Cilt: 20, Sayı: 4, Yıl: 2020
e-ISSN: 2667-8583

AÜSBD
Anadolu Üniversitesi Sosyal Bilimler Dergisi
Anadolu University Journal of Social Sciences

Comparison of Machine Learning Methods in Prediction of Financial Failure of Businesses in The Manufacturing Industry: Evidence from Borsa İstanbul[1,2,3,4]

Asst. Prof. Dr. Barış AKSOY[5] - Prof. Dr. Derviş BOZTOSUN [6]

Submitted by: 25.10.2019 **Accepted by:** 09.12.2020 **Article Type:** Research Article

Abstract

In this study, Artificial Neural Networks (NN), C5.0 Classification Algorithm, Classification and Regression Trees (CART) analyses were used to predict the financial success/failure of 126 businesses that are operating in the BIST (Borsa İstanbul) Manufacturing Industry Sector. The data contains the years 2006 to 2009. In the study, 25 quantitative variable and 4 qualitative variable were used. The overall classification accuracy from the highest to the lowest of 3 years prior to successful-failure year (for 2006) is 84.21% for CART, 81.58% for ANN and 76.32% for C5.0, respectively. The overall classification accuracy from the highest to the lowest of 2 years prior to successful-failure year (for 2007) is 86.84% for CART, 84.21% for ANN, 78.95% for C5.0, respectively. The overall classification accuracy from the highest to the lowest of 1 year prior to successful-failure year (for 2008) is 92.11% for CART, 92.11 for ANN and 86.84% for C5.0, respectively. ANN and CART models are notable in terms of their ability to predict upcoming financial failure of unsuccessful businesses with 100% classification accuracy from a year ago. The prediction of the financial success/failure by the three models obtained in the study more than one, two and three years ago shows that the models used in this study can be included in the model used by those concerned.

Keywords: *Financial Failure Prediction, Borsa İstanbul, Artificial Neural Networks, C5.0 Decision Rule Algorithm, CART Classification and Regression Trees*

Atıf: Aksoy, B. and Boztosun, D. (2020). Comparison of machine learning methods in prediction of financial failure of businesses in the manufacturing industry: Evidence from Borsa İstanbul. *Anadolu University Journal of Social Sciences, 20*(4), 237-268.

[1] Since the article was submitted for publication in the previous year, ethics committee approval is not required.

[2] This study was prepared in Erciyes University Institute of Social Sciences based on a PhD thesis entitled "Comparison of Data Mining Methods in Predicting Financial Failure in Businesses: An Application in BIST".

[3] PhD thesis titled "Comparison of Data Mining Methods in Predicting Financial Failure in Businesses: An Application in BIST" was supported by Erciyes University Scientific Research Projects Unit with SDK-2016-6988 coded project.

[4] This study is the final form of the paper "Comparison of Data Mining Methods in Financial Failure Prediction of Businesses Operating in the Manufacturing Industry: Borsa Istanbul Example" reviewed according to criticism and contributions which was orally presented at IV. International Symposium on Multidisciplinary Studies.

[5] Cumhuriyet University Faculty of Economics and Administrative Sciences Department of Banking and Finance, **baksoy@cumhuriyet.edu.tr**, ORCID: 0000-0002-1090-5693,

[6] Kayseri University, Faculty of Applied Sciences, Department of Accounting and Finance Management, **dboztosun@kayseri.edu.tr**, ORCID: 0000-0002-2656-2701

ANADOLU ÜNİVERSİTESİ
Cilt: 20, Sayı: 4, Yıl: 2020
e-ISSN: 2667-8583
AÜSBD
Anadolu Üniversitesi Sosyal Bilimler Dergisi
Anadolu University Journal of Social Sciences

İmalat Sanayi Sektöründeki İşletmelerin Finansal Başarısızlık Tahmininde Makine Öğrenmesi Yöntemlerinin Karşılaştırılması: Borsa İstanbul Örneği

Dr. Öğr. Üyesi Barış AKSOY - Prof. Derviş BOZTOSUN

Başvuru Tarihi: 25.10.2019 **Kabul Tarihi:** 09.12.2020 **Makale Türü:** Araştırma Makalesi

Öz

Bu çalışmada 2006-2009 yılları arasında BIST (Borsa İstanbul) İmalat Sanayi Sektöründe faaliyet gösteren 126 işletmenin finansal başarı/başarısızlığını tahmin etmek üzere Yapay Sinir Ağları (NN), C5.0 Karar Kuralı Türetme Algoritması, Sınıflama ve Regresyon Ağaçları (CART) analizi yapılmıştır. Çalışmada 25 mali tablolara dayalı nicel ve 4 mali tablolara dayalı olmayan nitel değişken kullanılmıştır. Analizler sonucunda başarılı-başarısızlık yılından 3 yıl öncesinde (2006 yılı için) en yüksekten düşüğe genel olarak (başarılı ve başarısız toplamı) sırasıyla; CART %84.21, Yapay Sinir Ağları %81.58 ve C5.0 %76.32 sınıflandırma doğruluğuna sahiptir. Başarılı-başarısızlık yılından 2 yıl öncesinde (2007 yılı için) en yüksekten düşüğe genel olarak sırasıyla; CART %86.84, Yapay Sinir Ağları %84.21 ve C5.0 %78.95 sınıflandırma doğruluğuna sahiptir. Başarılı-başarısızlık yılından 1 yıl öncesinde (2008 yılı için) en yüksekten düşüğe genel sınıflandırma doğruluğu sırasıyla; CART %92.11, Yapay Sinir Ağları %92.11 ve C5.0 %86.84 sınıflandırma doğruluğuna sahiptir. Çalışmada elde edilen üç modelin finansal başarı/başarısızlığı bir, iki ve üç yıl öncesinden yüksek oranda tahmin etmesi ilgililerin kullandıkları modele bu çalışmada elde edilen modellerin de dâhil edilebileceğini göstermektedir.

Anahtar Kelimeler: *Finansal Başarısızlık Tahmini, Borsa İstanbul, Yapay Sinir Ağları, C5.0 Karar Kuralı Türetme Algoritması, CART Sınıflama ve Regresyon Ağaçları*

Introduction

Today, competitive free market only gives the chance of survival to successful businesses, while businesses failing to comply with new market conditions are facing with financial failure and bankruptcy (Chen, 2011b, p. 11262). When a business pulls out of the market due to financial failure, it can affect the managers, the employees, the sector, investors, lenders and the economy of the country (Baixauli and Mo´dica-Milo, 2010, p. 76). For this reason, it is of great importance for all related groups to have a forecasting model that can notice the financial statements of the companies suffering financial distress on time (Gordini, 2014, p. 6433).

Business failure arises in two ways, economic and financial failure. The term business failure refers to the economic failure of businesses by excess of expenses over their revenues (Li & Sun, 2013, p. 186). Financial failure occurs in the form of technical bankruptcy and bankruptcy. The technical bankruptcy is the condition of a business that is unable to make payment on a liability but its business assets are generally sufficient to pay the liabilities. The bankruptcy represents the situation where the business debts can not be covered by the business assets (Sayılgan & Ece, 2016, s. 50).

One of the biggest problems in business financial failures is the lack of a consensus on defining the company's failure or financial failure (Muller, Steyn-Bruwer & Hamman, 2009, p. 22). Many people confuse the financial failure with concepts that are relevant but different in meaning, such as default, bankruptcy, and liquidation. Financial failure does not always result in bankruptcy (Bilir, 2015, s. 9). An enterprise in financial failure can experience a dynamic process of change between the two ends of the financial strains. A business can be said to fail when it faces any of the situations such as default, bank account lending, bankruptcy of the business, explicit declaration that debts can not be paid, agreement with creditors to reduce debts (Sun, Li, Huang & He, 2014, p. 43).

Internal reasons that may cause financial failure in the business include insufficient operating capital, excessive increase in short-term debt, increase in cost of capital, failure to control budgets, delay in payments and rejection of the loan (Jardin, 2009, p. 40). External factors of the financial failure in the business include growth rate of the economy, economic crisis or stagnation, inflation, high interest rate, exchange rate, tight monetary policies, changes in preferences, attitudes and consumer behaviors (Yakut & Elmas, 2013, s. 238).

In regards to the measures that can be taken against financial failure, businesses can use financial restructuring methods such as focusing on financing with equity and turning to external resources by offering stocks instead of internal resources for financing with equity (Tuncay, 2011, ss. 115-116). An organization that can not pay its debts may try to make an informal agreement with its creditors, known as a “special payment plan” (Brealey & Myers, 2007, pp. 590-591). By persuading creditors from giving up on certain parts of the credit balance, the business can reduce its financial failure. In addition to providing resources by selling tangible fixed assets, the business may be able to maintain its activities as it used to be by renting tangible fixed assets that it sold (Akgüç, 1997, ss. 950-951). It is also possible through mergers to expand the debt capacity and to perform better in comparison to the situation before the merger and acquisition of the company that sustains its existence (Aksoy & Ertaş, 2016, s. 773).

Since no valid model can be found in all circumstances considering the studies performed to date, studies in the field of financial failure prediction still remain popular in the financial literature. Among the financial failure prediction methods, logistic regression, discriminant analysis and artificial neural networks have been used extensively. The frequently used methods among the machine learning, CART and C5.0 algorithm, that have high classification success is emphasized to be used in future researches. The objectives of this study are as follows.

- ✓ To determine the impact of the sample selected based on financial statements and not based on financial statements of success/failure criteria on financial failure prediction performance.
- ✓ To create a comprehensive set of variables with superior prediction ability by using the financial ratios obtained from financial statements and qualitative variables obtained from company news and announcements.
- ✓ To develop models with high prediction accuracy, working in harmony with the variables determined in the study.
- ✓ To determine the methods and models with the highest prediction power of NN, CART, C5.0 algorithm methods used in the study and compare the obtained models within the framework of their distinctive features and limitations.

The study consists of introduction in the first part, the literature review in the second part, the research methods in the third part, the data mining methods (CART, C5.0 and ANN) analyzes and findings in the fourth part, comparison and evaluation of the results in the fifth part and conclusion in the sixth part.

Literature Review

Numerous studies based on statistical and machine learning methods have been done in the literature to predict business failure. The results emphasizes different or similar findings. It has seen that the similarity and difference for tehese results are due to the difference in the selection of the data set used in the prediction, the preprocessing done on the data, and the algorithm parameters. All methods claim a certain degree of accuracy in terms of prediction. However, none of these approaches predominates in terms of accuracy and reliability. There is no general consensus that a method is consistently superior to the others (Yip, 2006, p. 492).

Chen, Marshall, Zhang & Ganesh (2006) used Multiple Discriminant Analysis (MDA), Logistic Regression Analysis (LA), Decision Tree (DT), Artificial Neural Networks (NN) methods to predict financial failure in China. The list of ST (special treatment) companies available from the China Stock Star Database covers the period December 1999-June 2003. The sample of ST and non-ST companies is divided into training and test sub-samples. In prediction, 70% of the sample set was selected for training and 30% for testing. The training set was used to create a preliminary model, and the validation set was used to test and adjust the model weights during prediction. K-fold cross-validation was applied in the analyses. In the study, a single-hidden layer model with back-propagation learning algorithm was selected. The prediction accuracy of the methods used in the study ranged from 78% to 93%. As a result of the analysis, it was stated that LA and NN models were the best prediction models and provided the lowest total misclassification cost.

Chandra, Ravi & Bose (2009) used NN, SVM, CART techniques to predict the failure of dot-com companies. The data was obtained from Wharton Research Data Services (WRDS), that consist of 120 unsuccessful and 120 successful dot-com companies. Based on the financial statements of 240 companies included in the sample, 24 financial ratios were determined. A 10-fold cross-validation technique was used for validation of the data set for all methods. The results were supported by the ROC (Receiver Operating Characteristic) curve. According to the results of the analysis, the performance of the methods for classification accuracy were CART> SVM> NN, respectively.

