Determinants of Credit Ratings and Comparison of the Rating Prediction Performances of Machine Learning Algorithms

Ali İhsan Çetin[1,*] and *Syed Ejaz* Ahmed[2]

[1]Statistics Dep., Faculty of Science, Yildiz Technical University, 34220 Esenler, İstanbul, Turkey
[2]Faculty of Math & Science, Brock University, 1812 Sir Isaac Brock Way, ON L2S 3A1, Canada

Abstract. In the literature, new machine learning algorithms are dynamically produced in the field of artificial intelligence engineering and the algorithms are constantly updated with new parameter estimations. The performance of existing algorithms in various business areas is still an important topic of discussion. Also, machine learning algorithms are frequently used in long-term credit ratings, which is an crucially important sub-branch of finance. This study was conducted to determine which popular machine learning model performs better in credit scoring. Artificial Neural Network, Random Forest, Support Vector Machine and K Nearest Neighbor were used to determine the algorithm that is suitable for the structure, attribute content and distribution of the data, and the operating logic of the models. In the study, the long-term credit rating is the target variable and the remaining variables are the features, the prediction performances of these 4 algorithm, which are frequently used in previous studies such as credit rating, credit risk, fraud analysis were compared. After data preprocessing, a classification study was carried out using the features included in the model. The metrics used in the comparison are MSE, RMSE, MAE and accuracy. According to the metrics, RF algorithm showed the best performance in the credit scoring. **Keywords:** Machine Learning, Classification, Finance, Data mining.

1 Introduction

Credit rating plays a key role in meeting the needs of both countries, institutions and investors, such as the development of policies that will help the development and deepening of international markets, the elimination of information asymmetry, the strengthening of financial structures of institutions and organizations, and the standardization of risk determination.

The concept of credit rating has started to be pronounced a lot since the last century. The fact that an economic activity is in question has reached such a wide range of influence that it can mean that credit ratings have an effect on those activities in today's conditions. Credit ratings can also act as a guide for the future. Credit rating activity is considered as a decisive process in determining the economic situation of countries or businesses. The scope of the credit rating activity, which did not have such an effect when it first emerged, has expanded with the changing conditions around the world and the acceleration of globalization. The

*e-mail: aliihsancetin22@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

three major companies that have proven their existence in the credit rating processes and continue to prevail are Standard and Poor's, Moody's and Fitch companies.

1.1 Credit Rating Concept

Growing international capital markets, increasing cross-border activities and accepted Basel Standards have made the concept of credit rating vital today. Although there are various definitions of the concept of rating, some of them are presented below:

According to Akkaya and Demireli (2010), rating based on past and present qualitative and quantitative data is a classification system that use to predict whether an organization fulfill its financial obligations on time [1].

According to Standard and Poor's, one of the international rating companies, express rating as: "The borrower's opinion of the creditworthiness of a particular financial obligation or financial program [2]."

Rating is a tool that not only facilitates the activities of the markets and beneficiary investors, but also transforms the results obtained as a result of costly and time-consuming studies into understandable symbols for the markets [3].

In accordance with the Communiqué of the Capital Markets Board, the credit rating is an indicator that shows the risk status and solvency of the enterprises, or it is an independent, impartial and fair evaluation activity by the rating agencies in order to determine the ability to pay the principal, interest and similar liabilities of the capital market instruments representing the indebtedness at the end of maturity [4].

1.2 Scope and Purposes of Credit Rating

The rating process is applied to reveal the borrowing power of countries in international markets, as well as to evaluate the liabilities of national and international securities issuances arising from certain borrowings of commercial companies, financial institutions and banks and is a subjective process. Credit ratings are forward-looking as they assess the possible effects of foreseeable future events with the help of current and historical information [5].

Credit rating is a very broad process and the determinants in its scope are interrelated. The knowledge of how and according to which criteria the country rating is made plays a complementary role in the institution and securities rating process. Another important issue is what the objectives of the rating activity are in order to reveal the importance of the rating. At the same time, these objectives must be of a nature to support the reliability, integrity and dynamism of the markets.

As a result of the rating activity, it provides concentrated information about the businesses for investors who care about time constraints, enabling them to reach them as soon as possible [6]. In another matter; Since the rating process is not a one-off, the grade given is periodically verified or revised according to economic and other developments [7]. Even the rating agencies have the right to include the companies they have risk rated on the 'Risk Watch List' due to incomplete information.

