

Article

Not peer-reviewed version

---

# Application of Machine Learning Algorithm to Assess and Minimize Credit Risks

---

[Garnik Arakelyan](#) and [Armen Ghazaryan](#) \*

Posted Date: 12 August 2025

doi: 10.20944/preprints202508.0783.v1

Keywords: credit risk modeling; credit risk; machine learning; prediction; classification algorithm; scoring



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Application of Machine Learning Algorithm to Assess and Minimize Credit Risks

Garnik Arakelyan <sup>1</sup> and Armen Ghazaryan <sup>1,\*</sup>

<sup>1</sup> Armenian State University of Economics, Paruyr Sevaki str. 77, Yerevan 0069, Armenia

\* Correspondence: armenghazaryanq1@gmail.com

## Abstract

Banking system being the most important sector of the economy of every country often encounters a number of risks. Financial institutions of that system operate in an unstable environment and without having complete information about that environment, they may suffer significant losses. The main source of such losses is considered to be a credit risk, for the management of which various mathematical models are being developed, which allow to make decision on granting a loan. Lately on this level machine learning (ML) classification algorithms are often used for credit risk modeling. In this research work using the ideas of ML well-known algorithms a new algorithm for solving the binary classification problem was developed. By means of the created algorithm based on real data a classification model has been developed, qualitative indicators of that model such as: ROC AUC, PR AUC, Precision, Recall, F1 score, were evaluated. By modifying the resulting probabilities into a range of 300-850 score points a scoring model has been developed, the usage of which can mitigate the credit risk and protect financial organizations from major losses.

**Keywords:** credit risk modeling; credit risk; machine learning; prediction; classification algorithm; scoring

---

## 1. Introduction

Historically, a proper credit risk management plays a key role in the normal functioning of the economy and in the prevention of financial crises. Thus, for example, the global financial crisis of 2008 indicated how important it is to precisely manage the credit risk (Siddique et al., 2022). The Basel Committee on Banking Supervision also emphasizes the active and coordinated management of such type of risk, which contributes to the stability of the banking system (Curcio & Hasan, 2015).

At the time of lending credit organizations often apply various developed models in order to assess the credit risk of an individual, for example, FICO Scoring model (Addo et al., 2018). FICO credit scoring solutions provide a better understanding of credit risk in lending to consumers and businesses. The introduction of broad-based credit scores by FICO more than 30 years ago has transformed economic growth in the US and globally - making access to credit more efficient and objective while protecting the financial stability of lenders by enabling decisions that limit risk (Klein & Thompson, 2021).

This model and its analogues are "black box" and credit organizations often can not clearly understand the reason for the formation of this or that score point, as well as those models are often built based on data of borrowers from only few countries and in essence do not describe the social-economic situation of the operation region of a credit organization. For the same reasons for the accurate evaluation and credit risk management credit organizations often collect the information on their borrowers and the quality of provided loans themselves and based on all that data they build models (Addo et al., 2018). Historically, logistic regression model is used often in order to solve such

problems. Though, recently it is noticed that modern ML models are also often used in order to solve this problem, for example, XGBoost, LightGBM, Neural Network etc. (Zhang et al., 2020).

The aim of this research work is to investigate the behaviour of borrowers based on real data, applying the ideas of known ML models to develop new algorithms for solving binary classification problems. The aim is also to develop local score system of the given bank based on the data of the borrowers available to us by using new algorithms.

The results of achieving the goal of this research can be used by credit organizations, which will contribute to the optimal credit risk management and to the cost reduction in the given organizations.

## 2. Literature Review

Within the framework of this research many scientific works of others have been studied. We shall consider some of them.

In the article named "Advancing credit risk modelling with Machine Learning: A comprehensive review of the state-of-the-art" the authors studied the granting of the loans as the most important part of the banking activity and credit risk originating from it. In this research work the literature, which refers to modeling of consumer credit risk by means of ML, has been studied. The authors analyzed strengths and weaknesses of the ML, the pre-processing of the data, the choice of the variables and optimization and other important phases of the modeling. The results of the research work show, that the implementation of ML in the field of management of the financial risk is increasing. However, standardized modeling procedures are necessary before difficult ML classifiers can be deployed commercially (Montevecchi, et al., 2024).

