the prediction accuracy for both the majority and minority classes.

To provide further insight, a sample evaluation of the five experiments' results on the holdout test set, including the true and predicted labels, was conducted and compared.

The F1 score is a measure of accuracy for the model in academic language. Its mathematical formula is as follows:

F1 = 2 * (precision * recall) / (precision + recall)

Precision and recall are defined by the following formulas:

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

TP represents the number of true positives, which are the points that are truly classified as positive by the model (correctly identified as successful).

TN represents the number of true negatives. It is the complement of TP, assuming that if the failed class is considered positive and the non-failed class is considered negative, TN represents the number of non-failed points. The aim here is to ensure that the model makes correct predictions.

FP (false positive) and FN (false negative) are the number of data points that belong to one class but are predicted by the model as being in the other class.(Figure 6)

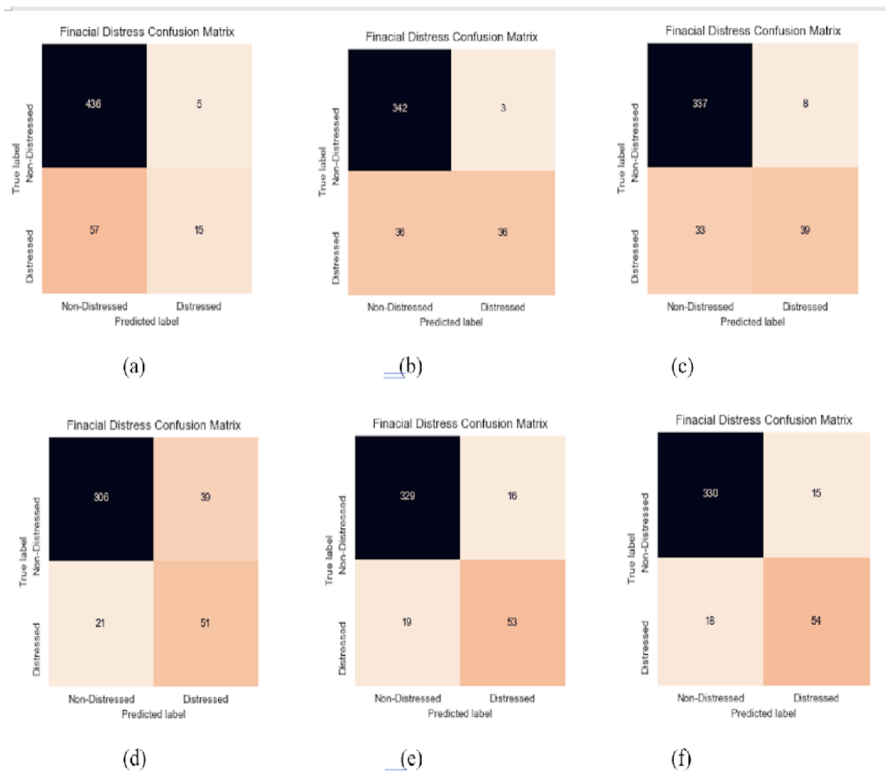

Figure 6. Confusion matrices of XGBoost models performance on the test set for each experiment: (a) 1st- on original data, (b) 2nd- after data cleaning, (c) 3rd- after random oversampling, (d) 4th- after applying SMOTE, (e)5th- time series feature engineering, (f) 6th- polynomial features derived.

Based on the results of the first modeling experiment conducted on the holdout test set (not presented in the tables), it was determined that all companies were financially non-distressed in both predicted and actual status. However, EKIZ and TUKAS were distressed despite being predicted as financially non-distressed. In the second modeling experiment on the holdout test set (not presented in the tables), it was found that all companies, except BRMEN and QUAGR, were financially non-distressed in both predicted and actual status. BRMEN and QUAGR were financially distressed in both predicted and actual status. Additionally, DESA and EMNIS were distressed even though they were predicted as financially non-distressed. The results of the third modeling experiment on the holdout test set (not presented in the tables) indicated that all companies, except VKING and TUCLK, were financially non-distressed in both predicted and actual status. VKING and TUCLK were financially distressed in both predicted and actual status. FRIGO and MNDRS were also distressed, despite being predicted as financially non-distressed. In the fourth modeling experiment on the holdout test set (not presented in the tables), it was observed that all companies, except DGNMO, FRIGO, and ORCAY, were financially non-distressed in both predicted and actual status. DGNMO, FRIGO, and ORCAY were found to be financially distressed in both predicted and actual status. IZFAS and PINSU were also non-distressed, despite being predicted as financially distressed. The results of the fifth modeling experiment on the holdout test set (not presented in the tables) showed that all companies, except ABANA, BMSCH, and ROYAL, were financially non-distressed in both predicted and actual status. ABANA, BMSCH, and ROYAL were financially distressed in both predicted and actual status. KRTEK, on the other hand, was non-distressed, even though it was predicted as financially distressed. Conversely, for MEGAP, the situation was the opposite. According to the results of the fifth modeling experiment on the holdout test set (not presented in the tables), it was determined that all companies, except KATMR, TETMT, and DGNMO, were financially non-distressed in both predicted and actual status. KATMR, TETMT, and DGNMO were found to be financially distressed in both predicted and actual status. The best outputs were selected from all the modeling experiment results on the holdout test. It was observed that the best results and the outputs obtained from the fifth modeling experiment on the holdout test were the same. Due to these similarities, the experiment was discontinued, and the obtained results were summarized in Table 5. All company names are presented in abbreviated form, as stated in KAP (2022b).