In measurement of creditworthiness, comparative analysis is used to prevent confusion in the sector by ensuring homogeneity and to perform the rating in a complete and realistic way. At the same time, the reliability of the ratings given by the rating agencies is ensured both by the methods they use and by their experience, knowledge and impartial analysis.

Credit rating eliminates information asymmetry and also provides important data between financial markets and decision makers [8]. However, even if the rating score affects the decisions and preferences of investors, market prices and marketability, it is not a recommendation to make or not to invest in the relevant field, to lend or not to lend/loan to the relevant person and/or institution [9].

1.3 Types of Credit Rating

In order to better understand the credit rating activities, the rating service is divided into different categories according to its scope, maturity, account units and application areas as seen table 1.

Table 1. Types of credit rating [10]

In Terms of Scope	In Terms of Maturities	In terms of Account Units	In Terms of Groups Rating Service Provided
Rating of Issuer (Debtor)	Rating of LT Liabilities	Currency Rating	Country Rating
Rating of Issuer (Debt)	Rating of ST Liabilities	Local Currency Rating	Institutional Rating
			Securities Rating
			Project Evaluation

When the rating activity is examined according to their scope; the debtor rating (issuer rating), which includes the country (sovereign rating) and corporate rating (institutional rating), and the debt rating (dept / issue rating), which is the rating of the obligation belonging to the country or institution, can be collected in two headings [11]. The debtor rating is the rating of a certain type of financial instrument of the issuer, while the debt rating is expressed as the rating of the issuer's ability to pay back the interest and principal arising from the debt [12]. Areas of application of the rating based on this description, it appears by being considered together with the debt and borrower rating classified according to their scope.

The maturity lengths of credibility follow the financial literature. It consists of two groups as short-term and long-term credit ratings. Short-term ratings are an assessment of the capacity of the country, bank, or commercial company to meet its obligations for up to one year, while long-term ratings refer to its ability to meet obligations with a maturity of more than one year. With this feature, it imitates the maturities of the income statement and balance sheet. The liquidity of the institution is important in the short-term rating. In the long-term rating, principles such as the financial situation in the sector, technological developments of the institution, changes in demand, legal regulations and management quality are taken into account. According to account units, a distinction is made between foreign currency credit ratings and local currency credit ratings. All country risks are included in the foreign currency note and the ability of the country or company to fulfill its foreign exchange obligations by creating foreign currency is evaluated. Considering the transfer risk and convertibility risk while making the rating shows the possibility of closing the foreign exchange market in the country by limiting the foreign exchange outflow of the state administration. In a local currency note, the country assesses the ability of the bank or company to pay its obligations in local currency by generating the local currency.

1.4 Country Rating

Country credit ratings are an indicator of the capacity of countries to fulfill their financial obligations, as well as their credit risk. It is expected that the country's ability to access more cheap and long-term foreign financing will increase with the increase in its credit rating. If the credit rating is at the "investment grade" level, it is an important threshold for countries' access to external finance [13]. Thus, the markets of developed countries accelerate and enable these countries to develop in a shorter time. On the other hand, it plays an important role in determining the credit worthiness in developing countries such as our country.

Country ratings are performed by Moody's, S&P and Fitch in three stages: [14]

- Evaluation of the economic situation: At least two analysts, seeking the opinions of key government officials, private sector analysts, journalists, universities and representatives of the opposition, prepare a report including macroeconomic data.
- Digitizing qualitative and quantitative factors with a point system: Since it is aimed to give similar grades among countries in similar situations, country scores are calculated by representatives from different regions, private sector analysts and relevant country experts in accordance with various criteria, points on the vertical axis and points on the horizontal axis. Based on the rating grades, graphs showing the countries are created.
- Thus, countries can be seen together and compared.

Final evaluation by the committee: Credit ratings are determined by evaluating the data obtained in the previous stages by the rating committee.

1.5 Credit Rating Agencies

After discussing the formation of the legal infrastructure related to credit rating and its harmonization according to the changing and developing markets over the years, the credit rating agencies and the letter grades they give are emphasized.