In the study of Doğrul (2009), LA, Classification and Regression Trees (CART) and ANN models were created by using 29 financial ratios on 70 successful and 70 unsuccessful manufacturing businesses selected according to certain criteria that were traded in BIST between 1997 and 2007. With the established models, it is aimed to estimate the financial status of the enterprises 1, 2 and 3 years in advance. As a result of the analysis, LA and CART predicted financial failure at 92.90%, NN at 90.00% classification accuracy one year ago, LA at 88.60%, CART at 94.30%, NN at 85.00% classification accuracy two years ago and LA at 75.70%, CART at 90.00% and NN at 82.80% classification accuracy three years ago.

Gepp, Kumar & Bhattacharya (2010) compared the classification performances of MDA, NN, C5.0 and CART models in their financial failure prediction. The classification and prediction ability of the C5.0 algorithm was clearly found to be the best classification technique. The C5.0 algorithm produced more complex trees compared to CART. The CART model had the most consistent predictability on misclassification costs. DT methods performed better than NN and LA in predicting with six financial ratios and continuous data. It was stated that all of the different DT techniques performed better than MDA in order to achieve the classification between successful and unsuccessful businesses.

Chen (2011b) analyzed with a total of 37 variables consisting of financial and non-financial variables, by taking 100 company data of 50 financially unsuccessful and 50 financially successful companies registered in Taiwan Stock Exchange between 2000-2007. C5.0, CART and CHAID and LA methods were used in the study. As the time of financial distress approached, the decision tree prediction model gave more accurate results. C5.0 algorithm's prediction accuracy was 88.80% for 8 pre-periods, while it was 97.01% for 2 pre-periods. For LA, the prediction accuracy before the 2- and 8-period prior to financial distress was 85.07% and 91.70%, respectively. It was concluded that the accuracy rate of the C5.0 algorithm was better than CART and CHAID.

Chen (2011a), the data of about 200 businesses registered in the Taiwan Stock Exchange were obtained and a total of 42 (including 33 quantities, 8 qualities and 1 combined macroeconomic index) rate was used. It is concluded that traditional statistical methods can better process small data sets without compromising the performance of predictions, whereas intelligent techniques perform better in large data sets. The analysis showed that C5.0 and CART had the best prediction performance for upcoming bankruptcies.
Aktan (2011) examined the effectiveness of 8 machine learning algorithms in the classification of financial distress, such as Naive Bayes, Bayesian Network, k-NN, NN, SVM, C4.5, CHAID and CART. In the study, it was aimed to present the effectiveness of machine learning algorithms in the field of financial distress prediction. 10-fold cross-validation method was preferred to avoid the problem of memorizing all the data used in training and validation processes. In the one-year period before failure, all algorithms except C4.5 achieved more than 90% or equal classification accuracy. CART algorithm was found to have the highest classification accuracy with 96.60%, and the C4.5 had the lowest classification accuracy with 87.2%. In the period before the failure, Naive Bayes showed the lowest performance in classification accuracy with 80.2%, while CART reached 99.5% accuracy in the period before failure.

In the study of Kılıç (2011), 6 years data of 137 companies operating in the manufacturing sector, which were traded in BIST, were obtained between 2005 and 2010 and analyzed with C5.0 algorithm and ANN. Financial ratios of each company were calculated using the 12-month balance sheet and income statement, and analyses were carried out with the C5.0 algorithm and NN. As a result of the analyses, the C5.0 technique achieved higher prediction success compared to NN in the training group, while NN in the test group produced more successful results. It was concluded that the performance of NN technique in predicting financial failure was higher than C5.0 technique.

Jardin (2012) found that MDA, LA and ANN analysis were applied to estimate financial failure. The 2002 data set is divided into two sub-samples as a learning sample (A) consisting of 450 companies and a verification-test sample (T) consisting of 50 companies. According to the results of the analysis ANN performs better than MDA and slightly less than LA. The best result in the classification was 88.92% with ANN in the test samples, followed by LA with 86.02% and MDA with 83.86%.

Yakut (2012) created financial failure prediction models by using the 2002-2010 data of 60 successful and 60 unsuccessful businesses trading in BIST. From data mining techniques, C5.0 algorithm, SVM and NN were compared with each other to determine the best prediction method. NN method gave better results in general compared to C5.0 and SVM methods. When classified in terms of classification accuracy, NN>C5.0>SVM was found. According to the three methods, prediction results of one year before the failure achieved higher prediction than the results of 2, 3 and 4 years before, respectively.

Delen, Kuzey & Uyar (2013) calculated 31 financial ratios covering the years 2005 and 2011. In this study, four popular decision tree algorithms (CHAID, C5.0, QUEST and CART) were used to investigate the impact of financial ratios on company performance. After the prediction models were developed, sensitivity analyzes were carried out to measure the relative importance of the independent variables. As a result of the analysis, CHAID and C5.0 decision tree algorithms were found to be the models with the best prediction accuracy.

Tsai, Hsu & Yen (2014) conducted a comprehensive study to compare three commonly used classification techniques: multilayer perceptron (MLP) NN, Support Vector Machines (SVM) and Decision Trees (DT). Experimental results with three public data sets indicated that DT communities of 80-100 classifiers using the enhancement method showed the best performance. The Wilcoxon signed rank test also concluded that DT communities performed significantly differently than other classifier communities.

Okay (2015) used 32 successful and 32 unsuccessful companies to investigate the failures of non-financial companies registered in BIST between 2000-2015 and to compare the accuracy of different prediction models. They compared the accuracy of different prediction models such as MDA, Quadratic Discriminant (QDA), LA, Probit Analysis, DT, NN and SVM. When validation samples are used, NN model was found to have the best prediction power among all the models used in this study. With all data, total accuracy rates were determined as DT 89.06%>MDA 79.69% = PROBIT 79.69%>LA 76.56%> QDA76.55, respectively. The test samples total accuracy rates were NN 81.30%> QDA 79.40%> SVM 78.80%> MDA 77.50%> LA 76.90% = PROBIT 76.90%> DT 68.10%. It was concluded that the tree model had the lowest accuracy rate with 68.00% in the test sample.

In the study of Gepp and Kumar (2015), Cox Survival Analysis Model, CART, LA and MDA methods were used for financial failure prediction. According to the data used in the study, CART and Cox analysis was found to be superior classifier than LA. As a result, DT, in particular the CART model, has better classification accuracy than other techniques. DTs can be stated to be the best compared to other methods to make accurate predictions without the risk of violating statistical assumptions.

Misund (2017) carried out LA and CART analysis to estimate the financial failure of the 1626 enterprises operating in the Norwegian salmon industry with two separate data from 1994-1999 data set 1 and 2000-2002 data set 2. The performance of the LA and CART model was compared with the benchmark comparison model developed by the National Bank of Norway. The data that comprise of the years 1994 to 1999 was carried out by using the LA and CART models, which consisted of salmon farming and fishing companies. The performance of the models was compared with a benchmark model developed by the National Bank of Norway for all industry sectors. It is concluded that LA and CART models are better in the salmon industry compared to the National Bank of Norway standard comparison model among high-risk and low-risk firms. According to the classification results for data set 1, CART (86.70%)> LA (71.80%)> Benchmark (57.70%) and for data set 2, LA (79.20%)> CART (75.20%) > Benchmark (53.50%).

Le & Viviani (2018) attempted to predict the financial failure of 3000 US banks, including 1438 failures and 1562 successful studies. Two conventional statistical methods, such as Discriminant analysis and logistic theorem, and three machine learning methods, such as Artificial neural network, Support Vector Machines and k-nearest neighbors, were used as methods. For each bank, the data were collected over a period of 5 years. The 31 financial ratios have been used obtained from the Bank's financial reports. The data were mixed to prevent the algorithm from memorizing the data. The data set is divided into 30% test samples for 70% training and data testing to learn the data. They observed that machine learning, ANNs and k-NN methods perform more effectively than traditional methods. ANN and the nearest neighboring algorithm proved to be extremely successful in accurately detecting financial failure, but in other methods they stated that this failure was low. The empirical result suggests that the neural network and the nearest neighboring methods are the most accurate. It was also found that SVM did not perform better than traditional statistical methods.

However, some authors (eg Muller, 2009, p. 29) stated that if the project was predicted to be unsuccessful as a result of the prediction made just one year before the year of failure, it would be higher for the cost of the decision-maker than foreseen 2 and 3 years before the failure. The earlier the financial failure can be determined with high accuracy, the earlier necessary measures will be taken and the lower the cost to the decision-maker. In the light of the given literatüre findings, the contribution of this study can be expressed as;

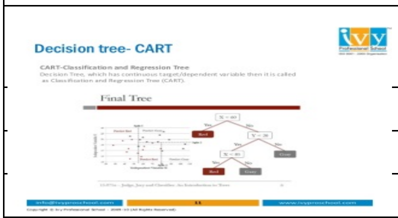
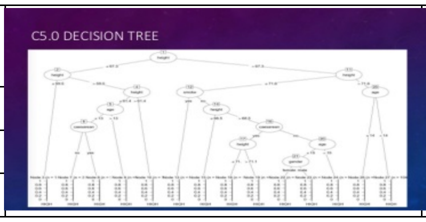
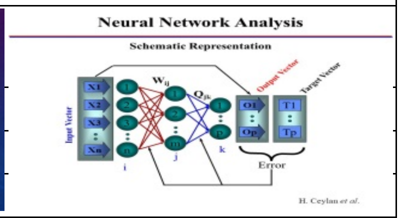
- ✓ To predict of financial failure in the study was carried out 1,2 and 3 years ago, which is a reasonable period for all interested parties to take necessary measures.
- ✓ In order to determine the financial success-failure, financial failure-indicators obtained from the financial statements as well as financial failure indicators obtained from BIST company news and announcements, special case disclosures were used.
- ✓ 25 financial ratios based on financial statements and 4 qualitative variables not based on financial statements were used.
- ✓ 10-fold cross validation method was used as a validation method in order to prevent the memorization of samples, to ensure better data distribution and to obtain more reliable results.
- ✓ In the analysis, parameter optimization has been used in order to test all of the minimum and maximum value ranges in the important parameter values and to obtain the highest performance model.

Research Method

Although the models related to financial failure prediction in businesses are significantly different from each other depending on the modeling method, variables or sample used, the correct classification rates of the models decrease significantly if the prediction success exceeds one year (Jardin & Séverin, 2011, p. 701). While it is sufficient to predict 1 year in advance for some decision-makers, longer-term business decisions or investment decisions can take a longer period of prediction (Gepp & Kumar, 2015, p. 398). Geng, Bose and Chen (2015) found that the prediction period up to 3 years yielded the best predictive results, and that the prediction time longer than 3 years had a negative impact on the correct classification success. Therefore, this study predicts one, two and three years before financial failure.

Since the number of methods used is being high and financial failure is predicted from 1,2 and 3 years before, not a single year, the study text is devoted to literature review, method, input of parameters, analysis phase, findings and interpretation of findings. In the study, SPSS 21 program was used for discriminant forward-step analysis. There are many commercial and open source programs to implement Data Mining applications. RapidMiner (YALE), WEKA and R programs are among the most widely used open source data mining programs among these programs. In 2007, Data Mining experts visited the site according to a survey, according to the WEKA, RapidMiner is more popular among experts (Dener & Orman, 2009, ss.1-2). For this reason, RAPIDMINER 7.6 program was used for artificial neural networks, CART and C5.0 analyzes. Table 1 shows the Methodology of the Study.

