

Predicting Company Credit Rating Using Artificial Intelligence Techniques from Publicly Available Financial Data^{*}

Juozas Širmenis^{1,*}, Mindaugas Kavaliauskas^{2,3,†} and Ingrida Lagzdinytė-Budnikė^{1,‡}

¹Faculty of Informatics, Kaunas University of Technology, Kaunas, Lithuania

²RiskPlanner, UAB Idėjų valda, Kaunas, Lithuania

³Faculty of Mathematics and Natural Sciences, Kaunas University of Technology, Kaunas, Lithuania

Abstract

The study focuses on predicting credit rating using statistical methods (Linear and Huber Regressions) and machine learning techniques (Artificial Neural Network and Random Forest) while using publicly available financial data with additionally calculated features. The results show that machine learning techniques outperformed statistical methods significantly. The best results were obtained using the ANN model: MSE reached 0.063, MAE – 0.1858, R² – 0.9065, and RMSE – 0.251. The notable performance improvement across all models was noticed when incorporating additionally derived financial ratios, notwithstanding their derivation from metrics already included in the analysis.

Keywords

Credit Rating, Linear Regression, Huber Regression, Artificial Neural Network, Random Forest, Financial Statements

Introduction

In today's world, where countries' borders are less of a barrier to international collaboration and diseases and military conflicts pose a threat to people and the environment, predicting a partner's financial behavior is critical. Some researchers work on projecting business defaults [1-3], while others focus on credit rating and scoring [4-6]. Credit ratings reflect how likely someone is to meet their financial responsibilities, however they are based on opinion rather than fact [7].

Based on research, data analysis techniques for evaluating, comparing, or predicting credit risk can be categorized into two groups: statistical methods and machine learning (ML) methodologies. According to some papers, combining methods and algorithms could lead to better results [5, 8-10].

Two of the most popular statistical methods are Logistic Regression (LG) and Discriminant Analyses (DA). In some papers, they are used to predict bankruptcy [1] or defaults [11] of the corporates. The others, forecasts the defaults of small and medium enterprise (SME) [2, 3, 12] or uses to create credit rating and scoring models [5, 10, 13-15]. Also, LG and DA was used to predict bond ratings [16] or evaluate credit risk in general [8].

The most popular ML methods for evaluating credit risk, predicting defaults or bankruptcy, forecasting ratings and scores are: Artificial Neural Network (ANN) [17, 18], Support Vector Machine (SVM) [15, 16], Decision Trees (DT) [14, 16], Genetic Algorithm [19], Random Forest (RF) [11, 13, 15], Bayesian techniques [18], Gradient Boosting techniques [19], Multilayer Perceptron [20], new approach of ANN - Convolutional Neural Networks (CNNs) [20, 21].

It is also worth noting the custom techniques and architectures that have been developed for solving similar credit risk problems: hybrid best-worst method (BWM) [22], combination of the deep neural network and decision tree classifier [23], model made from particle swarm optimization, random tree and Naïve Bayes techniques [24], the combination of decision trees and logistic

^{*} IVUS2024: Information Society and University Studies 2024, May 17, Kaunas, Lithuania

^{1,*} Corresponding author

[†] These author contributed equally.

✉ juozas.sirmenis@gmail.com (J. Širmenis); m.kavaliauskas@ktu.lt (M. Kavaliauskas); ingrida.lagzdinyte@ktu.lt (I. Lagzdinytė)

ORCID 0000-0002-4804-7102 (I. Lagzdinytė)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

regression - penalized logistic tree regression (PLTR) [25], the variables selection, regressor, and ordered probit model [26].

Two statistical methods were chosen for this study: Linear Regression and Huber regression. The first due to its simplicity and popularity for similar problems, such as modelling the dependence of bank ratings [27] or predicting companies' credit risk ratings [28]. And the second one, because of its improvement in terms of finding outliers [29]. Based on the literature analysis, machine learning methods were chosen: ANN and Random Forest, because they are among the most popular and promising models in this field.

For all models, data is crucial. The new regulation of European data aims to open more company data to the public [30] and it hopes that widely available free data will act as key element to developing the AI models [31]. We are going to use free and publicly available financial information of Lithuanian companies, although it is not complete financial statements, only the essentials are provided.