Rating agencies, which provide information to the market by measuring the risk of meeting the principal, interest and similar liabilities of debt instruments at maturity, aim to provide consistent and comparable assessments to national or international capital markets regarding the credit quality of companies, banks and countries.

Rating agencies are professional organizations that provide impartial (objective) service regardless of government patronage or any bank or similar institution. Established only in Tokyo, Japan Credit Rating (JCR) has a semi-public character with the Japanese Ministry of Finance being a partner in the company.

Rating agencies generally conduct their transactions on demand and analyze and evaluate other information, both public and non-public. It enables investors to make comparisons between various debt instruments by facilitating time-consuming and costly analyzes that they cannot do with the securities they issue. In addition, rating agencies reduce the problem of asymmetric information between the country bank or company and the investor. Thus, the investor can be aware of the risk. Based on the willingness of investors to accept the opinions expressed, these opinions are recognized and valuable in the market.

It is important for investors to trust the institutions and ratings to accept these views. The methods used by the organizations in the rating process, as well as the experience and knowledge, ensure the reliability of the rating grades given in the evaluation process [15].

On the other hand, the ratings given to companies, debt instruments, banks and countries by credit rating agencies not only help determine the company and country to invest in, but also determine the interest amount, which is the cost of debt, for companies and countries that want to borrow. A “speculative” or “default” rating of companies, debt instruments, banks or countries means that investment in them will be low because of the high cost of borrowing or the expectation that the investment is too risky.

1.6 Fitch Ratings Agency

Fitch, started its activities in 1913 with the “Fitch Publishing Company”, which was founded by John Knowles Fitch. It was established in New York and is the first European-based rating agency. Fitch, which provides different services in addition to credit rating activities, maintains a wide data set on fixed income securities with Fitch Solution, which markets its ratings, and offers services such as macroeconomic analysis, investment consultancy and risk

analysis. In addition, training on corporate finance and credit ratings is given in the Fitch Training section [16]. Fitch has analyzed credit ratings in two periods, long-term and short-term. Long-term investment grade ratings are given in table 2.

Table 2. S&P, Moody's and Fitch ratings

S&P	Fitch	Moody's	Description of the Note	Category	S&P	Fitch	Moody's	Description of the Note	Category
AAA	AAA	Aaa	Prime (Highest credit quality)	Investable level	BB+	BB+	Ba1	Speculative (Speculative)	The level that cannot be invested
AA+	AA+	Aa1	High note (Very high credit quality)		BB	BB	Ba2		
AA	AA	Aa2			BB-	BB-	Ba3		
AA-	AA-	Aa3			B+	B+	B1	Very Speculative (Quite Speculative)	
A+	A+	A1	Upper Middle Note (High credit quality)		B	B	B2		
A	A	A2			B-	B-	B3		
A-	A-	A3			CCC+	CCC	Caa	Extremely Risky (High Risk of Default)	
					CCC	CC	Caa3		
BBB+	BBB+	Baa1	Lower Middle Note (Good credit quality)		CC	C	Ca		
BBB	BBB	Baa2				DDD		Bad Loans (Default)	
BBB-	BBB-	Baa3			D	DD	D		

It is the common opinion of the three rating agencies that the credit rating given in line with the examinations and analyzes is only an opinion or judgment and not an investment advice. When determining the rating, the grade of at least one of the two large organizations is sufficient.

2 Methodology

2.1 Purpose of the Study

In the literature, machine learning models are used in many branches such as risk management, fraud analysis, anomaly detection, credit scoring, and pattern recognition in the field of finance. For example, in the study conducted in the field of Fake Online Reviews, popular machine learning methods were used. The purpose of this study is to propose a new way to predict fake reviews by implementing supervised machine learning models. To achieve this, a dataset consisting of 1,600 reviews was used, and features were extracted from this dataset. A variety of supervised learning classifiers were trained by using features from the dataset, and then tested. The performance of each prediction model was compared using certain metrics, and the best result was acquired using the Random Forest (RF) Classifier [17]. Other example, Bahnsen et al. conducted a fraud analysis study on credit card fraud. In this study, using a real credit card fraud dataset provided by a large European card processing company, state-of-the-art credit card fraud detection models were compared, and evaluated how the different sets of features have an impact on the results. By including the proposed periodic features into the methods, the results showed an average increase in savings of 13% [18]. In other study, an analysis of South Africa's (SA) sovereign credit rating (SCR) using Naïve Bayes, a Machine learning (ML) technique was conducted. Quarterly data from 1999 to 2018 of macroeconomic variables and categorical SCRs were analyzed and classified to predict and compare variables used in assigning SCRs. The findings showed that CRAs use different macroeconomic variables to assess and assign sovereign ratings. Household debt to disposable income, exchange rates and inflation were the most important variables for estimating and classifying ratings [19]. In other study, an approach has been developed to predict corporate credit ratings by analyzing public opinion of companies on social media to help financial institutions effectively assess and control corporate risk. The experimental results of this research showed that the accuracy of corporate credit rating prediction based