Table 1.
Methodology of the study

Obtaining Data of Companies for Business Failure Prediction		
Determination of Successful-Unsuccessful Businesses with Given Failure Criteria		
Recording Quantitative (Financial Ratio) and Qualitative Variables for Successful and Unsuccessful Businesses		
CART	C5.0	ANN
29 Variable Data Set	8 Variable Data Set Selected by Discriminant Stepwise Analysis	8 Variable Data Set Selected by Discriminant Stepwise Analysis
All data set was divided into 70% training and 30% test data and 10-fold cross validation method was used.	All data set was divided into 70% training and 30% test data and 10-fold cross validation method was used.	All data set was divided into 70% training and 30% test data.
Program used in analysis: Rapidminer 7.6		
		
Discussion of Analysis Results and Recommendations		

Data set and limitations

Because financial failure is defined differently by researchers, the reason for failure used in the definition is different. Firstly, it is necessary to make sure what is understood from failure. Once the criteria for business failure are determined, it is necessary to identify successful and unsuccessful firms in the mainstream. The most common sampling method is to identify successful samples of the same number and the same industry after sampling of unsuccessful firms (Gallego & Quintana, 2012, pp.159-160). The final criterion to make the final selection of financial ratios is that as the selected ratios are calculated for the four-year working period, the four-year financial ratio information of businesses in the sample is required.

The study periods whose data will be obtained are between 2005-2015. It is necesary to obtain as if of a business is successful in 2005 and 2006 but failed in 2007, the data of the 3 years before the failure are needed. For this reason, 2008 data were taken into consideration as the beginning of the year of failure. Considering the studies of Bee and Abdollahi (2013), retailers and finance companies were not included in the sample since they were highly dependent on economic conditions. Considering the studies of Glezakos, Mylonakis & Oikonomou (2010), investment companies, leasing companies, banks and insurance companies were not included in the sample because the financial ratios of the businesses in some sectors had certain characteristics. The most important limitation faced in almost all of the other studies in the literature is the year of success and failure year of the companies. In order to ensure that the successful and unsuccessful business sample was as high as possible, data between 2005 and 2015 were used and the year with the most failures was determined as 2009. As 2009 was determined as the year of failure due to the high number of samples, 63 businesses whose data for 3 years ago, 2006, 2007, 2008, could be accessed were included in the study. In other studies in the literature, for example, in the study of Yakut (2012), whichever year had most failure was accepted as the base year for successful businesses. Since 2009 was accepted as the year of failure, the same year was determined as the year of success for successful businesses, and 63 of the businesses that were successful in the period 2006-2009 were chosen randomly.

The data of 126 businesses determined to be used in the study were taken from BIST's website for 2009 and before, and from KAP’s news/announcements and market change announcements for 2010 and after. The list of companies whose trading order was closed, and the lists of companies whose stock market was changed and excluded from the exchange list due to financial status were obtained from the website of BIST. In other studies in the literature, for example, in the study of Yakut (2012), whichever year had most failure was accepted as the base year for successful businesses. Since the transactions of the businesses whose BIST trading order was closed at different periods of the year, the year with the balance sheet and income statement dated 31.12 a year before the transaction closure was determined as the failure year in order to eliminate the drawbacks of evaluating the financial statements for a different period. The list of businesses whose trading order was closed was obtained from the website of BIST and the data of the businesses whose financial statements were available for at least 4 years between 2005 and 2015 were evaluated. The fact that the data set contained a 10-year interval was due to the effort to increase the number of businesses that failed financially. In our country, it is not possible to find data about the companies that have failed financially to form the sample of such a study (Aktaş, Doğanay & Yıldız, 2004, ss. 11-12).

Glezakos et al. (2010) following the studies and the financial ratios of some companies in the sector have certain characteristics, therefore, investment companies, leasing companies, banks and insurance companies are not included in the sample. Researches on financial distress in enterprises are usually limited to the use of financial indicators disclosed in the accounting tables, and this neglects the role of non-financial indicators. In our study, company news and announcements for four years were examined in order to use the indicators that are not based on financial statements. Except for the financial statement failure criteria, the request of the bankruptcy of the bankruptcy, the demand for the restructuring of the commercial and financial debts, and the material disclosures not related with the financial statements were examined one by one. The criterion of not distributing dividends in material event disclosures is taken into account only in the case of not having a profit or a loss in the previous year's losses and has been used as a criterion of failure. They were evaluated as unsuccessful because they did not distribute dividends to privileged shares, not ordinary owners of ordinary shares.

The most important limitation faced by almost all other studies in the literature is the year of success and the year of failure. In Table 2, due to the high number of unsuccessful enterprises, 2009 was identified as the year of failure, and 63 enterprises that were able to reach the data of 2006, 2007, and 2008 from 3 years ago were taken. In other studies in the literature, such as Yakut (2012), the year of failure was the base year for successful enterprises. For this reason, 2009 has been determined as the success year for the successful enterprises, and 63 enterprises from the companies that are successful in the 2006-2009 period have been chosen by chance.

Table 2.
Distribution of failed companies by years

Years	Number of failed companies
2008	40
2009	63
2010	42
2011	44
2012	57
2013	55
2014	53
2015	50

Source: The study was created by the author according to the criteria determined in the study.

Before the financial distress in the businesses, there are many confidential information that are not only quantitative financial indicators but also candidates for being a non-quantifiable qualitative variable that is the source of financial distress (Sun et al., 2014, p. 53). In this study, successful and unsuccessful businesses were identified using financial statement based failure indicators as well as the financial failure indicators not based on the financial statements indicated by the material disclosure.

If the criteria are not clearly and clearly defined in the classification of the enterprises as successful or unsuccessful, the success rates obtained in the estimation may not be at the desired level. The first step in preventing financial failure is to determine the causes of financial failure. Businesses suffer from many different reasons. This situation differs from country to country and from sector to sector (Kılıç & Seyrek, 2016, s. 12). All cases between enterprises experiencing difficulty in paying their debts due to maturities in developing countries are referred to as financial failure (Selimoğlu & Orhan, s. 25). For this reason, as shown in Table 2, financial failure criteria based on financial statements between financial distress and bankruptcy and not based on financial statements were used. Table 3 shows the indicators of financial failure based on the financial statements and not based on the financial statements indicated by the material disclosure.

Table 3.
Indicators of Financial Failure Based on the Financial Statements and not based on the Financial Statements Indicated by the Material Disclosure

Indicators of Financial Failure Based on the Financial Statements	Negative equity
	Decrease in two third of the equity
	Decrease in total assets by 10% or more
	Company's making a loss over the last two or more years
Indicators of Financial Failure Not Based on the Financial Statements With Material Disclosure	Permanently closing the sequence of actions in BIST
	Restructuring of debts with financial institutions, creditor companies and asset management company
	Attachment and injunction
	Bond default and restructuring
	Not to distribute profit share to privileged share certificates due to period loss
	Refusal of the transfer to watchlist company market or withdraw from the watchlist company market
	Avoidance of auditor's opinion in independent auditing of financial statements
	Capital decrease against previous year losses
	Collective dismissal of employees
	Public seizure of private property
	Meeting with the lender banks
	Filing a bankruptcy lawsuit or rejection of suspension of bankruptcy
	Sale of tangible fixed assets (Machinery, buildings and land) or sale/renting of tangible fixed assets to financial leasing company
	Suspension of activities

Source: Created by the author in the direction of the literature review and recommendations given by members of the thesis surveillance committee.

There is no theory to date, which precisely states what financial ratios should be in the prediction of financial failure. Models depend on data set, data availability, data quality and method of analysis. The rates used in predicting financial failure may vary in different studies (Iwan, 2005, p. 42). Financial ratios are the most preferred indicators in determining financial failure. Jardin (2009) found that in 190 studies he reviewed, 93% of them used financial ratios and the remaining 7% used other variables (Jardin, 2009, p. 41). By using financial ratios, comparisons can be made among companies and sectors within an industry or in a company. Such a tool can also be used to compare the relative performance of companies of different sizes (Delen et al., 2013, s. 3970).

Since the relationship between financial ratio and business status is dynamic and the rates differ in each stage of countries, industries and economic cycles, it is not claimed that the financial rate set obtained in this study can be generalized, as it is in the study of Lussier (1995). The differences in accounting standards across countries limits the generalizability of the results to companies in different countries. Investors, lenders and suppliers can combine this model with their own techniques, and the model can be incorporated into other models and methods in the literature.

When determining the financial ratios that constitute the independent variables of the study, the ratios in the finance literature were taken into consideration. Appendix 4 provides information on the researchers using the financial ratios used in our study in their studies. The 25 financial ratio variables and the 4 non-financial variables are shown on Table 4. In CART analysis, all of the variables given in table 3 are used. In C5.0 and ANN analyzes, variables X4, X6, X13, X20, X21, X22, X24, X27, which were determined by discriminant stepwise analysis, were used.

Table 4.
Quantitative variables obtained from financial statements and independent variables not based on financial statements obtained from public disclosure platform

		Financial Ratios	Calculation
Liquidity Ratios	X1	Current Ratio	Current assets / Short-term liabilities
	X2	Acid-Test (Liquidity) Ratio	Current assets – Inventories / Short-term liabilities
	X3	Cash Ratio	Cash and Cash Equivalents / Short-term Liabilities
	X4	Inventory to Total Assets Ratio	Inventory / Total Assets
Leverage Ratios	X5	Long Term Debt to Total Assets Ratio	Long-Term Debt / Total Assets
	X6	Short Term Debt to Total Assets Ratio	Short-Term Debt / Total Assets
	X7	Liabilities to Assets Ratio	Total Liabilities / Total Assets
	X8	Short Term Debt to Equity Ratio	Short Term Debt / Stockholders' Equity
	X9	Fixed Assets to Equity Ratio	Fixed Assets / Stockholders' Equity
	X10	Current Assets to Total Assets Ratio	Current Assets / Total Assets
	X11	Debt to Equity Ratio	Total Liabilities / Total Stockholders' Equity
Efficiency Ratios	X12	Inventory Turnover Ratio	Cost of Goods Sold / Average Inventory
	X13	Receivables Turnover Ratio	Net Credit Sales / Average Accounts Receivable
	X14	Asset Turnover Ratio	Net Sales / Total Assets
	X15	Equity Turnover Ratio	Net Sales / Stockholders' Equity
	X16	Fixed Asset Turnover Ratio	Net Sales / Average Fixed Assets
	X17	Current Assets Turnover Ratio	Net Sales / Current Assets
	X18	Tangible Fixed Assets Turnover Ratio	Net Sales / Tangible Assets
Profitability Ratios	X19	Gross Margin Ratio	Gross Profit / Net Sales
	X20	Operating Margin Ratio	EBIT / Net Sales
	X21	Net Profit Margin	Net Profit / Net Sales
	X22	Return On Total Assets (ROTA)	EBIT / Total Assets
	X23	Return On Assets (ROA)	Net Income / Total Assets
	X24	Return On Equity (ROE)	Net Income / Stockholder's Equity
	X25	Expense Ratio	Total Fund Costs / Total Fund Assets
Variables Not Based on Financial Statements	X26	Company Operation Time	
	X27	Being Audited by Four Major Audit Firms (Pricewaterhousecoopers-Deloitte Touche Tohmatsu- Kpmg- Ernst And Young)	
	X28	Free Float Rate %	
	X29	Real and Legal Persons with a Direct or More Share of 5% in Capital - Foreign Capital Share in Non-Public Shares %	

Source: Created by the author as a result of the literature review.