This research project aims to perform a comprehensive analysis of the performance of different algorithms and techniques in credit rating prediction but using only publicly available and free-of-charge financial data on Lithuanian companies. This task is complicated by the fact that the amount of this type of data is highly limited and may be restricted to a few financial ratios per company. To determine the relative performance of traditional statistical methodologies and state-of-the-art machine learning algorithms in this type of dataset scenario, a comparative analysis is proposed. In addition, a new approach has been proposed: a combination of the classification and prediction model.

Materials and Methods

Financial data for this paper were obtained from Registrų Centras, an official Lithuanian publicly available data source (<https://www.registrucentras.lt/p/1094>). The credit rating of the corporations was determined using the credit risk management tool called RiskPlanner (<https://www.riskplanner.io/>). The two statistical methods and two machine learning algorithms were selected to be examined: Linear Regression (LR), Huber Regression (HR), Artificial Neural Network (ANN), and Random Forest (RF).

Dataset

The dataset consists 7 features from Financial Statement (FS) data: Sales (ISLT00001), Net Profit (-Loss) (ISLT00019), Profit/Loss Before Tax (ISLT00017), Short Term Assets (BSLT00021), Long Term Assets (BSLT00001), Amounts Payable and Liabilities (BSLT00055), and Net Worth (BSLT00040). The data is from the 2017-2022, with total of 8395 records (see Table 1), respectively, where the year mostly refers to the period of this FS year from January 1st to December 31st.

Table 1

Record quantity by Year

Year	2017	2018	2019	2020	2021	2022
Quantity	284	755	1884	2261	2240	971

Each record includes the credit rating calculated by RiskPlanner from company financials, along with register number and statement year. The rating value ranges from 1 to 5, with classes ranging from A to E, where A represents the best rating and E the worst. The difference in values between classes A/B, as well as D/E, is 0.5, while for the other classes, it is 1.

Data Preprocessing

In addition to publicly available financial features, the extra ratios were calculated and added to the dataset: Altman Z score [32], Current Ratio and Net Profit Margin [33], Return on capital employed (ROCE) [34], and Accounts Payable Turnovers in days [35]. In instances where the value required for the original formula was unavailable, it was identified as non-existent or the other value, which in

financial logic may be similar, was used instead. Furthermore, a calculation was performed to figure out the value of total assets (BSLT00039) by adding long-term and short-term assets.

The Altman Z score is calculated by multiplying X coefficients by weights (1):

$$X_1 = \frac{BSLT\ 00021 - BSLT\ 00055}{BSLT\ 00039}; X_2 = \frac{BSLT\ 00040}{BSLT\ 00039}; X_4 = \frac{BSLT\ 00040}{BSLT\ 00055}; X_5 = \frac{ISLT\ 00001}{BSLT\ 00039};$$

$$AltmanZ = 1.2 \cdot X_1 + 1.4 \cdot X_2 + 0.6 \cdot X_4 + 1 \cdot X_5, \quad (1)$$

where the codes refer to the financial characteristics, the same applies to other formulas. The current ratio is calculated by dividing short term assets and amounts payable and liabilities (2):

$$CurrentRatio = \frac{BSLT\ 00021}{BSLT\ 00055}, \quad (2)$$

The net profit margin was obtained by dividing the profit/loss before taxes by the sales amount (3):

$$NetProfitMargin = \frac{ISLT\ 00017}{ISLT\ 00001}, \quad (3)$$

The ROCE was computed by dividing the profit/loss before taxes by the net worth and amounts payable and liabilities totals (4):

$$ROCE = \frac{ISLT\ 00017}{BSLT\ 00040 + BSLT\ 00055}, \quad (4)$$

Accounts payable turnover in days was calculated by dividing payables and liabilities by sales and multiplying the total by 365 (5):

$$AccountsPayableTurnovers = \frac{BSLT\ 00055}{ISLT\ 00001} \cdot 365, \quad (5)$$

In addition to computing extra ratios and scores, data cleaning procedures were executed. Records containing null or infinity values were eliminated, removing 1146 records. Upon dataset analysis, significant noise was detected across all features. To fix this, 1180 records were deleted using Z-Score outlier detection, which involves subtracting the mean from the value, dividing the result by the standard deviation, and filtering the value [36]. Additionally, 224 records were removed after expert evaluation, leaving 5845 for further examination. The four different combinations of this final dataset used in the experiments are explained in the section "Experimental Setup".