on social media big data is higher than that of traditional financial report, corporate governance and macroeconomic indicators. Moreover, the adopted forecasting model, K-Nearest Neighbor (KNN), is superior to the other machine learning models in terms of accuracy [20]. An another study attempted to identify the long-term and short-term financial determinants of credit ratings issued to Indian companies and examines their impact on credit ratings on the basis of panel data and cross-sectional approaches. Results indicated that size, profitability and leverage have a significant relationship with corporate credit ratings in both panel data and cross-sectional approaches. Also, size has the highest impact on credit ratings followed by leverage and profitability in both the approaches [21]. A sovereign credit rating is a measurement of a sovereign government's ability to meet its financial debt obligations. The differences by Credit Rating Agencies on rating grades on similar firms and sovereigns have raised questions on which elements truly determine credit ratings. Also, the importance of machine learning models is increasing day by day in the world of dynamic algorithms, where new models are constantly produced and prediction parameters are updated. There are discussions about which of the machine learning algorithms developed and updated in the field of artificial intelligence engineering, which is one of the sub-branches of engineering, gives more meaningful results and the metrics work better. In this article, first of all, the results of a successful classification study were examined and the variables that were effective in achieving this success were determined. Afterwards, a cross-model study was carried out with the most obvious factors affecting the credit rating to contribute to the discussion. It was revealed which of the popular algorithms had a more successful classification on the sovereign credit rating.

The long-term credit ratings of the countries rated by Fitch Ratings for the years 2012-2020 and some macroeconomic and financial variables are used in this article. In which long-term credit rating is the dependent variable and other variables are independent variables, metric performances were measured using Random Forests, Artificial Neural Networks, Support Vector Machines and K Nearest Neighbor algorithms.

In the study, primarily taken from Fitch Ratings; 63 variables of 134 countries graded between 2012 and 2020 were taken from the databases of the World Bank and the United Nations and used. The "Continent" variable, which was evaluated by the expert opinion, which would have an effect on the study, was added later. Analysis was carried out with a total of 1075 data. The basis of this selection process is the cyclical and financial situation, which are also the cornerstones of long-term credit rating, the technological developments of the country, changes in demand, legal regulations and management quality.

The methodology of the study is structured as follows: the target variable in the study consists of 7 categories. While creating these 7 categories, a study in the literature was used [22]. The last 2 categories, which represent the worst grade, have been combined, waiving model sensitivity. The purpose of this is to increase the prediction success. Then, since scaling of the ANN, SVM and KNN algorithms used in the model will increase the prediction performance of these models, scaling has been done on the features first. In the second step, 29 features with null values above 70% were detected and excluded from the study, as there were many features with missing data. The missing data of the remaining features were imputed by the KNN imputation method, which has been proven successful in the literature [23]. With the Random Forest Feature Importance method, 30 features with an effect of more than 0.01% on the target variable were selected, and correlation analysis was performed to determine the relationships above 0.5 and these features were eliminated from the study. A test-train split of 0.8-0.2 was performed on the 13 final variables that were ready for modelling. Finally, modeling studies were completed with hyperparameter optimization. The independent variables used in the model are given in the table 3 below.