In the article named "Machine Learning for Credit Risk Prediction: A Systematic Literature Review" the authors raise the question of necessity of use of Artificial intelligence (AI) and ML by financial institutions in order to evaluate credit risk by analyzing large amounts of information. Analyzing popular databases the researchers have discovered 52 researches, which are related to the microfinance lending sector. Research has shown that boosted models are the most researched among ML models, and most often, in order to evaluate the quality of models, such metrics as Area Under Curve (AUC), Accuracy (ACC), Recall, precision measure F1 (F1) and Precision (Niu & Zheng, 2020) are used.

In the article named "Credit Risk Prediction Using Machine Learning and Deep Learning: A Study on Credit Card Customers" the researchers studied approaches predicting default of the credit card customers using various models of ML including neural networks, logistic regression, AdaBoost, XGBoost and LightGBM. The results of the research work show, that XGBoost outperforms other models reaching 99,4% accuracy. The study has shown that application of algorithms of ML and deep learning (DL) significantly improved accuracy of forecasts in the field of customer default (Chang, et al., 2024).

In the article named "Machine learning-driven credit risk: a systemic review" the researchers studied a number of major research works over the last eight years (76 articles), where in order to solve credit risk problems statistical, ML and DL methods were used. Using available datasets a new classification methodology of ML-based credit risk algorithms and their efficiency ranking was proposed. The research has shown, that deep learning models for the credit risk assessment are more accurate than statistical algorithms of ML. Authors also mention that ensemble methods provide higher accuracy than separate models (Shi, et al., 2022).

In the article named "Financial Inclusion in Emerging Economies: The Application of Machine Learning and Artificial Intelligence in Credit Risk Assessment" as a result of the conducted research, it was revealed, that ML and artificial intelligence (AI) have a significant impact on the credit risk assessment, during which various available data is also used. All this allows lenders to conduct in-depth research of the behavior of the borrowers. The authors advise financial organizations to invest in AI and ML fields (Mhlanga, 2021).

In the article named "Machine Learning for an Enhanced Credit Risk Analysis: A Comparative Study of Loan Approval Prediction Models Integrating Mental Health Data" the researchers studied

two subsets: the data about the mental health of people and choice about the loan approval. The data was generated in terms of 1991 people. This study work shows how chosen ML algorithms can distinguish risky and non-risky customers (Alagic, et al., 2024).

In the article named "Responsible Credit Risk Assessment with Machine Learning and Knowledge Acquisition" the authors mention, that human experience can be used for correction of an inaccurate data. In the research ML algorithms were used on a small amount of data, then errors in the data were corrected through human experience using Ripple-Down rules. As a result, it was revealed, that combined model, which learned from small amount of data works in the same way as the model, which learned from big amount of data. The presented approach is universal and can be applied to improve the validity of decisions on lending (Guan, et al., 2023).

In the article named "Machine Learning for Enhanced Credit Risk Assessment: An Empirical Approach" ten ML models are evaluated using more than 2,5 million observations. The results of the research show, that ensemble models, especially, XGBoost model qualitatively is better than traditional algorithms such as logistic regression (Suhadolnik, Ueyama, & Da Silva, 2023).

In the article named "Comparison of the impact of dimensionality reduction and data splitting on classification performance in credit risk assessment" it has been studied how changes in data dimension affect the efficiency of classification algorithms. In the research work data dimension was carried out using PCA method, and in order to solve the classification problem Random Forest, Logistic Regression, Decision Tree and Naive Bayes algorithms were used. The research showed, that the models, that were built based on that dataset, which consists of a total of 20 features, which are expressed with 13 PC, show same or better results, than the models, that are have been developed using the original database (Bulut & Arslan, 2024).