Models

In this section the used models were presented. It contains statistical: Linear Regression (LR), Huber Regression (HR), and ML techniques: Artificial Neural Network (ANN), and Random Forest (RF).

Linear Regression

Linear Regression (LR) is a statistical method used to model the relationship between a dependent variable and one or more independent variables. The model assumes that the relationship between the dependent and independent variables is linear [37]. Based on the linear relationship, the formula can be constructed to perform the prediction task. It is a simple and useful use of linear regression [38].

Due to the use of various data scenarios in this research, two different multiple linear regression formulas were created (see more in the Experimental Setup section). The feature significance analysis was done to determine which independent variables were best for each formulation [37].

Huber Regression

Huber Regression (HR) strikes a balance between squaring errors, like Linear Regression, and computing absolute errors, like Mean Absolute Error Regression, to handle outliers effectively. The primary goal of HR is to reduce the difference between the values predicted by the model and the actual observed values. When the errors are small, meaning the predictions are close to the actual values, HR behaves similarly to Linear Regression and squares these errors. On the other hand, when the errors are large, indicating a significant difference between the predictions and actual values, HR acts like Mean Absolute Error Regression and computes their absolute values. The shift from squaring to absolute at which the model switches from squaring errors to taking their absolute values [39].

Artificial Neural Network

ANNs, inspired by the human brain, are made up of interconnected neurons that process inputs and generate outputs. The network adjusts input weights using backpropagation and optimization algorithms, allowing it to learn complex data patterns and improve performance over time [40].

The main principle of ANN is that it learns by adjusting the connection between neurons. A training set consists of input patterns and associated labels encoding the characteristics the network should learn. The ANN adjusts connection strengths, learning to classify data accurately. Once trained, networks can generalize the results by extending their learning to other datasets. The new data must not significantly differ from the training set for this generalization to be made. In essence, the network's ability to classify new data accurately is dependent on the similarity between the training and new data [41].

In this research, a simple model architecture was selected, comprising 64, 32, and 1 (4 for classification) neurons in the input, hidden, and output layers. Additionally, the ANN was trained with varying: Batch Size, Epochs, Optimizers with different learning rates, and Activation functions.

Random Forest

A Random Forest (RF) is a tree-based model that systematically splits an input dataset into two subsets based on a specific rule, repeating this process until a certain condition is met. The end points of these trees, known as leaf nodes or leaves, represent the final divisions made by the model. In the context of a predicting credit ratings, RF uses the average prediction of all the trees to generate a result. This method is particularly effective as it reduces variance and prevents overfitting [42].

To optimize the model's hyperparameters, tuning will involve modifying N-Estimators, Max Depth, Min Samples (Split), and Min Samples (Leaf).

Proposed method

To validate a hypothesis that emerged during the review of related studies, we suggest a two-tiered approach: initially, we employ a classification algorithm to predict the rating class. Subsequently, depending on the predicted class, we selected a different structure of machine learning algorithm which was trained based on data from that class and attempted to predict the rating. For both classification and prediction tasks, an ANN model was chosen.

Performance Metrics

The performance of rating value **prediction** models will be evaluated on the following four main characteristics:

1. Mean Squared Error (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (6)$$

2. Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (7)$$

3. R-squared (R^2) score:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (8)$$

Where in all formulas: y_i is the actual value of the i -th value, \hat{y}_i is the predicted value of the i -th value, n – the total number of records and \bar{y} is the mean of the actual values in the test set.

4. Root Mean Square Error (RMSE).

$$RMSE = \sqrt{MSE}, \quad (9)$$

To obtain preliminary results for the evaluation of the proposed method, four main metrics have been chosen for the **classification** problem: Accuracy, Precision, Recall, and F1-Score (F1S).

Results

In this section, the datasets and experiments with methods were presented. Also, described each experiment scenario, hyper-parameter tuning, and the results obtained.