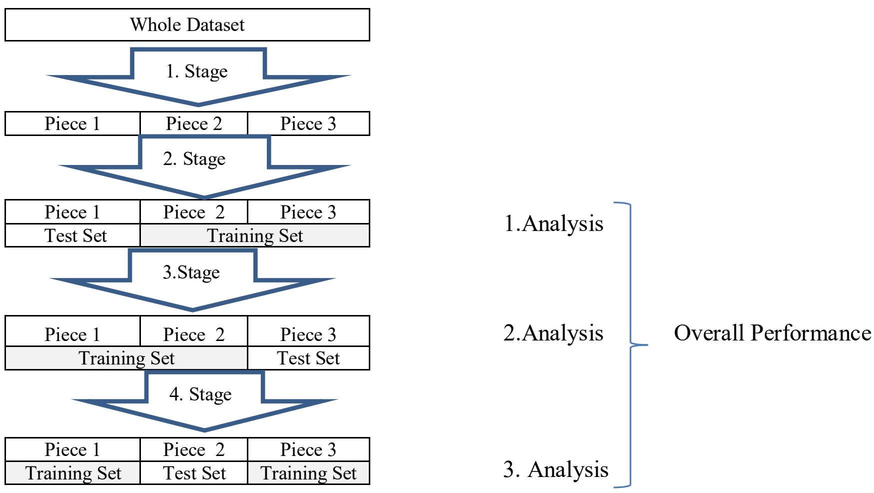
Machine Learning Methods (CART, C5.0 And ANN) Analysis and Findings

Data mining is the process of transforming confidential knowledge in the data into qualified knowledge by using statistical analysis techniques and artificial intelligence algorithms together (Uzar, 2013, s. 52). Against existing financial analysis methods, data mining provides some advantages, which are ability of obtaining important information from huge data and competence of obtaining previously unknown information. There exist two major constraints of data mining implementation that are lack of experts on both data mining and related domains and cost of computer software and hardware used (Özkan & Boran, 2014, p. 59). Machine learning can be defined as the ability of the computer to learn about an event and decide on similar events in the future (Öztemel, 2012, s. 17).

CART Analysis and Findings

One of the decision tree algorithms designed by Breiman et al., Classification and Regression Trees (CART), is used as a classification tool to predict continuous dependent variables using a set of independent predictors (Chuang, 2013, p. 176). Using the gini index as a branching criterion, the CART tree grows continually splitting without any stopping rule. When there is no new split, the process of pruning from the tip to the root is initiated. The most successful decision tree is determined by evaluating with a randomly selected test data after each pruning (Çalış, Kayapınar & Çetinyokuş, 2011, s. 6). In the stratified sampling, samples are taken from both classes (successful-unsuccessful) equally and randomly. It is successful when the dependent variable is categorical (Liang, Tsai & Wu, 2015, p. 291). Therefore, stratified sampling was used in sample selection of this study.

The K-fold cross-validation method is an effective method frequently used in the literature for appropriate data distribution. In k-fold cross-validation, the data is randomly divided into k equal parts. Analysis is done using a piece for testing and the rest for training, respectively. Then another piece is used for testing, the others are used for training. Data mining analysis is performed at each stage and after all parts are tested, overall performance is obtained. In experimental studies, the most appropriate value for k number was found to be 10 according to expert opinions (Çelik, Akçetin & Gök, 2017, p. 243). Gaganis (2009) stated that 10-fold cross-validation as a model validation type was one of the best methods to increase the detection accuracy, and over 75% detection accuracy was a good outcome in the social sciences. Figure 1 shows the k-fold cross-validation.



Source: Çelik et al.(2017) , Data Mining with Rapidminer, s. 244.

Figure 1. K-fold cross verification

In the literature, where different rates are used in the separation of the training and test data set. For example, Geng et al. (2015) reported standard deviation to increase when given 90% of the training rate and 10% of the test rate. In order to avoid this problem, 70% of all the data were divided into two as training data and 30% as test data and 10-fold cross validation was performed to avoid this problem. Table 5 gives the CART analysis parameters entered in the program.

Table 5
CART Classification and regression tree analysis parameters (2006, 2007, 2008 years)

Specified Parameters for Analysis	**Explanation**			
Data Set Separation	70% Training, 30% Test Data Set			
Validation Type	10 Fold Cross-Validation			
Number of Variables	29			
Sample Selection	Stratified Sample Selection			
Split Criterion	Gini Index			
Specified Parameters for Analysis	**Minimum**	**Maximum**	**Steps**	**Scale**
Minimal Size For Split	1.0	4.0	10	Linear
Minimal Leaf Size	1.0	2.0	10	Linear
Minimal Gain	1	20	10	Linear
Maximal Depth	1	20	10	Linear
Confidence Level	0	0.25		
Number of Pre-Pruning	0	10	-	-

The Kappa test is a statistical method that measures the reliability of aggrement between two or more observers. In the classification by Fleiss, the Kappa value of 0.75 and above is considered to be perfect aggrement, between 0.40-0.75 moderate-good aggrement, and below 0.40 as perfect disaggrement. As the kappa value in the study is 0.684 three years before failure, 0.737 two years before failure and 0.842 one year before failure, it can be said that there is a low differencebetween the predicted group and the actual / observed group performance, the financial failure is good for 3 years before, for two years ago there is moderate-good aggrement and the perfect aggrement for one year ahead. Table 6 gives parameters that give the highest classification accuracy determined by CART decision tree analysis performance measurement results and parameter optimization.

Table 6.
Parameters giving the highest classification accuracy determined by CART decision tree analysis performance mesurement results and parameter optimization (2006, 2007, 2008 Years)

Parameters	**2006**	**2007**	**2008**
Accuracy	84.21%	86.84%	92.11%
Classification Error	15.79%	13.16%	7.89%
Kappa	0.684	0.737	0.842
Weighted Mean recall	84.21%	86.84%	92.11%
Weighted Mean Precision	85.80%	86.94%	93.18%
Minimal Size For Split	4	2	3
Minimal Leaf Size	2	2	2
Minimal Gain	12,4	4,8	4,8
Maximal Depth	9	7	12

In Figure 2, the CART analysis was given the decision tree image 3 years before (2006). One of the aims of our study was to determine important variables in the creation of CART and C5.0 decision trees. For this reason, in order to show the variables that specify the root of the tree and the nodes in the formation of the tree, the figures were detailed for the years 2006 and only the shape was given for the other years. At the root of the CART decision tree, the independent variable "X23, Active Profitability Rate" which separates the classes from the first in the successful / unsuccessful business divide, is the root of the decision tree. "Active profitability ratio", Doğrul (2009) in the CART analysis in the study made one and two years before the power of the most variable as the predictive power of the tree formed the root. "The asset profitability ratio" is used in the analysis performed by using the C5.0 algorithm used in Yakut (2012) study. "X23 Active profitability rate" was found to be unsuccessful in 100% of 22 enterprises with an active profitability ratio of less than or equal to 0.003, and "X18, Tangible Asset Transfer Speed" was found as the second most important variable in decision making for enterprises larger than 0.003. 34 of the 43 enterprises with "X18, tangible fixed asset turnover" 43 less than or equal to 3.562 were successful and 9 were unsuccessful. "X7, Financial Leverage Ratio" was found as the third important variable in the decision of the tree for enterprises larger than 3.562. X7, financial leverage ratio was found to be 100% successful in 2 enterprises with a value less than or equal to 0.158. "X7, financial leverage ratio" of the 21 enterprises with greater than 0.158, 13 were unsuccessful and 8 were successful.

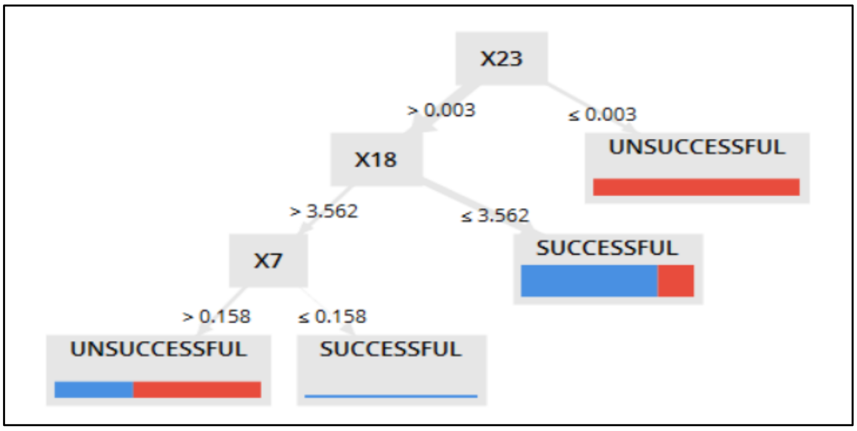


Figure 2. 2006 year CART decision tree image

Figure 3 shows the CART decision tree of 2007

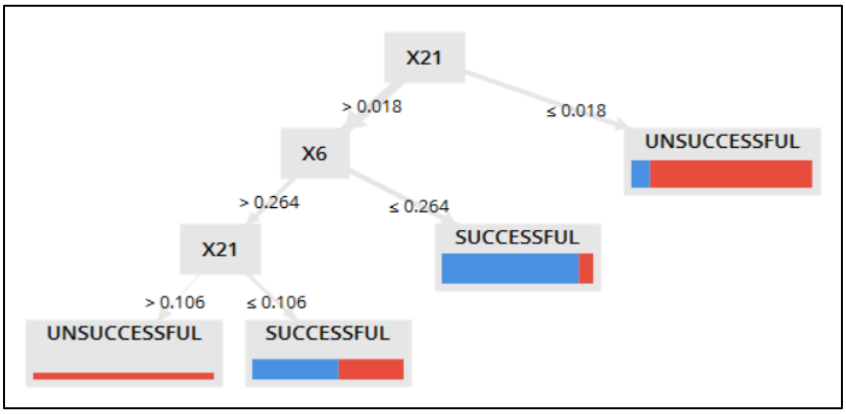


Figure 3. 2007 year CART decision tree image

Figure 4 shows the CART decision tree of 2008.

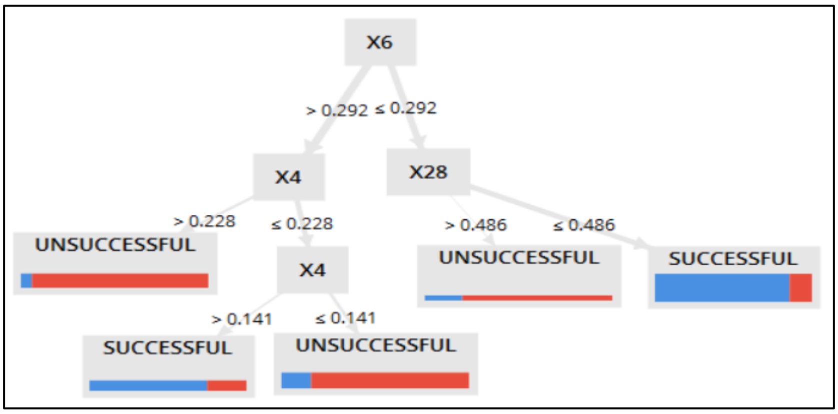


Figure 4. 2008 year CART decision tree image

C5.0 Decision-making algorithm analysis and findings

C5.0 is a new decision tree algorithm designed by Quinlan based on C4.5 and containing all the functionality of C4.5 (Chen, 2011a, p. 4515). The C5.0 algorithm is a supervised classification algorithm. The decision tree is trained with a data set that knows the target variable and a model is created and the performance of the model is measured by testing the generated dependent model on a new data set (Yakut & Elmas, 2013, p. 246). 70% training 30% test, 90% training 10% test data set separation, 80% training 20% test data set separation analyzes performed with different parameter values of data set. Examined whether or not all variables or with variables obtained from discriminant analysis will be performed, stratified sample selection or shuffle sample selection model tests. It is made evaluation that about the number of branches and all parameter values.

Low prediction results were obtained for all models with all variables. The model with the highest classification rate after a large number of model tests was the model operated with 8 independent variables which were variables determined separately for each year from the forward stepwise discriminant analysis. As in the study of Yakut (2012), it was observed that the exlusion of some variables from the model increased the prediction power of methods used in the analysis. As a final model, 70% training, 30% test data set, 10-fold cross-validation with stratified sampling model has been used. The decision tree algorithm was first performed at different pruning levels and the tree that gave the best performance was chosen. Table 7 gives the independent variables determined in the forward stepwise discriminant analysis and used in the C5.0 decision tree.