In the article named "A study on credit scoring modeling with different feature selection and machine learning approaches" it was demonstrated, that when building a credit score it is possible to improve the forecast results by choosing the right combination of different variable selection methods and ML algorithms. In this research work available credit data from Germany was used. In this study work an analysis was carried out, as a result of which the best combination between ML models and variable selection methods was revealed: random forest and Chi-Square (Trivedi, 2020).

Summarizing the reviewed researches we can mention, that in the scientific world various literature is often studied and in practical model building known models of ML are used. For that reason, within the framework of this research work the results of the conducted research are even more up-to-date.

### 3. Research Methodology

Within the framework of this research various data processing methods were used. In order to identify patterns in the data the correlation analysis method was used.

Correlation is a bivariate analysis that measures the strength of association between two variables and the direction of the relationship (Smith & Brown, 2020): This coefficient shows a linear relationship between variables.

Taking into account that the dataset contains categorical variables, that data was processed using One Hot Encoding method.

One Hot Encoding is a method for converting categorical variables into a binary format. It creates new binary columns (0s and 1s) for each category in the original variable. Each category in the original column is represented as a separate column, where a value of 1 indicates the presence of that category, and 0 indicates its absence (Smith & Brown, 2020).

Taking into account that the dataset contains numeric variables, the data was processed using StandardScaler method.

StandardScaler is a versatile and widely used preprocessing technique that contributes to the robustness, interpretability, and performance of machine learning models trained on diverse datasets.

StandardScaler operates on the principle of normalization, where it transforms the distribution of each feature to have a mean of zero and a standard deviation of one. This process ensures that all features are on the same scale, preventing any single feature from dominating the learning process due to its larger magnitude (Zhang & Wang, 2020).

In order to solve the classification problem within the framework of this research a new binary classification algorithm has been developed, which consists of the following steps mentioned below:

1. Depending on the dependent variable ( $y$ ) it is necessary to divide the existing dataset ( $S$ ) into two subsets. First subsets will consist of ( $S_0$ ) those observations, for which  $y$  takes the value of 0, and second ( $S_1$ )  $y$  takes the value of 1. It is appropriate

$$S_0 = \{x_i \in S | y_i = 0\}, \quad S_1 = \{x_i \in S | y_i = 1\},$$

where  $x_i$ ;  $i$  is the vector of the observation variable,  $S$  is the initial dataset.

$$S = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} & y_1 \\ x_{21} & x_{22} & \dots & x_{2j} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{31} & x_{32} & \dots & x_{ij} & y_i \end{pmatrix} \quad y \in (0,1),$$

$$S_0 = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} \\ x_{21} & x_{22} & \dots & x_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ x_{31} & x_{32} & \dots & x_{ij} \end{pmatrix} \quad y = 0,$$

$$S_1 = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} \\ x_{21} & x_{22} & \dots & x_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ x_{31} & x_{32} & \dots & x_{ij} \end{pmatrix} \quad y = 1$$

2. For each subset obtained in point 1 we find the centroid of the given group according to the formulas given below:

$$C_0 = \left( \frac{1}{S_0} \sum_{i=1}^n x_{i1}, \frac{1}{S_0} \sum_{i=1}^n x_{i2}, \dots, \frac{1}{S_0} \sum_{i=1}^n x_{ij} \right),$$

$$C_1 = \left( \frac{1}{S_1} \sum_{i=1}^n x_{i1}, \frac{1}{S_1} \sum_{i=1}^n x_{i2}, \dots, \frac{1}{S_1} \sum_{i=1}^n x_{ij} \right),$$

where  $C_0$  and  $C_1$  are the centroids of the choices, and  $S_0$  and  $S_1$  is the amount of the relevant subsets.

3. We calculate the distances from the point subject to classification to the centroid of each group and between given selected point and  $N$  neighbors. The distance can be measured using various methods. Within the framework of this study Euclidean distance was applied according to the formula given below (Gallet & Gowanlock, 2022):

$$p = (p_1, p_2, \dots, p_n) \quad q = (q_1, q_2, \dots, q_n)$$

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{k=1}^n (p_k - q_k)^2},$$

where  $p$  and  $q$  are two points in  $n$ -dimensional space.