Experimental Setup

Datasets vary based on whether all financial features or just the initial ones are used, and whether the value for 2022 is predicted or done randomly. The data is randomly split to match the number of 2022 records as the test data:

- **FinDataRandom**- only financial features from database (for Linear Regression - without BSLT00021 and ISLT00017) and random split (89% training, 11% test data)
- **FinData2022** - only financial features from database (for Linear Regression - without BSLT00021 and ISLT00017) and manual split (2022 years data as test)
- **AllRandom** - all financial features with calculated ratios/scores and random split (89% training, 11% test data)
- **All2022** - all financial features with calculated ratios/scores and manual split (2022 years data as test)

All models were implemented using the SKLearn library, except for only one ANN - from Keras.

The first model in the experimental part is **Linear Regression (LR)**. The LR functions used in the experiments were constructed in two ways: from the SKLearn library (SK-LR) and manually created using coefficients received from the *statsmodels* module (StatsM-LR). Using the *statsmodels* with all publicly available financial features and calculated ratios – with every feature there was high enough to be included in the linear regression formula, although, whenever we used a dataset containing only publicly available features, the BSLT00021 and ISLT00017 were removed due to high p value [37].

The other statistical method in the experiments is **Huber Regression (HR)**. The model includes a parameter named *epsilon*. A smaller *epsilon* – less sensitive model to extreme data points. The GridSearchCV method from the SKLearn library was used to find the optimal *epsilon*. In further experiments, the Huber Regression model will be marked as HR-XX, while XX means the optimal *epsilon* parameter for that dataset (testing range 1-100, step - 1).

One of the machine learning models included in this research was **Artificial Neural Network (ANN)**. Using grid search, various parameters were systematically tweaked to optimize the model's performance. These parameters included batch sizes (b), the number of epochs (e), various optimizers with differing learning rates, and a range of activation functions (see Table 2).

Table 2

Configuration of ANNs hyperparameters

Batch size (b)	Epochs (e)	Optimizer (lr)	Activation (f)
[10, 20, 40, 60, 80, 100, 140, 220]	[20, 50, 80, 120, 160, 210, 250]	SGD (0.1, 0.05, 0.01, 0.001), Adam (0.1, 0.05, 0.01, 0.001)	relu, tanh, sigmoid

In total of 1344 different scenarios of ANN use were performed for each dataset. The scenarios that achieved the best performance of evaluation metrics for each dataset are shown in Table 3.

Table 3

Configuration of ANN hyper-parameters which achieved the best performance

Model Scenario	Batch size	Epochs	Optimizer (lr)	Activation (f)
ANN1	100	120	Adam (0.01)	sigmoid
ANN2	80	120	Adam (0.01)	tanh
ANN3	80	50	Adam (0.01)	sigmoid
ANN4	220	250	Adam (0.01)	tanh

The other machine learning approach which was used in this research is **Random Forest** (RF) algorithm. Similarly to an ANN, the four different hyper-parameters were experimented with. It included n-estimators, max depth, minimum samples for the split and of the leaf (see Table 4).

Table 4

Configuration of RF hyperparameters

N-Estimators	Max Depth	Min Samples (Split)	Min Samples (Leaf)
[100, 200, 300, 500, 750, 1000]	[None, 5, 10, 20, 35]	[2, 3, 5, 7, 10, 15, 20, 30]	[1, 2, 4, 6, 8]

In total, a 1200 experiments with each dataset were conducted using RF model. The Table below displays the scenarios, that obtained the best performance of evaluation metrics for each dataset.

Table 5

Configuration of RF hyperparameters which achieved the best performance

Model Scenario	N-Estimators	Max Depth	Min Samples (Split)	Min Samples (Leaf)
RF1	200	20	7	1
RF2	300	35	2	1
RF3	200	10	5	1
RF4	200	20	2	1

For the proposed method the ANN were chosen to perform both: classification and prediction tasks. Due to small amount of data for E rating, the two classes were combined: D and E. The experiments were conducted with the same parameters as shown in Table 2. For the classification problem, the scenarios of the ANN's parameters that achieved the best results for each dataset are shown in Table 10 and for the rating prediction, the best result for each four rating classes (D and E combined) and for each dataset is shown in Table 11. A total of 33,000 different experiments were conducted for the classification and prediction task to achieve preliminary results of the proposed method.