Table 7.
Independent Variables Determined in the Forward Stepwise Discriminant Analysis

X4	Inventory to Total Assets Ratio
X6	Short Term Debt to Total Assets Ratio
X13	Receivables Turnover Ratio
X20	Operating Margin Ratio
X21	Net Profit Margin
X22	Return On Total Assets (ROTA)
X24	Return On Equity (ROE)
X27	Being Audited by Four Major Audit Firms (Pricewaterhousecoopers-Deloitte Touche Tohmatsu- Kpmg- Ernst And Young)

Table 8 gives analysis parameters entered in the program by C5.0 classification algorithm parameter optimization.

Table 8.
C5.0 Classification Algorithm Parameter Optimization Analysis Parameters

Specified Parameters for Analysis	Explanation			
Data Set Separation	70% Training, 30% Test Data Set			
Validation Type	10 Fold Cross-Validation			
Number of Variables	8 Variables Determined by Stepwise Discriminant Analysis			
Sample Selection	Stratified Sample Selection			
Split Criterion	Information Gain (Entropy)			
Specified Parameters for Analysis	**Minimum**	**Maximum**	**Steps**	**Scale**
Minimal Size For Split	1.0	4.0	10	Linear
Minimal Leaf Size	1.0	2.0	10	Linear
Minimal Gain	0	20	10	Linear
Maximal Depth	-1	20	10	Linear
Confidence Level	0	0.25		
Number of Pre-Pruning	0	3	-	-

Table 9 gives the selected parameters which constitute the best classification determined by parameter optimization of the C5.0 classification algorithm.

Table 9.
The Parameters Giving the Highest Classification Accuracy Determined by the C5.0 Algorithm Performance Results and Parameter Optimization

Parameters	2006	2007	2008
Accuracy	76.32%	78.95%	86.84%
Classification Error	23.68%	21.05%	13.16%
Kappa	0.526	0.579	0.737
Weighted Mean Recall	76.32%	78.95%	86.84%
Weighted Mean Precision	76.99%	79.27%	89.58%
Minimal Size For Split)	3	2	3
Minimal Leaf Size	1	1	1
Minimal Gain	0	0	0
Maximal Depth	16	-1	3

Figure 5 shows the C5.0 decision tree 3 years before the failure (year 2006). The independent variable is the “X21 Net Profit Margin”, which separates the classes first at the root of the C5.0 decision tree at the root of a successful / failed business. “X21 net profit margin” was found to be unsuccessful in 100% of 22 enterprises whose net profit margin was less than or equal to 0.002. Of the 66 enterprises with a net profit margin of X21 greater than 0.002, 44 were successful and 22 were unsuccessful. Geng et al. (2015) found da Net Profit Margin da among important variables.

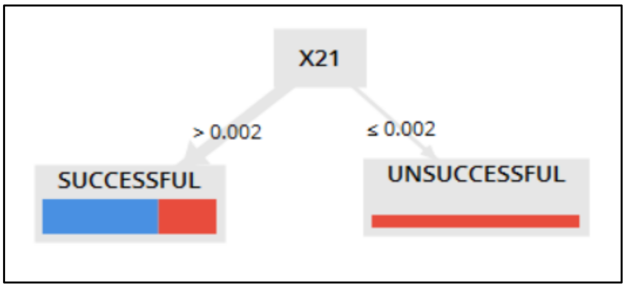


Figure 5. 2006 year C5.0 decision tree image

Figure 6 shows the C5.0 decision tree of 2007.

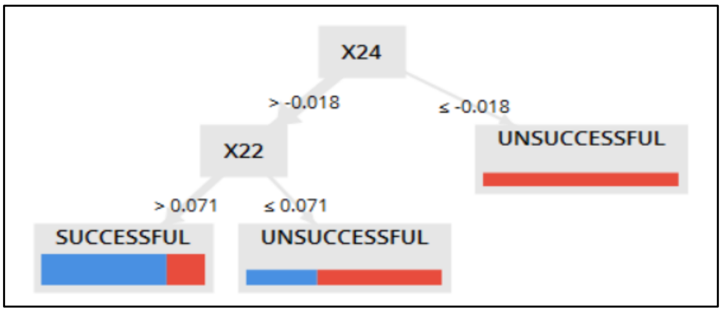


Figure 6. 2007 year C5.0 decision tree image

Figure 7 shows the C5.0 decision tree of 2008.

Figure 7. 2008 year C5.0 decision tree image

Artificial Neural Network Analysis and Findings

Artificial Neural Network (ANN) method is used to learn or adapt brains simulated based on a computer model as a response to external inputs. The network responds correctly to all samples during training does not show that the performance is good. The expected performance of the learning network should be measured against the sample that it has not seen before (Öztemel, 2012, ss. 90-91). In order for ANN not to misunderstand, the number of successful and unsuccessful businesses must be equal (Çelik ve diğerleri 2017, s. 149). Therefore, the number of successful and unsuccessful businesses are equal in this study. All data values were scaled by minimum-maximum normalization (0, 1) interval by following the Li et al., (2010) study for ANN analysis. ANN analysis was carried out using a hidden layer to be used for predictions with the normalized data set and the number of optimized training cycles, learning rate, momentum, sigmoid function. This classifier with a hidden layer achieved better performance than multiple hidden layers. Thus, the findings of the single hidden layer model are presented.

Mathematically, the process performed by a hidden or output neuron with n inputs is defined as:

$$\text{Output} = f\left(\sum_{i=1}^{n} w_i x \; Input\ i\right) \qquad (1)$$

Here w_i represents the weight given for input i and f represents logistic transformation. In essence, the weights of the inputs at each node are adjusted using a recursive process to match the inputs closer to the output in the training example (Jackson & Wood, 2013, p. 190).

In order to better understand the effect of the different ratios between the training and test set on the generalization capacity, the highest classification success was found as 70:30 composition in the analyzes with 60:40, 75:25, 70:30 samples. In the determination of training and test set rates in the literature, it was divided into 70% of the training and 30% of the test set (eg. Koç and Ulucan (2016)).

In the ANN model created in this study, weights stored in neurons for each independent variable are given in the table for 3 years. In this study, parameter optimization was performed and the lowest and highest values required for the individual parameters were entered in the program.

In order to determine the best ANN network model, trial and error method is widely used and many tests are performed. As in Yakut (2012), it is observed that the removal of some variables from the model increases the predictive power of the methods used in the analysis. The ANN model, which was constructed with all variables, obtained lower classification performance than the 8 variable model obtained from discriminant stepwise analysis. Therefore, 8 variables were used for 3 years by discriminant stepwise analysis.The model with the highest classification percentage after tens of different model trials; a hidden layer, 70% training, 30% is the model established with the separation of the test data set.

Determining high values for the learning coefficient in ANN will cause the system to memorize, which is not desirable. In our study, parameter optimization was performed and the parameters of the parameters with the highest classification result were obtained by entering the minimum and maximum values desired in the parameters determined separately. Özdağoğlu, Özdağoğlu, Gümüş & Kurt Gümüş (2017) studies were carried out based on the values of three important parameters such as artificial neural network model, learning speed, momentum and training speed. Thus, the highest performance predictions have been obtained based on the given input set. For all learning algorithms, training and test datasets are based on 10-fold cross-validation. All data values were scaled between 0 and 1 with minimum-maximum normalization following the study of Li et al., (2010) for ANN analysis. This classifier with a hidden layer has achieved better performance than multiple hidden layers. Thus, the findings of the single hidden layer model are presented. In the study of Çelik (2009), it was observed that the 3-tier structure produced more accurate results than the more or less layered structures. In this study, a multilayered perceptron (MLP) with a single layer was used to perform the classification task. In the NN model created in this study, the weights stored in neurons for each independent variable were removed and given as table in the appendix of the study for 3 years. In this study, the lowest and highest values, desired to be tried separately in the parameters determined by parameter optimization, were entered into the program. The model parameters are given in Table 10.

Table 10.
Artificial Neural Network Parameters with the Highest Performance

Network Type	Multilayer Perseptron			
Learning Algorithm	Back Propagation			
Learning Rule	Momentum			
Number of Nodes in Input Layer	8			
Number of Hidden Layers	1			
Number of Nodes in Hidden Layer	6			
Number of Output Layer Nodes	2; Successful, Unsuccessful Categorical			
Variable Selection	8 Independent Variables Selected by Stepwise Discriminant Method			
Classification of Data	70% Training Set			
	30% Test Set			
Sample Selection Type	Stratified Sample Selection			
Activation Function	Sigmoid			
Learning Ratio	Minimum: 0,00	Maximum: 0,80	Steps: 10	Scale: Linear
Momentum	Minimum: 0,00	Maximum: 0,80	Steps: 10	Scale: Linear
Number of Training Cycle	Minimum: 1,00	Maximum: 500	Steps: 10	Scale: Linear

Compared to statistical methods, a large number of sample data is needed to create a stable ANN model. ANN is often criticized by practitioners for its complicated network structure that is difficult to understand because it looks like a black box for decision makers. In order to overcome this disadvantage, a decision table in the form of explanatory rules can be created by obtaining the learned information embedded in the networks (Sun et al., 2014, p. 45). In our study, ANN model was added to each of the independent variables by adding weights stored in neurons for the years 2006, 2007 and 2008 respectively.

Figure 8 shows the model image of the artificial neural network 2006, 2007, 2008.

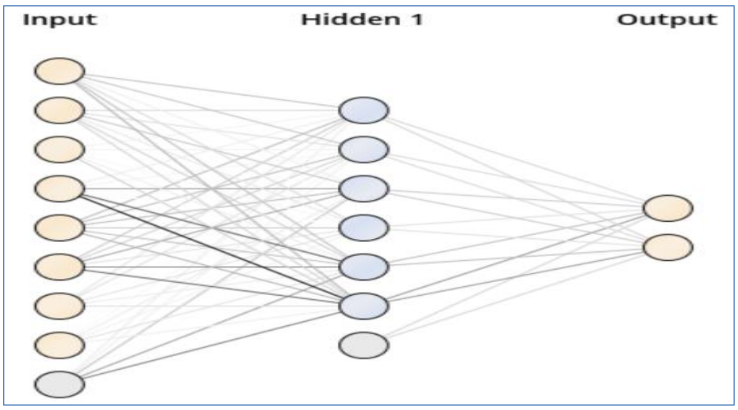


Figure 8. Artificial neural network 2006, 2007, 2008 years model image

Table 11 gives the parameters selected by the artificial neural network with the highest performance.

Table 11.
Parameters Selected by the Artificial Neural Network with the Highest Performance

Parameters	2006	2007	2008
Accuracy	81.58%	84.21%	92.11%
Classification Error	18.42%	15.79%	7.89%
Kappa	0.632	0.684	0.842
Weighted Mean Recall	81.58%	84.21%	92.11%
Weighted Mean Precision	82.39%	84.21%	93.18%
Learning Rate	0.16	0.72	0.80
Momentum	0.80	0.24	0.00
Training Cycles	151	101	151

Comparison and Evaluation of the Results

The overall classification accuracy from the highest to the lowest of 3 years prior to successful-failure year (for 2006) is 84.21% for CART, 81.58% for ANN and 76.32% for C5.0, respectively. The overall classification accuracy from the highest to the lowest of 2 years prior to successful-failure year (for 2007) is 86.84% for CART, 84.21% for ANN, 78.95% for C5.0, respectively. The overall classification accuracy from the highest to the lowest of 1 year prior to successful-failure year (for 2008) is 92.11% for CART, 92.11 for ANN and 86.84% for C5.0, respectively. When become distant from failure year, the CART classification and regression tree has the highest classification accuracy 3 years ago. When Table 12 is analyzed, ANN and CART models are notable in terms of their ability to predict the upcoming financial failure of unsuccessful businesses with a 100.00% classification accuracy from a year ago.The CART decision tree can be regarded as positive in all years, with higher accuracy.