4. Applying the distances obtained in point 3 we calculate the influence of each centroid and each nearest neighbor according to the formula mentioned below:

$$Impact = \frac{1}{d(p, q)^k + e},$$

where  $d(p, q)$  is the distance between  $p$  and  $q$  points, and  $k$  is the impact factor,  $e$  is a the rational number.

5. The probability of belonging to the class 0 and 1 of the point subject to classification are calculated according to the formulas given below:

$$Impact(0) = Impact(C_0) + \sum_{i=1}^{N_0} Impact(d_{0_i}), N_0 \in N$$

$$Impact(1) = Impact(C_1) + \sum_{i=1}^{N_1} Impact(d_{1_i}), N_1 \in N$$

$$Probability(0) = \frac{Impact(0)}{Impact(0) + Impact(1)}$$

$$Probability(1) = \frac{Impact(1)}{Impact(0) + Impact(1)}$$

where  $Impact(C_0)$  and  $Impact(C_1)$  are the influence of class centroids 0 and 1,  $Impact(d_{0_i})$  and  $Impact(d_{1_i})$  are the influences from a given point to the N nearest neighbors (0 and 1 class neighbors),  $Probability(0)$  and  $Probability(1)$  the probabilities of belonging to groups 0 and 1 of the point subject to classification.

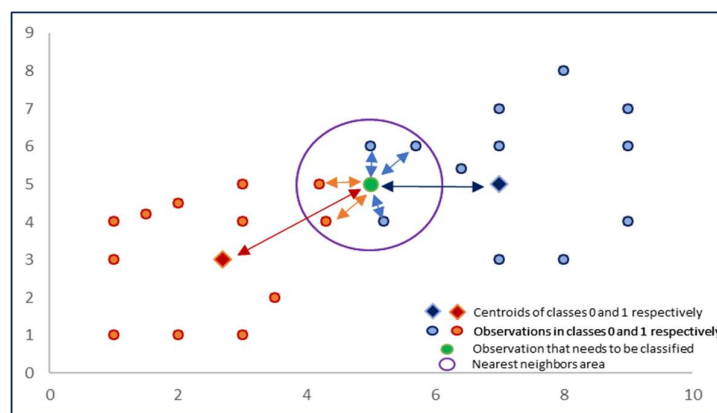


Figure 1. Classification algorithm diagram.

#### 4. Analysis

The research is based on real credit information by example of “Unibank” OJSC. In this dataset there are 13 variables:

1. Marital Status. In the original data, this variable had the values M1, M2, M3, and M4. Each value has been changed to “Married”, “Widow”, “Divorced”, and “Single”, respectively;
2. Education. In the original data, this variable had the values E1, E2, E3, and E4. Each value has been changed to “Academic degree”, “Higher education”, “Medium professional”, and “High school”, respectively;
3. Availability of property. In the original data, this variable had the values P1, P2, P3, and P4. Each value has been changed to “Availability of real estate”, “Availability of movable property”, “Availability of real estate and movable property”, and “Absence”, respectively;
4. Sex. In the original data, this variable had the values S1 and S2. Each value has been changed to “Female”, and “Male”, respectively;
5. Borrower’s Age. The variable is numeric and contains values from 20 to 65;
6. Number of days past due in the last 12 months. The variable is numeric and contains values from 0 to 1600;
7. Number of delays. The variable is numeric and contains values from 1 to 67. This variable contains missing data. Based on the information received from the bank, it became clear that these clients do not have overdue debts. Based on this, the missing values were changed to 0;