Experimental Results

In this section the results of each dataset and each method with different scenarios are presented.

Testing Results for each Dataset

In the table below, we can see the dataset called All2022 results. The machine learning models outperformed the statistical ones. These two models did not differ much from each other and the same may be said about ANN and RF models.

Table 6

Results of All2022 dataset

Model	MSE	MAE	R ²	RMSE
SK-LR	0.1746	0.3222	0.7409	0.4178
StatsM-LR	0.1737	0.3218	0.7422	0.4168
HR-5	0.1746	0.3222	0.7409	0.4179
ANN1	0.0630	0.1858	0.9065	0.2510
RF1	0.0669	0.1881	0.9007	0.2587

The results of the dataset known as AllRandom is displayed in the table below. It performed slightly worse than previous dataset.

Table 7

Results of AllRandom dataset

Model	MSE	MAE	R ²	RMSE
SK-LR	0.1789	0.3332	0.7420	0.4229
StatsM-LR	0.1776	0.3321	0.7438	0.4214
HR-32	0.1789	0.3332	0.7420	0.4229
ANN3	0.0671	0.1932	0.8996	0.2591
RF3	0.0680	0.1943	0.8983	0.2607

The table below presents results from the FinData2022 dataset. It is interesting that the results of all statistical methods significantly decreased compared with the other two datasets where data included all the financial features (initial and calculated ones). In this dataset, the difference between the results of ANN and RF models was highest.

Table 8

Result of FinData2022 dataset

Model	MSE	MAE	R ²	RMSE
SK-LR	0.3951	0.4914	0.4137	0.6286
StatsM-LR	0.3905	0.4891	0.4206	0.6249
HR-18	0.3951	0.4914	0.4137	0.6286
ANN2	0.0707	0.1996	0.8951	0.2659
RF2	0.0795	0.2106	0.8820	0.2820

The table below shows the outcomes of the FinDataRandom dataset. The first time when RF outperformed the ANN, admittedly, very little.

Table 9

Result of FinDataRandom

Model	MSE	MAE	R ²	RMSE
SK-LR	0.4031	0.4943	0.4186	0.6349
StatsM-LR	0.4029	0.4948	0.4189	0.6347
HR-4	0.4031	0.4943	0.4186	0.6349
ANN4	0.0734	0.2055	0.8902	0.2709

RF4	0.0728	0.2055	0.8911	0.2698
------------	---------------	---------------	---------------	---------------

The best results were achieved with the ANN1 scenario, with a batch size of 100, number of epochs of 120, optimizer Adam with a learning rate of 0.01, and an activation function being sigmoid. The model obtained significant quite good results: MSE being 0.0630, MAE - 0.1858, R² - 0.9065, and RMSE - 0.2510. The model achieved the best results using all data from features and calculated ratios while splitting the dataset to train as previous years and test as the newest one.

The results of the proposed model's classification task are presented in Table 10. All the indicators are around 0.8, which is quite a high score. It can be concluded that there is no significant difference between the results of each dataset.

Table 10

Results of ANN hyper-parameters configuration for a classification task for each dataset

Dataset	Parameters	Accurac y	Precisio n	Recal l	F1S
All2022	b=220, e=250, Adam (0.001), relu	0.798	0.822	0.798	0.796
AllRandom	b=220, e=20, Adam (0.001), relu	0.793	0.799	0.793	0.792
FinData2022	b=220, e=50, Adam (0.001), relu	0.802	0.804	0.802	0.800
FinDataRando m	b=220, e=20, Adam (0.001), relu	0.780	0.779	0.780	0.777

The preliminary results of the prediction task using the proposed method for each dataset and class, with parameter configuration, are shown below. These results appear promising and demonstrate a consistent pattern: incorporating all financial features has a slight positive effect on outcomes.