Table 5. The best results of the holdout test

	COMPANY	DATE	ACTUAL LABEL	PREDİCTED LABEL
1147	KATMR	2015-12-01	Distressed	Distressed
1205	KONYA	2019-12-01	Non-Distressed	Non-Distressed
1333	LUKSK	2017-12-01	Non-Distressed	Non-Distressed
975	GUBRF	2019-12-01	Non-Distressed	Non-Distressed
1698	SAMAT	2015-12-01	Non-Distressed	Non-Distressed
1787	SILVR	2011-12-01	Non-Distressed	Non-Distressed
140	ARSAN	2015-12-01	Non-Distressed	Non-Distressed
1941	ULKER	2011-12-01	Non-Distressed	Non-Distressed
1900	TTRAK	2019-12-01	Non-Distressed	Non-Distressed

1833	TETMT	2018-12-01	Distressed	Distressed
1077	IZMDC	2011-12-01	Non-Distressed	Non-Distressed
1876	TOASO	2012-12-01	Non-Distressed	Non-Distressed
1454	OLMK	2014-12-01	Non-Distressed	Non-Distressed
1032	IPEKE	2010-12-01	Non-Distressed	Non-Distressed
709	EGGUB	2012-12-01	Non-Distressed	Non-Distressed
587	DGNMO	2016-12-01	Distressed	Distressed
1245	KRATL	2014-12-01	Non-Distressed	Non-Distressed
461	CELHA	2011-12-01	Non-Distressed	Non-Distressed
198	AVOD	2015-12-01	Non-Distressed	Non-Distressed
487	CEMTS	2015-12-01	Non-Distressed	Non-Distressed

4. Conclusion

The dataset used in our study was collected from the financial data of 233 companies listed on the Borsa Istanbul SINAI Index between 2010 and 2020. Each company's financial data during this period was labeled according to its financial distress, with 25 independent financial variables (X1 to X25) defining each company. Both financial ratios and non-ratio-based disclosures from the Public Disclosure Platform (KAP) were utilized to determine whether the companies were in financial distress. The dataset consisted of a total of 2563 company-year observations, with 359 observations indicating financial distress and 2204 observations indicating no financial distress, creating an imbalance in the dataset that required addressing. To address this imbalance, the SMOTE technique was employed. The purpose of balancing the imbalanced data was to distinguish the minority class, not to classify the failed ones as non-failed. In summary, the classifier tends to classify all examples as the majority class, which achieves high accuracy but may miss minority examples. In this study, undersampling and SMOTE methods were used to mitigate this issue. Subsequently, Future Engineering was employed to make the best predictions. The primary goal was to prepare the data for machine learning models, enabling us to achieve better performance results. In this study, data processing was performed using lagging and rolling features. The maximum F1 score (85.1%), recall (84.5%), precision (85.7%), and accuracy (91.6%) were obtained in the 5th machine learning experiment. A total of 417 predictions were made regarding whether companies would experience financial distress, and 92% of these predictions were accurate.

The study demonstrated that XGBOOST, a machine learning method, is highly capable of predicting business financial failures. Therefore, managers, who must always consider the fact that the failure of businesses adversely affects multiple stakeholder groups, can utilize these methods to predict the financial status of their companies. In this context, managers should be aware that financial ratios, particularly financial structure and profitability ratios, which were found to be significant in the study, can be used to predict a company's failures in advance and prevent potential risks.

One of the main limitations of this study can be attributed to the use of a 10-year dataset. The primary reason for this is the availability of Non-Financial Financial Distress indicators from KAP disclosures only from 2010 onwards. The performance of the study is expected to improve with the utilization of a larger dataset. Additionally, future studies could consider incorporating macro variables.

Lastly, the XGBOOST algorithm used in this study is pioneering in measuring financial distress in the national literature, thereby contributing to the literature and setting a precedent for similar studies in the future.

References

Abu-Mostafa, Y. S., & Atiya, A. F. (1996). Introduction to financial forecasting. *Applied Intelligence*, *6*(3), 205–213.

Akkaya, G., Demireli, E. ve Yakut, Ü. H. (2009). İşletmelerde finansal başarısızlık tahminlemesi: Yapay sinir ağları modeli ile İMKB üzerine bir uygulama. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, *10*(2), 187-216.

Aktaş, R. (1993). *Endüstri işletmeleri için mali başarısızlık tahmini: Çok boyutlu model uygulaması*. Ankara: Türkiye İş Bankası Kültür Yayınları.