8. Number of changes in risk classes. The variable is numeric and contains values from 0 to 28;
9. Credit load. The variable is numeric and contains values from 0 to 2,984,303. Data are presented in AMD;
10. Credit history length. The variable is numeric and contains values from 0 to 488. Data is presented in days;
11. Maximum repaid loans. The variable is numeric and contains values from 0 to 25,131,048. Data are presented in AMD;
12. Contract sum. The variable is numeric and contains values from 30,000 to 2,300,000. Data are presented in AMD. There are some outliers in the data. These are the clients to whom the bank has issued a loan in the amount of more than AMD 1,000,000. There are 7 such observations in the sample. It was calculated that the removal of observational data does not significantly affect the quality of the models. Therefore, when creating models, these observations were not removed from the dataset;
13. Default. Binary data is presented. 1 - default, 0 - no default. Based on the information received from the bank, a default is considered the presence of overdue obligations for 90 days or more at least once in the last year.

In the Dataset there are 4603 observations. For modeling purposes Default variable was chosen as the dependent variable.

Both the analysis and the model building were carried out with the help of the Excel program, Python programming language and its appropriate libraries, which are Pandas, Matplotlib, Seaborn, Scikit-learn libraries.

In the table 1 a descriptive statistics of numerical variables is presented. It can be seen that there are no omissions in the data, because in all variables the quantity is mentioned to be 4603. The statistical average for all the variables is represented in the table, minimum and maximum values, as well as 25%, 50% u 75% percentiles (Thompson & Lee, 2022).

**Table 1.** Descriptive statistics.

Indicator	Contract sum	Age	Number of days past due in the last 12 months	Number of delays	Number of changes in risk classes	Credit load	Credit history length	Maximum repaid loans	Default
<b>count</b>	4 603	4 603	4 603	4 603	4 603	4 603	4 603	4 603	4 603
<b>mean</b>	254229,24	37,54	4,03	2,71	0,55	471551,5	59,39	332959,85	0,43
<b>std</b>	144770,86	12,09	38,54	4,28	1,46	537158,02	61,27	920340,39	0,49
<b>min</b>	30000	20	0	0	0	0	0	0	0
<b>25%</b>	181900	27	0	0	0	0	13	0	0
<b>50%</b>	200000	36	0	1	0	300000	40	150000	0
<b>75%</b>	377050	48	0	4	0	825100	86	376000	1
<b>max</b>	2300000	65	1600	67	28	2984303	488	25131048	1

A correlation matrix was constructed for the numerical variables (Table 2), which shows that there is a moderate connection between "Number of changes in risk classes" and "Number of delays", "Credit history length" and "Credit load" variables.

**Table 2.** Correlation analysis.

	Contract sum	Age	Number of days past due in the last 12 months	Number of delays	Number of changes in risk classes	Credit load	Credit history length	Maximum repaid loans	Default
<b>Contract sum</b>	1,00	0,07	0,04	0,01	0,01	0,07	0,10	0,13	0,10

Age	0,07	1,00	0,01	0,05	0,08	0,14	0,21	0,12	-0,08
Number of days past due in the last 12 months	0,04	0,01	1,00	0,34	0,38	-0,01	0,09	0,14	0,08
Number of delays	0,01	0,05	0,34	1,00	0,57	0,05	0,35	0,18	0,04
Number of changes in risk classes	0,01	0,08	0,38	0,57	1,00	0,00	0,22	0,12	-0,05
Credit load	0,07	0,14	-0,01	0,05	0,00	1,00	0,50	0,17	0,15
Credit history length	0,10	0,21	0,09	0,35	0,22	0,50	1,00	0,26	-0,05
Maximum repaid loans	0,13	0,12	0,14	0,18	0,12	0,17	0,26	1,00	0,00
Default	0,10	-0,08	0,08	0,04	-0,05	0,15	-0,05	0,00	1,00

Within the framework of this research work various variables were analyzed, which shows that default is accumulated in in the age group of 20-25 (50,49%, Table 3, Figure 2), in the "widow" group of customers (54,14%, Table 4, Figure 3), as well as in the group of customers that have real estate (44,36%, Table 5, Figure. 4) and in the group of customers that have medium professional education level (42,83%, Table 6, Figure 5).

Table 3. Analysis of age groups.