Table 11

Preliminary result of ANN hyper-parameters configuration for a prediction task for each dataset/class

Dataset/Class	Parameters	MSE	MAE	R ²
All2022_A	b=140, e=250, Adam (0.01), sigmoid	0.0122	0.0850	0.5406
All2022_B	b=60, e=160, Adam(0.01), relu	0.0288	0.1320	0.6059
All2022_C	b=220, e=250, Adam (0.01), sigmoid	0.0402	0.1613	0.4603
All2022_DE	b=80, e=120, Adam (0.01), tanh	0.0561	0.1896	0.5413
AllRandom_A	b=10, e=50, Adam (0.01), relu	0.0133	0.0956	0.4710
AllRandom_B	b=10, e=80, Adam (0.01), relu	0.0293	0.1381	0.6310
AllRandom_C	b=220, e=50, Adam (0.01), sigmoid	0.0399	0.1593	0.4489
AllRandom_DE	b=10, e=120, SGD (0.001), tanh	0.0460	0.1811	0.5680
FinData2022_A	b=60, e=160, Adam (0.01), tanh	0.0143	0.0949	0.4619
FinData2022_B	b=80, e=210, Adam (0.01), tanh	0.0314	0.1403	0.5704
FinData2022_C	b=140, e=210, Adam (0.01), sigmoid	0.0433	0.1657	0.4176
FinData2022_DE	b=10, e=210, Adam (0.01), sigmoid	0.0582	0.1881	0.5240
FinDataRandom_A	b=20, e=210, Adam (0.01), sigmoid	0.0143	0.0952	0.4320
FinDataRandom_B	b=60, e=160, Adam (0.01), sigmoid	0.0315	0.1415	0.6031
FinDataRandom_C	b=20, e=120, SGD (0.1), relu	0.0435	0.1708	0.3990
FinDataRandom_D E	b=100, e=160, Adam (0.01), sigmoid	0.0563	0.1865	0.4708

Conclusions

This research found a big improvement in the performance of statistical methods and a noticeable increase in machine learning results when using all financial features rather than just the initial ones

for prediction (datasets: All vs FinData). In terms of classification, there was not much difference in accuracy between the datasets.

The most favorable outcomes for prediction task were achieved using the All2022 dataset. This suggests that utilizing all available features, even if derived from each other, and training on past data while testing on the latest data is preferable when dealing with annual financial information.

Machine learning algorithms outperformed statistical methods significantly. Linear Regression slightly outperformed Huber Regression, and in most cases, Artificial Neural Network performed better than the Random Forest model. The prediction task improved with rating class data training but note the significant decrease in training data size due to class splitting, which may impact the results.

Future work may include additional publicly available company data as well as other calculated ratios. Furthermore, additional datasets from other countries would be highly advantageous. In addition, it would be useful to experiment with the architecture of ANN itself. Preliminary results of the proposed method are promising, but further experiments are needed in the proposed prediction process.