Table 12.
Performance Results of the Classification Methods Used in the Study 1, 2, 3 Years Ago

Classification Method	Condition	3 Years Ago (2006)%	2 Years Ago (2007)%	1 Year Ago (2008)%
CART	Unsuccessful	94.74	89.47	100.00
	Successful	73.68	84.21	84.21
	Total	84.21	86.84	92.11
C5.0	Unsuccessful	84.21	73.68	73.68
	Successful	68.42	84.21	100.00
	Total	76.32	78.95	86.84
Artificial Neural Network	Unsuccessful	89.47	84.21	100.00
	Successful	73.68	84.21	84.21
	Total	81.58	84.21	92.11

The CART model, which was established in our study, showed a higher performance in 3 years than ANN in the same way as the findings of the study of Doğrul (2009). When the results of our study obtained pursuing the study of Özdağoğlu et al. (2017) was compared with the results of previous studies, it was revealed that the performance of the model depended on the data set, the algorithm and the variables, that is, the selected financial ratios.

Doğrul (2009) found the best models for predicting financial failure as a result of analyses as LA=CART>NN one year ago, CART>LA>NN two years ago, CART>NN>LA three years ago. When this study was examined in terms of classification accuracy, CART=NN>C5.0 was found one year ago, CART>NN>C5.0 two years ago and CART>NN>C5.0 three years ago. In addition, Doğrul found that CART had the highest classification accuracy in 3 years, while CART had the highest classification accuracy for only 2006 in this study.

Gepp et al. (2010) concluded that the CART model has the most consistent predictive capability over misclassification costs. This result is consistent with our findings. Related work in contrast to the conclusion that the C5.0 algorithm produces more complex trees than CART, the C5.0 algorithm has formed a more simple tree for three years according to CART. Chen (2011a) found that CART analysis predicted with high accuracy one year ago is consistent with our analysis results. In the related study, it was concluded that the accuracy rate of C5.0 algorithm was better than CART. When this finding was compared with our study results, CART decision tree showed superior prediction performance over the C5.0 decision tree within 3 years in our analyses.

In the study of Chen (2011a) it was found that C5.0 and CART provided the best prediction performance for upcoming bankruptcy and this result is also valid for CART decision tree in our study. When this finding was compared with the results in this study, the CART decision tree in our analysis showed superior prediction performance in three years than the C5.0 decision tree.

Sun et al. (2014) found that CART had a higher prediction accuracy than C5.0. This result is compatible with our study within 3 years. This result is compatible with these study results within three years. In the related study, it was stated that C5.0 and CART obtained higher classification accuracy than ANN. This result applies to the CART decision tree in this study. In this study, it was seen that CART method had higher prediction power two and three years before NN. However, in this study, C5.0 has never had higher classification power than NN. In our study, CART method seems to have higher prediction power than ANN for 2 and 3 years ago. However, in our study, C5.0 has never had higher classification power than ANN.

In the study of Yakut (2012), similar to the findings of our study, when the ANN and C5.0 methods went backwards from the years of failure, the estimation results decreased and the estimation results reached higher rates as the approached the failure year. In the study of Yakut (2012), the C5.0 decision tree comes immediately after ANN in all periods in terms of accurate classification ratio, which is consistent with our analysis results. In the study of Kılıç (2011), that the ANN prediction performance is higher than the C5.0 decision tree performance is consistent with our analysis results.

Conclusion

In this study, Artificial Neural Networks (NN), C5.0 Classification Algorithm, Classification and Regression Trees (CART) analyses were used to predict the financial success/failure of 126 businesses that are operating in the BIST (Borsa İstanbul) Manufacturing Industry Sector. The data contains the years 2006 to 2009.

According to the analysis results, the highest classification accuracy was obtained by CART (84.21%) analysis 3 years prior to the successful-failure year (for 2006). CART analysis has the highest classification accuracy of 86.84% 2 years prior to the successful-failure year (for 2007). The highest classification accuracy was obtained by CART (92.11%) and ANN (92.11%) analysis 1 year prior to the successful-failure year (for 2008). The CART had the highest classification accuracy three years ago. ANN and CART models are notable in terms of their ability to predict upcoming financial failure of unsuccessful businesses with 100% classification accuracy from a year ago.

According to the decision trees formed in the CART and C5.0 analyzes, as the important variables in predicting 1, 2 and 3 years prior to failure "Return On Assets Ratio, Tangible Fixed Assets Turnover Rate, Financial Leverage Ratio, Ratio of Short-Term Liabilities to Total Assets, Free Float Rate, Ratio of Stocks to Total Assets, Net Profit Margin, Return On Equity Ratio" were found.

CART predicted unsuccessful businesses with the highest classification accuracy 1,2 and 3 years ago (100%, 89.47% and 94.74%, respectively) considering the fact that the Type I error (unsuccessful in reality but mispredicted as successul) costs 20 to 38 times (Type I error is the lowest) more of the Type II error (successful in reality but mispredicted as unsuccessful).

Among the important variables determined by forward stepwise discriminant analysis are the qualitative independent variable "X27, Supervised by the Four Major Audit Companies". In addition, in the creation of 2008 CART Decision Tree, qualitative independent variable "X28, Free Float Rate" was found as the second important variable after the root node. According to the relevant results, "supervision by four major audit companies and free float rate" were identified as important non-financial variables in the analysis for classification and prediction.

Early warnings of financial failure in businesses is an interesting issue for both practitioners and academicians. In order to predict the failure of the business, many studies have been done in the literature that analyze public disclosure by using classification techniques. Numerous conclusions have been found from these studies. However, the reliability of the presented findings is limited because very little data is used. Future researches are thought to require larger data sets to be analyzed in order to avoid this problem. In the future studies, more non-financial indicators can be used in the models to be created in order to increase the prediction accuracy. In the future, researchers can build models with data from different countries and compare these models for financial failure prediction. In future research, more sensitive variables can be obtained with more dependent variables by using fuzzy logic instead of binary classification of dependent variable as financial success and financial failure. It would also be interesting to apply the methods used in this study to large amounts of data from service industries and trading companies and to see how different the prediction values in the rates are. In future studies, filter and wrapper feature selection methods can be used in the variable selection stage as they are the most used methods in bankruptcy prediction and credit scoring. In addition to using single classification techniques to develop prediction models, combining multiple classifiers with bagging and boosting combination methods and examining the performance of community classifiers can be considered.

List of abbreviations

MDA	Multiple Discriminant Analysis
ANN	Artificial Neural Networks
CART	Classification and Regression Trees
C5.0	C5.0 decision rule derivation algorithm
LA	Logistic Regression Analysis
DT	Decision Tree
ST	special treatment
BIST	stock market istanbul
SVM	Support Vector Machine

References

Akgüç, Ö. (1997). *Financial Management*, İstanbul: Avcıol publications.

Aksoy, B. & Ertaş, F. C. (2016). Effects of corporate merger and acquisition on target company financial performance. *Academic View Review, 54,* 772-786.

Aktaş, Ramazan, M. Mete Doğanay & Birol Yıldız (2003). Forecasting financial failure: comparison of statistical methods and artificial neural network. *Ankara University SBF Journal*, *58*(4), 1-24. Access address: https://app.trdizin.gov.tr/publication/paper/detail/TWpZME1URT0

Altman, E. I., T. Kant & T. Rattanaruengyot (2009). Bankruptcy performance: Avoiding", *Journal of Applied Corporate Finance*, *21*(3), 50-65. Access address: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1495296

Baixauli, J. S. & Mo´dica-Milo, A. (2010). The bias of unhealthy smes ın bankruptcy prediction models. *Journal of Small Business and Enterprise Development*, *17*(1), 60-77. Access address: https://www.emerald.com/insight/content/doi/10.1108/14626001011019134/full/pdf

Beaver, William H. (1968). Market prices, financial ratios, and the prediction of failure. *Journal of Accounting Research*, *6*(2), 179-192. Access address: https://www.jstor.org/stable/2490233?seq=1#metadata_info_tab_contents

Bee, Thai Siew & Mehdi Abdollahi (2013). Corporate failure prediction: Malaysia's emerging market. *The International Journal of Finance*, *25*(4), 7985-8008. Access address: https://www.econbiz.de/Record/corporate-failure-prediction-malaysia-s-emerging-market-bee-thai-siew/10010382959

Bilir, H.. (2015). Definition and market oriented solution of financial distress: Debt structuring, asset sales and new capital injection. *Socioeconomic*,1, 9-24. Access address: http://eds.a.ebscohost.com/eds/pdfviewer/pdfviewer?vid=0&sid=8f6841cb-c699-4e01-959d-f18901710107%40sessionmgr4007

Brealey, R. & Myers, S. C., Trans. Bozkurt, A. J., Arıkan T., & Doğukanlı, H. (2007), *Fundamentals of Business Finance,* İstanbul: MC Graw Hill and Literature publications.

Chandra, D. K., Ravi, V. & Bose, I. (2009). Failure prediction of DOTCOM companies using hybrid intelligent techniques. *Expert Systems with Applications*, *36,* 4830-4837. Access address: https://www.sciencedirect.com/science/article/pii/S0957417408002741

Chen, J., Marshall, B. R., Zhang, J., & Ganesh, S. (2006). Financial distress prediction in China. *Review of Pacific Basin Financial Markets and Policies*, *9*(2), 317-336. Access address: https://econpapers.repec.org/article/wsirpbfmp/v_3a09_3ay_3a2006_3ai_3a02_3an_3as0219091506000744.htm

Chen, Mu-Y.. (2011). Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches. *Computers and Mathematics with Applications*, 62, 4514-4524. Access address: https://www.sciencedirect.com/science/article/pii/S0898122111008947

Chen, Mu-Y.. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. *Expert Systems with Applications*, *38,* 11261-11272. Access address: https://www.sciencedirect.com/science/article/pii/S0957417411003976

Chen, Mu-Y.. (2014). Using a hybrid evolution approach to forecast financial failures for Taiwan-listed companies. *Quantitative Finance*, *14*(6), 1047-1058. Access address: https://www.tandfonline.com/doi/full/10.1080/14697688.2011.618458

Chuang, Chun-L.. (2013). Application of hybrid case-based reasoning for enhanced performance in bankruptcy prediction. *Information Sciences*, *236,* 174-185. Access address: https://www.researchgate.net/publication/256721212_Application_of_hybrid_case-based_reasoning_for_enhanced_performance_in_bankruptcy_prediction

Çalış, A., Kayapınar, S. & Çetinyokuş, T. (2011). An application on computer and internet security with decision tree algorithms in data mining. *Journal of Industrial Engineering*, *25*(3-4), 2-19. Access address: https://dergipark.org.tr/en/download/article-file/752270

Çelik, M. K. (2010). Prediction of Financial Failure of Banks with Traditional and New Methods. *Management and Economy*, *17*(2), 129-143. Access address: https://dergipark.org.tr/tr/pub/yonveek/issue/13693/165725

Çelik, U., Akçetin, E., & Gök, M.. (2017), *Data mining with rapidminer*, İstanbul: Pusula publications.

Delen, D., Kuzey, C. & Uyar, A.. (2013). Measuring firm performance using financial ratios: A decision tree approach. *Expert Systems with Applications*, *40,* 3970-3983. Access address: https://www.sciencedirect.com/science/article/pii/S0957417413000158

Dener, M. & Orman, A. (2009). Open source data mining programs: sample practice in WEKA", Conference: Conference: XI. Academic Informatics Conference, Şanlıurfa, January 2009, 1-11. Access address: https://ab.org.tr/ab09/kitap/dener_dorterler_AB09.pdf

Doğrul, Ü. (2009), *Financial failure and prediciton of financial failure: An application on industry companies trading in İstanbul Stock Exchange,* Mersin University, Institute of Social Sciences, Mersin, (Unpublished Master's Thesis).