Age range	Default		Total	Default %
	No	Yes		
1. 20-25	457	466	923	50,49%
2. 26-35	760	589	1349	43,66%
3. 36-50	880	558	1438	38,80%
4. 51+	545	348	893	38,97%

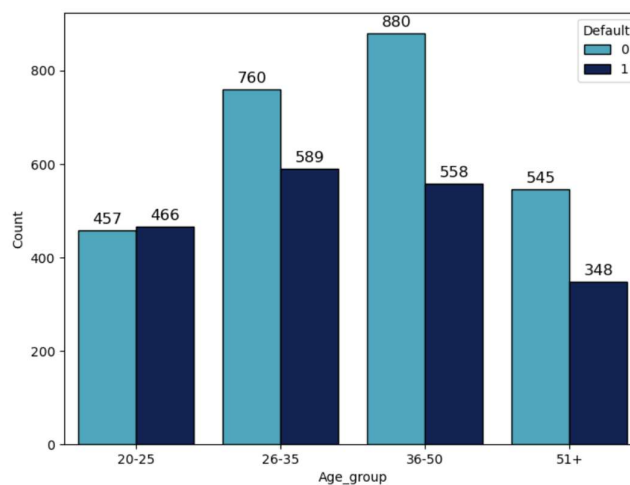


Figure 2. Distribution by age group with division by default.

Table 4. Analysis of clients by marital status.

Marital Status	Default		Total	Default %
	No	Yes		
Divorced	46	13	59	22,03%
Married	2139	1428	3567	40,03%
Single	42	30	72	41,67%

<b>Widow</b>	415	490	905	54,14%
--------------	-----	-----	-----	--------

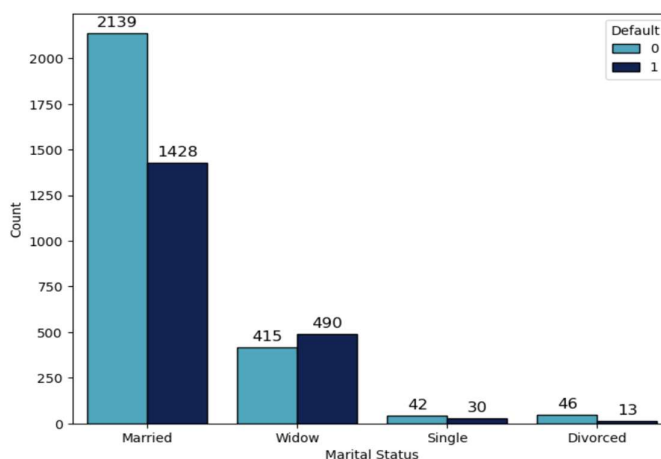


Figure 3. Distribution by marital status with division by default.

Table 5. Analysis of customers depending on availability of property.

Availability of property	Default		Total	Default %
	No	Yes		
Availability of real estate and movable property	133	99	232	42,67%
Availability of real estate	153	122	275	44,36%
Availability of movable property	611	464	1075	43,16%
Absence	1745	1276	3021	42,24%