Aktaş, R., Doğanay, M. ve Yıldız, B. (2003). Finansal başarısızlığın öngörülmesi: İstatistiksel yöntemler ve yapay sinir ağı karşılaştırması. *Ankara Üniversitesi SBF Dergisi*, *58*(4), 1-24.

Albayrak, A. S. ve Yılmaz, Ş. K. (2009). Veri madenciliği: Karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 14*(1), 31-52.

Alkhatib, K., Najadat, H., Hmeidi, I., & Shatnawi, M. K. A. (2013). Stock price prediction using k-nearest neighbor (kNN) algorithm. *International Journal of Business, Humanities and Technology*, *3*(3), 32–44.

Altan, G., & Demirci, S. (2022). Makine öğrenmesi ile nakit akış tablosu üzerinden kredi skorlaması: XGBoost yaklaşımı. *Journal of Economic Policy Researches*, *9*(2), 397–424.

Altman, E. I., & Hotchkiss, E. (2010). *Corporate financial distress and bankruptcy: Predict and avoid bankruptcy, analyze and invest in distressed debt* (Vol. 289). John Wiley & Sons.

Baş, M., & Çakmak, Z. (2012). Gri ilişkisel analiz ve lojistik regresyon analizi ile işletmelerde finansal başarısızlığın belirlenmesi ve bir uygulama. *Anadolu Üniversitesi Sosyal Bilimler Dergisi, 12*(3), 63–82.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Conference: the 22nd ACM SIGKDD International Conference*, 785–794.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., & Chen, K. (2015). Xgboost: Extreme gradient boosting. *R Package Version 0.4-2*, *1*(4), 1–4.

Civan, M., & Dayı, F. (2014). Altman Z skoru ve yapay sinir ağı modeli ile sağlık işletmelerinde finansal başarısızlık. *Akademik Bakış Dergisi*, *41*.

Coelho, L. P., & Richert, W. (2015). *Building machine learning systems with Python*. Packt Publishing Ltd.

Goecks, J., Jalili, V., Heiser, L. M., & Gray, J. W. (2020). How machine learning will transform biomedicine. *Cell*, *181*(1), 92-101.

KAP (2022a). *BİST bildirim*. Erişim Adresi: https://www.kap.org.tr/tr/

KAP (2022b). *BİST şirketler*. Erişim Adresi: https://www.kap.org.tr/tr/bist-sirketler

Keskin, Y. (2002). İşletmelerde finansal başarısızlığın tahmini, çok boyutlu model önerisi ve uygulaması (Yayımlanmamış Doktora Tezi). Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

Kılıç, Y. ve Seyrek, İ. H. (2012). *Finansal başarısızlık tahmininde yapay sinir ağlarının kullanılması: İmalat sektöründe bir uygulama*. *ISAF 2012*. Paper presented at the 1st International Symposium on Accounting and Finance (Tam Metin Bildiri/Sözlü Sunum) (Yayın No: 3438181).

Kılıç, Y., & Seyrek, İ. H. (2012). Finansal başarısızlık tahmininde yapay sinir ağlarının kullanılması: imalat sektöründe bir uygulama. *1st International Symposium on Accounting and Finance* içinde (677–689, ss.).

Li, H. and Sun, J. (2011). Empirical research of hybridizing principal component analysis with multivariate discriminant analysis and logistic regression for business failure prediction. *Expert Systems with Applications, 38(*5), 6244-6253.

Marček, D. (2004). Stock price forecasting: Statistical, classical and fuzzy neural network approach. Torra, V., Narukawa, Y. (Eds). *Modeling Decisions for Artificial Intelligence. MDAI 2004. Lecture Notes in Computer Science*, vol 3131 içinde (41-48, pp). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-27774-3_5

Özdemir, F.S. (2011). *Finansal başarısızlık ve finansal tablolara dayalı tahmin yöntemleri.* Ankara: Siyasal Kitabevi.

Raschka, S. (2015). *Python machine learning*. Packt Publishing Ltd.

Shirata, C. Y. (1998, August). Financial ratios as predictors of bankruptcy in Japan: An empirical research. *In Proceedings of the second Asian Pacific interdisciplinary research in accounting conference*, Vol. 1, 17.

Torun, T. (2007). Finansal başarısızlık tahmininde geleneksel istatistiki yöntemlerle yapay sinir ağlarının karşılaştırılması ve sanayi işletmeleri üzerinde uygulama. (Yayımlanmamış Doktora Tezi). Erciyes Üniversitesi Sosyal Bilimler Enstitüsü, Kayseri.

Ustalı, N. K., Tosun, N., & Tosun, Ö. (2021). Makine öğrenmesi teknikleri ile hisse senedi fiyat tahmini. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi, 16*(1), 1–16.

Yakut, E. ve Elmas, B. (2013). İşletmelerin finansal başarısızlığının veri madenciliği ve diskriminant analizi modelleri ile tahmin edilmesi. *Afyon Kocatepe Üniversitesi İİBF Dergisi, XV*(I), 237-254.