References

- [1] Y. Shi and X. Li, "An overview of bankruptcy prediction models for corporate firms: A systematic literature review," *Intangible Capital*, vol. 15, no. 2, pp. 114–127, 2019.
- [2] F. Ciampi, A. Giannozzi, G. Marzi, and E. I. Altman, "Rethinking SME default prediction: a systematic literature review and future perspectives," *Scientometrics*, vol. 126, no. 3, pp. 2141–2188, 2021, doi: 10.1007/s11192-020-03856-0.
- [3] H. Kim, H. Cho, and D. Ryu, "Corporate Default Predictions Using Machine Learning: Literature Review," *Sustainability*, vol. 12, no. 16, 2020, doi: 10.3390/su12166325.
- [4] P. Golbayani, I. Florescu, and R. Chatterjee, "A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees," *The North American Journal of Economics and Finance*, vol. 54, p. 101251, 2020, doi: <https://doi.org/10.1016/j.najef.2020.101251>.
- [5] P. Ubarhande and A. Chandani, "Elements of Credit Rating: A Hybrid Review and Future Research Agenda," *Cogent Business & Management*, vol. 8, no. 1, p. 1878977, 2021, doi: 10.1080/23311975.2021.1878977.
- [6] J. Kaur, M. Vij, and A. K. Chauhan, "Signals influencing corporate credit ratings—a systematic literature review," *DECISION*, vol. 50, no. 1, pp. 91–114, 2023, doi: 10.1007/s40622-023-00341-4.
- [7] A. TUOVILA, "Corporate Credit Rating: What it is, How it Works," Investopedia.
- [8] S. Shi, R. Tse, W. Luo, S. D'Addona, and G. Pau, "Machine learning-driven credit risk: a systemic review," *Neural Comput Appl*, vol. 34, no. 17, pp. 14327–14339, 2022, doi: 10.1007/s00521-022-07472-2.
- [9] A. Bhattacharya, S. Kr. Biswas, and A. Mandal, "Credit risk evaluation: a comprehensive study," *Multimed Tools Appl*, vol. 82, no. 12, pp. 18217–18267, 2023, doi: 10.1007/s11042-022-13952-3.
- [10] M. R. Machado and S. Karray, "Assessing credit risk of commercial customers using hybrid machine learning algorithms," *Expert Syst Appl*, vol. 200, p. 116889, 2022, doi: <https://doi.org/10.1016/j.eswa.2022.116889>.
- [11] M. Moscatelli, F. Parlapiano, S. Narizzano, and G. Viggiano, "Corporate default forecasting with machine learning," *Expert Syst Appl*, vol. 161, p. 113567, 2020, doi: <https://doi.org/10.1016/j.eswa.2020.113567>.
- [12] Z. Khemais, D. Nesrine, and M. Mohamed, "Credit scoring and default risk prediction: A comparative study between discriminant analysis & logistic regression," *Int J Econ Finance*, vol. 8, no. 4, p. 39, 2016.
- [13] L. Munkhdalai, T. Munkhdalai, O.-E. Namsrai, J. Y. Lee, and K. H. Ryu, "An empirical comparison of machine-learning methods on bank client credit assessments," *Sustainability*, vol. 11, no. 3, p. 699, 2019.
- [14] N. Muñoz-Izquierdo, M. J. Segovia-Vargas, M.-M. Camacho-Miñano, and Y. Pérez-Pérez, "Machine learning in corporate credit rating assessment using the expanded audit report," *Mach Learn*, vol. 111, no. 11, pp. 4183–4215, 2022, doi: 10.1007/s10994-022-06226-4.
- [15] M. Wallis, K. Kumar, and A. Gepp, "Credit rating forecasting using machine learning techniques," in *Managerial perspectives on intelligent big data analytics*, IGI Global, 2019, pp. 180–198.
- [16] S. Ben Jabeur, A. Sadaoui, A. Sghaier, and R. Aloui, "Machine learning models and cost-sensitive decision trees for bond rating prediction," *Journal of the Operational Research Society*, vol. 71, no. 8, pp. 1161–1179, 2020.
- [17] C. Daniel, J. Hančlová, and H. el Woujoud Bousselmi, "Corporate rating forecasting using Artificial Intelligence statistical techniques," *Investment Management & Financial Innovations*, vol. 16, no. 2, p. 295, 2019.
- [18] G. Teles, J. Rodrigues, R. A. L. Rabê, and S. A. Kozlov, "Artificial neural network and Bayesian network models for credit risk prediction," *Journal of Artificial Intelligence and Systems*, vol. 2, no. 1, pp. 118–132, 2020.
- [19] A. R. Provenzano *et al.*, "Machine learning approach for credit scoring," *arXiv preprint arXiv:2008.01687*, 2020.
- [20] S. Pol and S. S. Ambekar, "Predicting Credit Ratings using Deep Learning Models—An Analysis of the Indian IT Industry," *Australasian Accounting, Business and Finance Journal*, vol. 16, no. 5, pp. 38–51, 2022.