Table 3. Features

	Indicator Name		Indicator Name
1	Current account balance (% of GDP)	33	General government final consumption expenditure (% of GDP)
2	Use of IMF credit (DOD, current US$)	34	General government final consumption expenditure (annual % growth)
3	Total debt service (% of GNI)	35	Air transport, passengers carried
4	Short-term debt (% of total reserves)	36	CPIA macroeconomic management rating (1=low to 6=high)
5	Real interest rate (%)	37	CPIA property rights and rule-based governance rating (1=low to 6=high)
6	Interest payments (% of expense)	38	CPIA quality of public administration rating (1=low to 6=high)
7	Lending interest rate (%)	39	Deposit interest rate (%)
8	Urban population (% of total population)	40	Firms that spend on R&D (% of firms)
9	Urban population growth (annual %)	41	GDP (constant 2010 US$)
10	Population growth (annual %)	42	High-technology exports (% of manufactured exports)
11	Population, total	43	Gross value added at factor cost (constant 2010 US$)
12	Net financial account (BoP, current US$)	44	Human capital index (HCI) (scale 0-1)
13	CPIA financial sector rating (1=low to 6=high)	45	Imports of goods and services (% of GDP)
14	Inflation, consumer prices (annual %)	46	Income share held by highest 10%
15	Gross savings (% of GDP)	47	Manufacturing, value added (constant 2010 US$)
16	Gross capital formation (% of GDP)	48	Manufacturing, value added (annual % growth)
17	GDP per capita growth (annual %)	49	Natural gas rents (% of GDP)
18	GDP per capita (current US$)	50	Net lending (+) / net borrowing (-) (% of GDP)
19	Foreign direct investment, net inflows (% of GDP)	51	Oil rents (% of GDP)
20	Final consumption expenditure (% of GDP)	52	Other taxes (% of revenue)
21	Final consumption expenditure (annual % growth)	53	Profit tax (% of commercial profits)
22	External debt stocks (% of GNI)	54	Rail lines (total route-km)
23	Domestic credit to private sector (% of GDP)	55	Renewable energy consumption (% of total final energy consumption)
24	Gross domestic savings (% of GDP)	56	Revenue, excluding grants (% of GDP)
25	Death rate, crude (per 1,000 people)	57	Researchers in R&D (per million people)
26	Cause of death, by injury (% of total)	58	Services, value added (% of GDP)
27	Consumer price index (2010 = 100)	59	Services, value added (annual % growth)
28	Central government debt, total (% of GDP)	60	Tax revenue (% of GDP)
29	Bank capital to assets ratio (%)	61	Trade (% of GDP)
30	Adjusted net national income (annual % growth)	62	Unemployment, total (% of total labor force) (modeled ILO estimate)
31	Adjusted net national income (constant 2010 US$)	63	Human Development Index (HDI)
32	Bank nonperforming loans to total gross loans (%)	64	Continent

2.2 Machine Learning Models

In this section, machine learning techniques used in the study are discussed. In addition, reference studies on which other financial studies have successfully used these 4 popular machine learning algorithms are also pointed out.

(1)Random forest

A Random Forest (RF) is an ensemble of decision trees [24], i.e. K decision trees are built on bootstrapped samples with m observations. Each decision tree is developed using a subset of randomly chosen k features. Each decision tree will give a class of a new feature vector. Thereafter, for overall classification, a RF assigns the class of the new feature vector by using majority vote based on the outputs from the decision trees.

West created a credit scoring model using five artificial neural network techniques (multi-layer perceptron, expert systems, radial basis function, learning vector quantization, fuzzy adaptive resonance), linear discriminant analysis, logistic regression analysis, k nearest neighbor, kernel density estimation and decision tree techniques [25].

Ha Van Sang et al. proposed another known credit risk model in their study, a model based on the random forest (RF) parallel credit scoring model. They achieved average accuracy of 76.2% and 89.4% by testing two real data, and the model significantly reduced run time by using the optimum mean median and minimum standard deviation as variable scoring rules [26].

(2)Artificial neural networks

An Artificial Neural Network (ANN) [27] is a system which is motivated by a biological neural network system. An ANN emulates the way in which a biological neural network

of the brain processes information by means of interconnected neurons [28, 29]. Typically, a neural network consists of three layers, namely an input, hidden and output layers [30]. Essentially, training a neural network involves the process of finding optimal weights that map the input and output layers by means of back-propagation.