Gaganis, C. (2009). Classification techniques for the identification of falsified financial statements: a comparative analysis. *Intelligent Systems in Accounting, Finance and Management*, *16*, 207-229. Access address: https://www.researchgate.net/publication/ 220613891

Gallego-Garcia, A., Mures-Quintana, M.-J.. (2012). Business failure prediction models: finding the connection between their results and the sampling method. *Preparation of Electronic Manuscripts for Publication*, 157-168. Access address: https://www.researchgate.net/publication/286361899

Geng, R., Bose, I. & Chen, X. (2015). Prediction of financial distress: an empirical study of listed chinese companies using data mining. *European Journal of Operational Research*, *241*, 236-247. Access address: https://www.sciencedirect.com/science/article/pii/S0377221714006511

Gepp, A., Kumar, K. & Bhattacharya, S. (2010). Business failure prediction using decision tree. *Journal of Forecasting*, 29, 536-555. Access address: https://ideas.repec.org/a/jof/jforec/v29y2010i6p536-555.html

Gepp, A. & Kuldeep K. (2015). Predicting financial distress: a comparison of survival analysis and decision tree techniques. *Procedia Computer Science*, *54,* 396-404. Access address: https://www.sciencedirect.com/science/article/pii/S1877050915013708

Glezakos, M., Mylonakis J. & Oikonomou K. (2010). An empirical research on early bankruptcy forecasting models: Does logit analysis enhance business failure predictability? *European Journal of Finance and Banking Research*, *3*(3), 1-15. Access address: http://globip.com/contents/articles/european-vol3-article1.pdf

Iwan, M. (2005). Bankruptcy prediction model with zeta optimal cut-off score to correct type ı errors. *Gadjah Mada International Journal of Business*, *7*(1), 41-68. Access address: https://www.researchgate.net/publication/277071919_Bankruptcy_prediction_model_with_zetac_optimal_cut-off_score_to_correct_type_I_errors

Jackson, Richard H.G. & Wood, A. (2013). The performance of insolvency prediction and credit risk models in the UK: a comparative study. *The British Accounting Review*, *45,* 183-202. Access address: https://www.researchgate.net/publication/259133786

Jardin, P. D. (2009). Bankruptcy prediction models: how to choose the most relevant variables. *Bankers, Markets and Investors*, 98, 39-46. Access address: https://www.researchgate.net/publication/235643766

Jardin, P. D. (2012). The influence of variable selection methods on the accuracy of bankruptcy prediction models. *Bankers, Markets and Investors*, *116,* 20-39. Access address: https://www.researchgate.net/publication/235643793

Kılıç, Y. (2011), *Financial faılure prediction using data mining: an application in İstanbul Stock Exchange*, Gaziantep University, Institute of Social Sciences, Gaziantep, (Unpublished Master's Thesis). Access address: https://tezarsivi.com/finansal-basarisizlik-tahmininde-veri-madenciliginin-kullanilmasi-imkbde-bir-uygulama

Kılıç, Y. & Seyrek, İ. H. (2016). Use of artificial neural networks in forecasting financial failure: an application in manufacturing sector. *www.researchgate.net.* Berlin, 2016, s. 1-15. Access address: https://www.researchgate.net/publication/296705043

Koç, S. & Ulucan, S.. (2016). Testing of Altman z methods which is used for detecting of financial failures with fuzzy logic (anfis) technique: a case study on technology and textile sector. *Revenue and Finance Articles,* 106, 147-167. Access address: https://dergipark.org.tr/en/download/article-file/259955

Le, H. H. & Viviani, J.-L. (2018). Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. *Research in International Business and Finance*, 44, 16-25. Access address: https://ideas.repec.org/a/eee/riibaf/v44y2018icp16-25.html

Li, H. & Sun, J. (2013). Predicting business failure using an rsf-based case-based reasoning ensemble forecasting method. *Journal of Forecasting*, *32,* 180-192. Access address: https://www.researchgate.net/publication/263113267

Liang, D., Tsai, C.-F. & Wu, H.-T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems*, *73,* 289-297. Access address: https://isslab.csie.ncu.edu.tw/download/publications/5.pdf

Lussier, R. N. (1995). A nonfinancial business success versus failure prediction model for young firms. *Journal of Small Business Management*, *1995,* 8-20. Access address: https://www.questia.com/library/journal/1G1-16787401/a-nonfinancial-business-success-versus-failure-prediction

Misund, B. (2017). Financial ratios and prediction on corporate bankruptcy in the Atlantic salmon industry. *Aquaculture Economics & Management*, *21*(2), 241-260. Access address: https://www.researchgate.net/publication/304611502

Muller, G.H., Steyn-Bruwer, B.W. & Hamman, W.D. (2009). Predicting financial distress of companies listed on the JSE- a comparison of techniques. *S.Afr.J.Bus.Manage*, *40*(1), 21-32. Access address: https://www.researchgate.net/publication/286881166

Okay, K. (2015). *Predicting business failures in non-financial turkish companies*, İhsan Doğramacı Bilkent University Graduate School of Economics and Social Sciences, Ankara (Unpublished Masters Thesis). Access address: http://www.thesis.bilkent.edu.tr/0006970.pdf

Öcal, N., Ercan, M. K. & Kadıoğlu, E. (2015). Predicting financial failure using decision tree algorithms: an empirical test on the manufacturing industry at Borsa İstanbul. *International Journal of Economics and Finance*, *7*(7), 189-206. Access address: https://www.researchgate.net/publication/281269848

Özdağoğlu, G., Özdağoğlu, A., Gümüş, Y. & Kurt-Gümüş, G. (2017). The application of data mining techniques in manipulated financial statement classification: the case of Turkey. *Journal of AI and Data Mining*, *5*(1), 67-77. Access address: http://jad.shahroodut.ac.ir/article_664.html

Özkan, M. & Boran, L. (2014). Usage of data mining at financial decision making. *Çankırı Karatekin University Faculty of Economics and Administrative Sciences Journal*, *4*(1), 59-82. Access address: https://dergipark.org.tr/en/download/article-file/382288

Öztemel, E. (2012). *Artificial Neural Networks*, İstanbul: Papatya publications.

Sayılgan, G. & Ece, A. (2016). Bankruptcy postponement and panoramic analysis on bankruptcy postponement litigations in Turkey between 2009-2013. *Revenue and Finance Articles, 105*, 47-74. Access address: https://dergipark.org.tr/tr/download/article-file/259954

Selimoğlu, S. & Orhan, A. (2015). Measurement of financial failure by using ratio analysis and discriminant analysis: a research on woven, clothing and leather enterprises in BIST. *Journal of Accounting and Finance*, *April*, 21-40. Access address: https://www.researchgate.net/publication/323406251

Sun, J., Li, H., Huang, Q. H. & He, K-Yu. (2014). Predicting financial distress and corporate failure: a review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge-Based Systems*, *57*, 41-56. Access address: https://www.sciencedirect.com/science/article/pii/S0950705113003869

Tuncay, M. (2011). Corporate financial restructuring after financial crises. *Journal of Accounting and Finance*, 103-118. Access address: https://dergipark.org.tr/tr/download/article-file/426903

Tsai, C. F., Hsu, Y. F. & Yen D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, *24*, 977-984. Access address: https://www.sciencedirect.com/science/article/pii/S1568494614004128

Uzar, C. (2013). *The usage of data mining technology in financial ınformation system: an application on Borsa Istanbul*, Dokuz Eylül University Social Sciences Institute, İzmir (Unpublished PhD Thesis).

Wu, W., Lee, V. C. S., & Tan T. Y. (2006). Data preprocessing and data parsimony in corporate failure forecast models: evidence from Australian materials industry. *Accounting and Finance*, *46*, 327-345. Access address: https://www.researchgate.net/publication/4737947

Yakut, E. & Elmas, B. (2013). Estimating financial failure of enterprises with data mining and discriminant analysis. *Afyon Kocatepe University İİBF Journal*, *15*(1), 237-254. Access address: https://dergipark.org.tr/tr/pub/eab/issue/39985/475267

Yakut, E. (2012). *The Comparison of the classification successes of the artifical neural networks through data mining techniques of c5.0 algorithm and supporting vector machines: an application in manufacturing sector,* Atatürk University Social Sciences Institute, Erzurum (Unpublished PhD Thesis). Access address: https://tezarsivi.com/veri-madenciligi-tekniklerinden-c50-algoritmasi-ve-destek-vektor-makineleri-ile-yapay-sinir-aglarinin-siniflandirma-basarilarinin-karsilastirilmasi-imalat-sektorunde-bir-uygulama

Yapraklı, T. Ş. & Erdal, H.. (2016). Firm failure prediction: a case study based on machine learning. *Journal of Information Technologies*, *9*(1), 21-31. Access address: https://dergipark.org.tr/tr/pub/gazibtd/issue/26691/280780

Yip, A. Y. N. (2006). Business failure prediction: a case-based reasoning approach. *Review of Pacific Basin Financial Markets and Policies*, *9*(3), 491-508. Access address: https://www.Business failure prediction: a case-based reasoning approach. Review of Pacific Basin Financial Markets and Policiesworldscientific.com/doi/abs/10.1142/S021909150600080X

Appendix 1. Artificial neural network weights (Year 2006)

Input Layer	Hidden Layer						Output Layer	
Independent Variable	Node1 (Sigmoid)	Node2	Node3	Node4	Node5	Node6	Class'1'	Class'0'
X4	X4: -2.341	X4: -2.174	X4: -0.578	X4: -1.396	X4: 2.295	X4: 2.788	Node1: -1.610	Node1: 1.590
X6	X6: 1.383	X6: 1.181	X6: 2.082	X6: 0.844	X6: 1.726	X6: 1.620	Node2: -1.468	Node2: 1.504
X13	X13: -0.393	X13: -0.471	X13: 0.115	X13: -0.243	X13: 0.619	X13: 1.731	Node3: -2.029	Node3: 2.021

X20	X20: - 0.632	X20: - 0.345	X20: - 3.203	X20: - 0.003	X20: - 5.975	X20: - 8.895	Node4: - 1.159	Node4: 1.141
X21	X21: - 2.579	X21: - 2.445	X21: - 2.503	X21: - 1.822	X21: - 2.319	X21: - 2.900	Node5: - 2.438	Node5: 2.438
X22	X22: - 1.952	X22: - 1.705	X22: - 3.282	X22: - 1.055	X22: - 4.062	X22: - 5.565	Node6: - 3.753	Node6: 3.761
X24	X24: - 0.940	X24: - 1.100	X24: - 0.105	X24: - 1.044	X24: 0.693	X24: 1.050	Threshold: 1.569	Threshold: - 1.569
X27	X27: 0.090	X27: 0.237	X27: - 0.678	X27: 0.466	X27: - 1.090	X27: - 0.576		
	Bias: 0.667	Bias: 0.479	Bias: 2.232	Bias: - 0.091	Bias: 3.516	Bias: 4.766		

Appendix 2. Artificial neural network weights (Year 2007)

Input Layer	**Hidden Layer**						**Output Layer**	
Independent Variable	**Node1 (Sigmoid)**	**Node2**	**Node3**	**Node4**	**Node5**	**Node6**	**Class'1'**	**Class'0'**
X4	X4: - 4.080	X4: 0.360	X4: 1.660	X4: 4.518	X4: - 3.884	X4: - 0.665	Node1: - 3.428	Node1: 3.437
X6	X6: 0.061	X6: - 0.274	X6: - 3.081	X6: 4.391	X6: - 8.156	X6: 0.463	Node2: 0.940	Node2: - 0.904
X13	X13: 2.520	X13: 0.457	X13: - 0.608	X13: 1.495	X13: 0.236	X13: 0.862	Node3: 0.808	Node3: - 0.821
X20	X20: 0.457	X20: 1.199	X20: 0.277	X20: - 0.735	X20: 1.707	X20: 0.312	Node4: - 3.925	Node4: 3.927
X21	X21: 0.490	X21: 1.112	X21: 0.095	X21: - 0.732	X21: 1.092	X21: 0.387	Node5: 3.061	Node5: - 3.055
X22	X22: - 5.203	X22: 1.098	X22: 3.029	X22: - 3.261	X22: 11.523	X22: - 1.002	Node6: - 0.546	Node6: 0.486
X24	X24: - 3.384	X24: 0.507	X24: 1.164	X24: - 5.686	X24: 3.692	X24: - 0.715	Threshold: - 1.205	Threshold: 1.216
X27	X27: 1.580	X27: 0.514	X27: - 0.478	X27: 0.777	X27: 2.470	X27: 0.885		
	Bias: - 2.280	Bias: - 0.773	Bias: - 0.035	Bias: 0.582	Bias: - 1.495	Bias: - 0.788		