- [21] M. Tavakoli, R. Chandra, F. Tian, and C. Bravo, "Multi-Modal Deep Learning for Credit Rating Prediction Using Text and Numerical Data Streams," *arXiv preprint arXiv:2304.10740*, 2023.
- [22] P. K. Roy and K. Shaw, "A multicriteria credit scoring model for SMEs using hybrid BWM and TOPSIS," *Financial Innovation*, vol. 7, no. 1, p. 77, 2021, doi: 10.1186/s40854-021-00295-5.
- [23] A. Kumar, D. Shanthi, and P. Bhattacharya, "Credit Score Prediction System using deep learning and K-means algorithms," in *Journal of Physics: Conference Series*, IOP Publishing, 2021, p. 012027.
- [24] V. Veeramani and M. Jeyakarthic, "A Futuristic Framework for Financial Credit Score Prediction System using PSO based Feature Selection with Random Tree Data Classification Model," in *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2019, pp. 826–831. doi: 10.1109/ICSSIT46314.2019.8987814.
- [25] E. Dumitrescu, S. Hué, C. Hurlin, and S. Tokpavi, "Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects," *Eur J Oper Res*, vol. 297, no. 3, pp. 1178–1192, 2022, doi: <https://doi.org/10.1016/j.ejor.2021.06.053>.
- [26] P. Gogas, T. Papadimitriou, and A. Agrapetidou, "Forecasting bank credit ratings," *The Journal of Risk Finance*, vol. 15, no. 2, pp. 195–209, 2014.
- [27] P. L. Hammer, A. Kogan, and M. A. Lejeune, "A logical analysis of banks' financial strength ratings," *Expert Syst Appl*, vol. 39, no. 9, pp. 7808–7821, 2012, doi: <https://doi.org/10.1016/j.eswa.2012.01.087>.
- [28] E. K. Laitinen, "Predicting a corporate credit analyst's risk estimate by logistic and linear models," *International Review of Financial Analysis*, vol. 8, no. 2, pp. 97–121, 1999, doi: [https://doi.org/10.1016/S1057-5219\(99\)00012-5](https://doi.org/10.1016/S1057-5219(99)00012-5).
- [29] Jason Brownlee, "Robust Regression for Machine Learning in Python," <https://machinelearningmastery.com/robust-regression-for-machine-learning-in-python/>.
- [30] "A European Strategy for data," <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>.
- [31] "European legislation on open data," <https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data>.
- [32] E. I. Altman, "17 Predicting financial distress of companies: revisiting the Z-Score and ZETA® models1," *Handbook of research methods and applications in empirical finance*, p. 428, 2013.
- [33] D. Mulyadi and O. Sinaga, "Analysis of Current Ratio, Net Profit Margin, and Good Corporate Governance against Company Value.," *Systematic Reviews in Pharmacy*, vol. 11, no. 1, 2020.
- [34] B. Camelia, "Analysis model for return on capital employed," *Annals of the "Constantin Brâncuși" University of Târgu Jiu, Economy Series*, vol. 1, pp. 82–87, 2013.
- [35] Lisa Schwarz, "Accounts Payable Turnover Ratio Defined: Formula & Examples | NetSuite," <https://www.netsuite.com/portal/resource/articles/accounting/accounts-payable-turnover-ratio.shtml#:~:text=The%20accounts%20payable%20turnover%20ratio,during%20a%20specified%20time%20period.>
- [36] H. Abdi, "Z-scores," *Encyclopedia of measurement and statistics*, vol. 3, pp. 1055–1058, 2007.
- [37] Ankita Banerji, "Hypothesis Testing On Linear Regression | by Ankita Banerji | Nerd For Tech | Medium," <https://medium.com/nerd-for-tech/hypothesis-testing-on-linear-regression-c2a1799ba964>.
- [38] D. and H. T. and T. R. and T. J. James Gareth and Witten, "Linear Regression," in *An Introduction to Statistical Learning: with Applications in Python*, Cham: Springer International Publishing, 2023, pp. 69–134. doi: 10.1007/978-3-031-38747-0_3.
- [39] Y. Wang, B. Wang, C. Peng, X. Li, and H. Yin, "Huber Regression Analysis with a Semi-Supervised Method," *Mathematics*, vol. 10, no. 20, 2022, doi: 10.3390/math10203734.
- [40] O. A. Montesinos López, A. Montesinos López, and J. Crossa, "Fundamentals of Artificial Neural Networks and Deep Learning," in *Multivariate Statistical Machine Learning Methods for Genomic Prediction*, O. A. Montesinos López, A. Montesinos López, and J. Crossa, Eds., Cham: Springer International Publishing, 2022, pp. 379–425. doi: 10.1007/978-3-030-89010-0_10.
- [41] B. Mehlig, *Machine learning with neural networks: an introduction for scientists and engineers*. Cambridge University Press, 2021.
- [42] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *Stata J*, vol. 20, no. 1, pp. 3–29, 2020, doi: 10.1177/1536867X20909688.