For a given input feature vector x, a three-layer ANN computes the output $\hat{y}$ according to

$$\hat{y} = a_2\left(a_1\left(a^{(1)}x + a_0^{(1)}\right)a^{(2)} + a_0^{(2)}\right). \tag{1}$$

where $\left(a_0^{(1)}, a^{(1)}\right)$, $\left(a_0^{(2)}, a^{(2)}\right)$ are a_1, a_2 weights and are activation functions between input and hidden layer, hidden layer and output layer, respectively. The parameters are learned on a training set. The ANN performs final decision by applying a decision function, such as soft-max, on $\hat{y}$. The ANN was first applied in credit scoring by Odom and Sharda [31].

Bektaș and Gökçen conducted an estimation study on twelve banks operating in the Turkish banking sector and awarded a financial strength rating by Moody's between 2007 and 2010 by using feedforward artificial neural networks, self-organizing maps technique and support vector machines technique. They concluded that the neural network model is superior to other techniques [32].

Chaveesuk et al. compared the classifi cation performances of back propagation neural networks, radial basis function, learning vector quantification and logistic regression analysis. It was determined that back propagation neural networks and logistic regression analysis showed higher performance than other methods with 51.9% and 53.3% success, respectively [33].

(3)Support vector machines

A Support Vector Machine (SVM) [34] uses an idea of a hyperplane (which is a decision boundary) that separates classes in a high dimensional feature space. The linear SVM focuses on maximizing the margin $\| \alpha \| = \sqrt{\sum_{i=1}^{m} a_i^2}$ between the negative and positive hyperplanes. The correct class is assigned by using the following equation:

$$y = \begin{cases} +1, if\ b + a^T x \geq +1 \\ -1, if\ b + a^T x \leq -1. \end{cases} \tag{2}$$

where b is the bias. For non-linear cases, b kernel trick [35] is used to project features into a high dimensional space.

Huang et al. in their study, two data sets consisting of data from financial institutions in Taiwan and commercial banks in the USA were prepared and credit rating predictions were made using back propagation neural networks and support vector machine techniques from artificial neural network systems. It is concluded that the performance of support vector machines is high [36].

(4)K-nearest neighbor

A k-Nearest Neighbor (k-NN) [37] assigns to an input feature vector x the class of the majority of its x nearest neighbors in the training dataset. The nearest neighbors are determined by calculating the Euclidean distance or Mahalanobis distance between the input feature vector x and the training dataset $\{x_k\}_{k=1}^m$. Thus the class for the new data point is

$$y = \begin{array}{c} majority \\ y \in \{+1, -1\} \end{array} \left[\begin{array}{c} \arg\min \left\| \{x_k\}_{k=1}^m - x \right\| \\ k \end{array} \right]. \tag{3}$$

A successful KNN algorithm example has been included in the literature as a pattern recognition study on railway outlines [38].

3 Findings and Future Studies

Feature Importance values of the features made ready for modeling according to the target variable are given in figure 1 below.

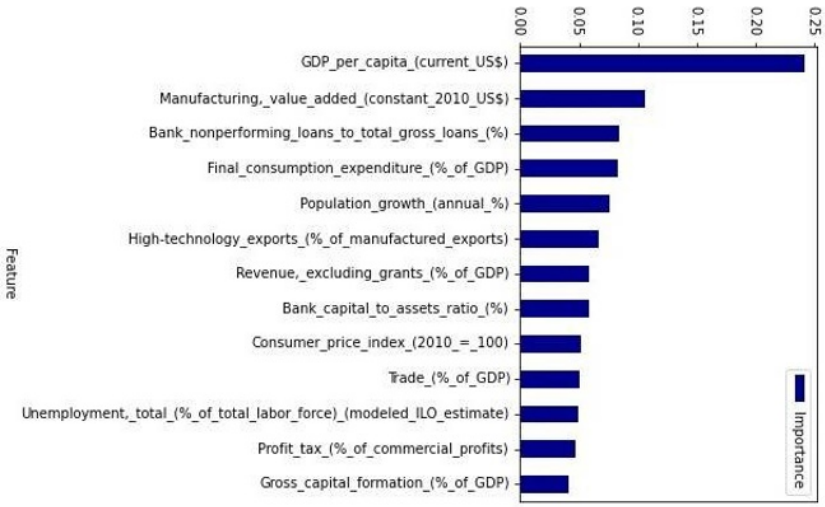


Figure 1. Feature importance

According to the figure 1, the variables with the highest impact on a country's credit rating are GDP per capita, Manufacturing value added, Bank nonperforming loans to total gross loans, Final consumption expenditure, Population growth, High-technology exports, Revenue excluding grants, Bank capital to assets ratio, Consumer price index, Trade, Unemployment, Profit tax and Gross capital formation. GDP per capita has a high degree of importance compared to other variables.