Appendix 3. Artificial neural network weights (Year 2008)

Input Layer	**Hidden Layer**						**Output Layer**	
Independent Variable	**Node1 (Sigmoid)**	**Node**	**Node3**	**Node4**	**Node5**	**Node6**	**Class'1'**	**Class'0'**
X4	X4: 4.873	X4: - 0.786	X4: - 2.105	X4: 2.235	X4: - 0.668	X4: 0.035	Node1: - 3.147	Node1: 3.141
X6	X6: 3.581	X6: 6.297	X6: - 2.313	X6: 3.558	X6: 3.739	X6: 1.143	Node2: - 4.611	Node2: 4.665

X13	X13: -2.038	X13: -0.631	X13: 0.137	X13: 1.149	X13: 0.000	X13: 0.551	Node3: -4.138	Node3: 4.136
X20	X20: -0.211	X20: 0.087	X20: -1.196	X20: -1.490	X20: 0.129	X20: -0.188	Node4: -3.711	Node4: 3.720
X21	X21: -2.721	X21: -0.538	X21: -9.941	X21: -4.633	X21: -0.761	X21: -0.733	Node5: -2.601	Node5: 2.535
X22	X22: 3.140	X22: 8.514	X22: -9.569	X22: -4.040	X22: 4.714	X22: 1.069	Node6: -0.753	Node6: 0.717
X24	X24: -0.944	X24: 1.604	X24: -4.674	X24: -2.658	X24: 0.587	X24: -0.323	Threshold: 3.402	Threshold: -3.398
X27	X27: 1.588	X27: 0.163	X27: 0.628	X27: -2.937	X27: -0.246	X27: -0.269		
	Bias: 1.327	Bias: 0.473	Bias: -0.800	Bias: 1.967	Bias: 0.310	Bias: -0.671		

Appendix 4. Information on the researchers using the financial ratios used in our study in their studies.

Independent variables	Financial Ratios	Author name and Publication Year
X1	Current Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Torun (2007), Alfaro et al (2008), Ekinci vd. (2008), Li and Sun (2009), Çelik (2009), Akkaya vd. (2009), Doğrul (2009), Korol and Korodi (2010), Glezakos et al (2010), M.Y. Chen (2011), Li and Sun (2011), Terzi (2011), Elmas vd. (2011), Kılıç (2011), Jardin (2012), Galego et al (2012), Kılıç ve Seyrek (2012), Yakut (2012), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Geng et al (2015), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Okay (2015), Kaygın vd. (2016)
X2	Acid-Test (Liquidity) Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Torun (2007), Ekinci vd. (2008), Akkaya vd. (2009), Doğrul (2009), Çelik (2009), Korol and Korodi (2010), Chen (2011), Kılıç (2011), Terzi (2011), Elmas vd. (2011), Galego et al (2012), Yakut (2012), Kılıç ve Seyrek (2012), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Geng et al (2015), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Kaygın vd. (2016)
X3	Cash Ratio	Jianguo Chen et al (2006), Torun (2007), Akkaya vd. (2009), Çelik (2009), Doğrul (2009), Glezakos et al (2010), Divsalar et al (2011), M.Y. Chen (2011), Terzi (2011), Elmas vd. (2011), Jardin (2012), Yakut (2012), Galego et al (2012), Yakut ve Elmas (2013), Ural vd. (2015), Kaygın vd. (2016)
X4	Inventory to Total Assets Ratio	Torun (2007), Akkaya vd. (2009), Doğrul (2009), Chen (2011), Terzi (2011), Yakut (2012), Ural vd. (2015), Kaygın vd. (2016)
X5	Long Term Debt to Total Assets Ratio	Torun (2007), Akkaya vd. (2009), Doğrul (2009), Çelik (2009), Elmas vd. (2011), Jardin (2012), Galego et al (2012), Yakut ve Elmas (2013), Ural vd. (2015), Kaygın vd. (2016)

X6	Short Term Debt to Total Assets Ratio	Torun (2007), Akkaya vd. (2009), Korol and Korodi (2010), Elmas vd. (2011), Kılıç (2011), Jardin (2012), Yakut ve Elmas (2013), Geng et al (2015), Ural vd. (2015), Kaygın vd. (2016)
X7	Liabilities to Assets Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Torun (2007), Ekinci vd. (2008), Alfaro (2008), Li and Sun (2009), Doğrul (2009), Çelik (2009), Akkaya vd. (2009), Divsalar et al (2011), M.Y.Chen (2011), Kılıç (2011), Li and Sun (2011), Elmas vd. (2011), Jardin (2012), Galego et al (2012), Kılıç ve Seyrek (2012), Yakut (2012), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Geng et al (2015), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Okay (2015), Kaygın vd. (2016)
X8	Short Term Debt to Equity Ratio	Doğrul (2009), Li and Sun (2011), Elmas vd. (2011), Galego et al (2012), Yakut ve Elmas (2013), Kaygın vd. (2016)
X9	Fixed Assets to Equity Ratio	Jianguo Chen et al (2006), Torun (2007), Akkaya vd. (2009), Doğrul (2009), Li and Sun (2011), Elmas vd. (2011), Yakut ve Elmas (2013), Ural vd. (2015), Kaygın vd. (2016)
X10	Current Assets to Total Assets Ratio	Torun (2007), Alfaro et al (2008), Ekinci vd. (2008), Akkaya vd. (2009), Li and Sun (2009), Doğrul (2009), Divsalar et al (2011), M.Y.Chen (2011), Li and Sun (2011), Jardin (2012), Yakut (2012), Galego et al (2012), Lin and Liang (2014), Geng et al (2015), Kaygın vd. (2016)
X11	Debt to Equity Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Ekinci vd. (2008), Li and Sun (2009), Çelik (2009), Doğrul (2009), Li and Sun (2011), M.Y. Chen (2011), Elmas vd. (2011), Yakut (2012), Chuang (2013), Yakut ve Elmas (2013), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Kaygın vd. (2016)
X12	Inventory Turnover Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Torun (2007), Li and Sun (2009), Çelik (2009), Akkaya vd. (2009), Li and Sun (2011), Terzi (2011), Elmas vd. (2011), Jardin (2012), Yakut (2012), Kılıç ve Seyrek (2012), Chuang (2013), Yakut ve Elmas (2013), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Kaygın vd. (2016)
X13	Receivables Turnover Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Torun (2007), Li and Sun (2009), Çelik (2009), Akkaya vd. (2009), Li and Sun (2011), Terzi (2011), Elmas vd. (2011), Jardin (2012), Kılıç ve Seyrek (2012), Yakut (2012), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Kaygın vd. (2016)
X14	Asset Turnover Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Torun (2007), Alfaro et al (2008), Ekinci vd. (2008), Li and Sun (2009), Doğrul (2009), Akkaya vd. (2009), Çelik (2009), Korol and Korodi (2010), M.Y.Chen (2011), Li and Sun (2011), Elmas vd. (2011), Kılıç (2011), Galego et al (2012),
X15	Equity Turnover Ratio	Ko and Lin (2006), Torun (2007), Akkaya vd. (2009), Doğrul (2009), Çelik (2009), Divsalar et al (2011), Chen (2011), Terzi

		(2011), Elmas vd. (2011), Yakut (2012), Chuang (2013), Yakut ve Elmas (2013), Kaygın vd. (2016)
X16	Fixed Asset Turnover Ratio	Torun (2007), M.Y.Chen (2011), Li and Sun (2011)
X17	Current Assets Turnover Ratio	Torun (2007), Alfaro et al (2008), Li and Sun (2009), Doğrul (2009), Akkaya vd. (2009), Divsalar et al (2011), M.Y.Chen (2011), Elmas vd. (2011), Li and Sun (2011), Jardin (2012), Galego et al (2012), Yakut ve Elmas (2013), Lin and Liang (2014), Geng et al (2015),
X18	Tangible Fixed Assets Turnover Ratio	Ko and Lin (2006), Jianguo Chen et al (2006), Li and Sun (2009), Doğrul (2009), Divsalar et al (2011), M.Y. Chen (2011), Elmas vd. (2011), Li and Sun (2011), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Ural vd. (2015)
X19	Gross Margin Ratio	Torun (2007), Li and Sun (2009), Doğrul (2009), Çelik (2009), Akkaya vd. (2009), Elmas vd. (2011), Yakut (2012), Yakut ve Elmas (2013), Lin and Liang (2014), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Kaygın vd. (2016)
X20	Operating Margin Ratio	Torun (2007), Çelik (2009), Doğrul (2009), Li and Sun (2011), Divsalar et al (2011), M.Y.Chen (2011), Kılıç (2011), Elmas vd. (2011), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Kaygın vd. (2016)
X21	Net Profit Margin	Torun (2007), Çelik (2009), Doğrul (2009), Li and Sun (2009), Akkaya vd. (2009), Glezakos et al (2010), Divsalar et al (2011), Li and Sun (2011), Kılıç (2011), Elmas vd. (2011), Kılıç ve Seyrek (2012), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Geng et al (2015), Ural vd. (2015), Öcal ve Kadıoğlu (2015), Öcal et al (2015), Kaygın vd. (2016)
X22	Return On Total Assets (ROTA)	Jianguo Chen et al (2006), Torun (2007), Alfaro et al (2008), Li and Sun (2009), Çelik (2009), M.Y.Chen (2011), Li and Sun (2011), Divsalar et al (2011), Terzi (2011), Jardin (2012), Galego et al (2012), Kılıç ve Seyrek (2012), Lin and Liang (2014), Öcal ve Kadıoğlu (2015), Kaygın vd. (2016)
X23	Return On Assets (ROA)	Ko and Lin (2006), Torun (2007), Ekinci vd. (2008), Li and Sun (2009), Çelik (2009), Korol and Korodi (2010), Glezakos et al (2010), Divsalar et al (2011), Li and Sun (2011), Galego et al (2012), Yakut (2012), Chuang (2013), Lin and Liang (2014), Geng et al (2015), Öcal ve Kadıoğlu (2015)
X24	Return On Equity (ROE)	Ko and Lin (2006), Torun (2007), Ekinci vd. (2008), Li and Sun (2009), Çelik (2009), Akkaya vd. (2009), Doğrul (2009), Glezakos et al (2010), M.Y. Chen (2011), Li and Sun (2011), Galego et al (2012), Yakut (2012), Chuang (2013), Yakut ve Elmas (2013), Lin and Liang (2014), Ural vd. (2015), Okay (2015), Öcal ve Kadıoğlu (2015), Kaygın vd. (2016)
X25	Expense Ratio	Ko and Lin (2006), Torun (2007), Li and Sun (2009), Çelik (2009), Doğrul (2009), Divsalar et al (2011), Li and Sun (2011),

		Kılıç (2011), Kılıç ve Seyrek (2012), Yakut (2012), Öcal ve Kadıoğlu (2015), Öcal et al (2015)
X26	Company Operation Time, M.Y.Chen (2011)	
X27	Being Audited by Four Major Audit Firms (Pricewaterhousecoopers-Deloitte Touche Tohmatsu- Kpmg- Ernst And Young) It was added by the researcher upon the recommendation of the members of the thesis monitoring committee.	
X28	Free Float Rate % It was added by the researcher upon the recommendation of the members of the thesis monitoring committee.	
X29	Real and Legal Persons with a Direct or More Share of 5% in Capital - Foreign Capital Share in Non-Public Shares % It was added by the researcher upon the recommendation of the members of the thesis monitoring committee.	