Table 4 shows the performance metrics resulting from the hyperparameter optimization made with the GridsearchCV method.

Table 4. Metric performance results

	Random Forest	**Artificial Neural Network**	**Support Vector Machines**	**K Nearest Neighbor**
Mean Absolute Error	**0,15**	0,21	0,23	0,26
Mean Squared Error	**0,23**	0,33	0,39	0,39
Root Mean Squared Error	**0,48**	0,57	0,62	0,62
Accuracy	**0,87**	0,82	0,83	0,78

According to the MAE, MSE, RMSE and accuracy results of the models in table 4, it is seen that the Random Forest algorithm has the highest classification success with 87% of the countries' credit ratings. Likewise, the model with the lowest error compared to other metrics was Random Forest. Therefore, when the performance of the most popular machine learning models on credit scoring data is measured, it can be said that the best performing algorithm is Random Forest.

There are over 200 countries in the world. The aim of these countries is to survive by keeping their economic situation as strong as possible. Countries that want to hold the economic power produce and trade goods and services, sell them both inside and outside the

country, and also engage in import activities. Production is the most important factor in ensuring the economic freedom of a country. Investment is required for the realization of production activities. Countries aim to attract both internal and external investors and make an economic contribution to them. On the other hand, investors prefer to invest in countries with low costs, high profits and less procedures. There are also ideological reasons. The need to evaluate the financial adequacy of the countries within the framework of general criteria emerges and this opens the way for credit rating. Such assessments are made by various credit rating agencies described above and can be a reference to investors. In this respect, this study will shed light on the literature on which machine learning model will be used for credit scoring studies in finance.

Estimated values and actual values are compared in the article. As a result, the Random Forest algorithm classified 187 of 214 real outputs (dependent variables) correctly and the success rate was 87%. The reason for the success of Random Forest is that it explains the extent to which the variables affect the target variable, it chooses the most votes systematically by creating many trees, it does not need standardization and it is superior in terms of being able to work with original data.

In addition, the two variables that draw attention in the variable used in the study and have a high impact on the determination of the credit rating are GDP per capita and Manufacturing value added features. Accordingly, these financial and macroeconomic variables constitute an important reference when estimating a country's credit rating for a certain year.

The reason for the low performance of the Artificial Neural Networks algorithm in the study is that it generally creates a learning neural network that establishes inter-layer relations with big data and imitates the human brain. This study has data limitations. In addition, there is no feature structure suitable for distance-based KNN and SVM algorithms in this study. For this reason, it can be said that the working logic of Random Forest algorithms in scoring studies in the field of finance is more suitable for the data structures in this field. Thus, KNN, SVM, ANN and RF algorithms, which are popular algorithms in the field of artificial intelligence engineering and data mining, which strengthen their place in the field of software every day, are compared in the application of financial data and the results are interpreted.

Another perspective, in terms of shedding light on future studies, is to include features with different structures in the study and performance comparisons can be made. For example, variables with a high degree of importance in the qualitative structure will positively affect the performance of other algorithms. In addition, a target variable with 7 categories was used in this study. By increasing the sensitivity of this variable, a new methodology can be constructed from the beginning. Finally, since the credit score variable is a variable related to the grades taken by countries from year to year, this study can be compared with the performance metrics of other models by using time series analysis, which is one of the classical statistical methods.

I would like to thank Professor Syed Ejaz Ahmed for his support and guidance in the emergence, development and conclusion of this study idea. Words cannot express my gratitude for her invaluable patience and feedback. I could not have undertaken this journey without her generous provided knowledge and expertise. Additionally, this endeavor would not have been possible without the generous support from the professor, who provide me to participate in this conference.

I would be remiss in not mentioning my family, especially my wife. His belief in me has kept my spirits and motivation high during this process. I would also like to thank my son for all the entertainment and emotional support.

References

[1] G.C. Akkaya, E. Demireli, Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi **19**, 319 (2010)
[2] P. Standart, *Corporate ratings criteria* (1988)
[3] P. Çelik, Central Bank of the Republic of Turkey **4**, 1 (2004)
[4] C.M.B. of Türkiye, *Communiqué on principles regarding rating activities and rating agencies in the turkish capital markets* (2014)
[5] S.. Poor's, *Kredi derecelendirmelerini tanima rehberi, pdf* (2012)
[6] F. Ratings, *Rating definitions* (2012), *http* : *//www.fitchratings.com/corporate/fitchresources.cfm?detail*
[7] S. Melek, Yayımlanmamış Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul (2011)
[8] H.H. An, K.C. Chan, Journal of Corporate Finance **14**, 584 (2008)
[9] Y. Akbulak, Mali Çözüm Dergisi (2012)
[10] M. Demir, *Kredi derecelendirme kuruluşları derecelendirme süreçleri ve türkiye* (2014)
[11] N.S. Halici, *Kredi derecelendirme şirketleri, kredi derecelendirmenin belirleyicileri ve etkileri* (2005)
[12] M.A. Boyacıoğlu, *Bankalarda derecelendirme (rating) ve türk bankacılık sektörü üzerine ampirik bir çalışma,(selçuk üniversitesi: Doktora tezi 2005: Konya)* (2005)
[13] İ.B. Kanlı, Y. Barlas, Türkiye Cumhuriyet Merkez Bankası, Çalışma Tebliği (2011)
[14] L.A. Mukatel, *Ülke derecelendirme notlarının borsa endekslerine etkisi: İmkb 100 endeksi üzerinde bir model denemesi* (2006)
[15] P. Deb, G. Murphy, London School of Economics working paper pp. 1–27 (2009)
[16] F. Ratings, *Fitch ratings about us* (2012), *http* : *//www.fitchratings.com/web/en/dynamic/about − us/sector − coverage.jsp*
[17] B.D. Kurtcan, T. Kaya, *Classification of Authentic and Fake Online Reviews with Supervised Machine Learning Techniques* (2022)
[18] A.C. Bahnsen, D. Aouada, A. Stojanovic, B. Ottersten, Expert Systems with Applications **51**, 134 (2016)
[19] O. Takawira, W.M. Mwamba, Eurasian Journal of Economics and Finance **8**, 279 (2020)
[20] Y.J. Chen, Y.M. Chen, Expert Systems with Applications **207**, 118042 (2022)
[21] R. Gupta, International Journal of Finance & Economics (2021)
[22] E.G. Genc, O.D. Basar, Business and Economics Research Journal **10**, 391 (2019)
[23] S.E. Ahmed, D. Aydin, E. Yılmaz, *Nonparametric regression estimates based on imputation techniques for right-censored data* (2020)
[24] L. Breiman, Machine learning **45**, 5 (2001)
[25] D. West, Computers & operations research **27**, 1131 (2000)
[26] H. Van Sang, N.H. Nam, N.D. Nhan, Indian Journal of Science and Technology **9**, 1 (2016)
[27] T.M. Mitchell, M. Learning, Engineering/Math **1**, 27 (1997)
[28] F. Barboza, H. Kimura, E. Altman, Expert Systems with Applications **83**, 405 (2017)
[29] C.F. Tsai, Y.F. Hsu, D.C. Yen, Applied Soft Computing **24**, 977 (2014)
[30] C.F. Tsai, J.W. Wu, Expert systems with applications **34**, 2639 (2008)
[31] M.D. Odom, R. Sharda, *A neural network model for bankruptcy prediction* (1990)
[32] H. Bektaş, A. Gökçen, Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi **31**, 345 (2011)

[33] R. Chaveesuk, C. Srivaree-Ratana, A.E. Smith, Journal of Engineering Valuation and Cost Analysis **2**, 117 (1999)
[34] C. Cortes, V. Vapnik, Machine learning **20**, 273 (1995)
[35] B. Schölkopf, Advances in neural information processing systems **13** (2000)
[36] Z. Huang, H. Chen, C.J. Hsu, W.H. Chen, S. Wu, Decision support systems **37**, 543 (2004)
[37] W. Henley, D.J. Hand, Journal of the Royal Statistical Society: Series D (The Statistician) **45**, 77 (1996)
[38] A.M. del Rio, I.S. Ramirez, F.P.G. Marquez, *Machine Learning Techniques for Pattern Recognition in Railway Switches: A Real Case Study* (2022)