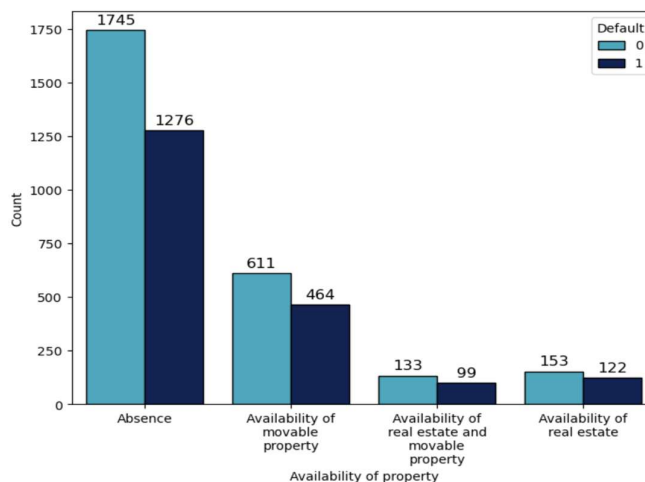
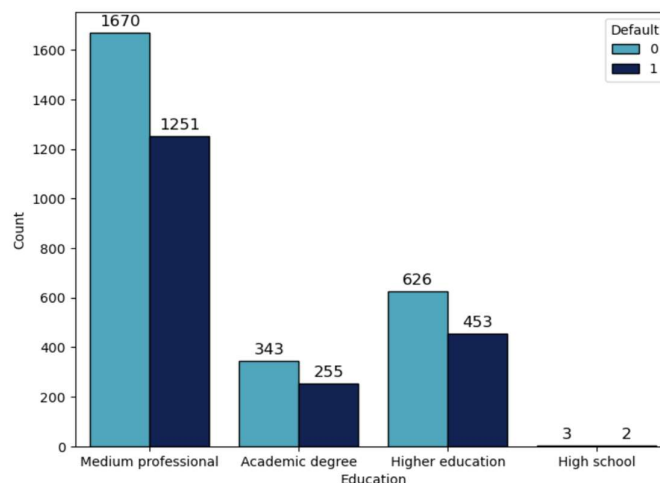


Figure 4. Distribution by availability of property with division by default.

Table 6. Analysis of customers depending on the level of education.

Education	Default		Total	Default %
	No	Yes		
Academic degree	343	255	598	42,64%
High school	3	2	5	40,00%
Higher education	626	453	1079	41,98%
Medium professional	1670	1251	2921	42,83%



**Figure 5.** Distribution by education with division by default.

After completing the analysis it is necessary to develop a classification model according to the methodology mentioned in this research, to evaluate the qualitative indicators of that model and to modify the obtained probabilities into 300-850 score points.

For the modeling purposes certain modifications have been made to the existing data, in accordance with the points mentioned below:

1. The entire dataset was divided into training (80%) and testing (20%) subsets;
2. For each subset obtained in the point 1 - "Contract sum", "Age", "Credit load", "Credit history length", "Number of days past due in the last 12 months", "Number of delays" and "Maximum repaid loans" variables were standardized, using StandardScaler instrument of Sklearn library of Python programming language (Pedregosa et al., 2011).
3. For all categorical variables dummy variables were created, using get\_dummies instrument of Pandas library of Python programming language (Pandas Development Team, 2023).

In the Research methodology section of this research the proposed methodology was implemented using Python programming language. The amount of N neighbors was chosen from the range of 2 - 15, and k coefficient from the range of 0.0001 - 0.2001. N was chosen to be 12, and k to be 0.0001, because that combination provided the best result for the model.

In the figure 6 ROC and PR curves and appropriate AUC coefficients are presented (Bradley, 1997).

As the best threshold of the model the threshold of 0.4286 was chosen, in which case the accuracy is equal to 0.6938, Precision is equal to 0.7005, Recall is equal to 0.6938, and F1 Score is equal to 0.6957: In the figure 7 the Confusion matrix of the built model is presented, from which it is evident that the model revealed around 68% of defaults (258/378), but at the same time it refuses 162 good customers, which makes around 30% (162/543) of the good customers in the test subset (Dyrland et al., 2023).

After the above as a result of the prediction of the test subset the probability of obtained 0 class was modified into score points, which can be in the range of 300 – 850. The modification is carried out according to the following formula:

$$Score = \left( \ln \left( \frac{P(y=0)}{1-P(y=0)} \right) * \frac{850-300}{|Percentile_{1}| + |Percentile_{99}|} \right) + \left( 850 - \frac{850-300}{|Percentile_{1}| + |Percentile_{99}|} * Percentile_{1} \right),$$

where Score is a score point,  $P(y=0)$  is the probability of class 0,  $Percentile_{1}$  and  $Percentile_{99}$  are the 1st and 99th percentiles of the logarithmic row.

In the Figure 8 the distribution of the resulting score groups is presented.

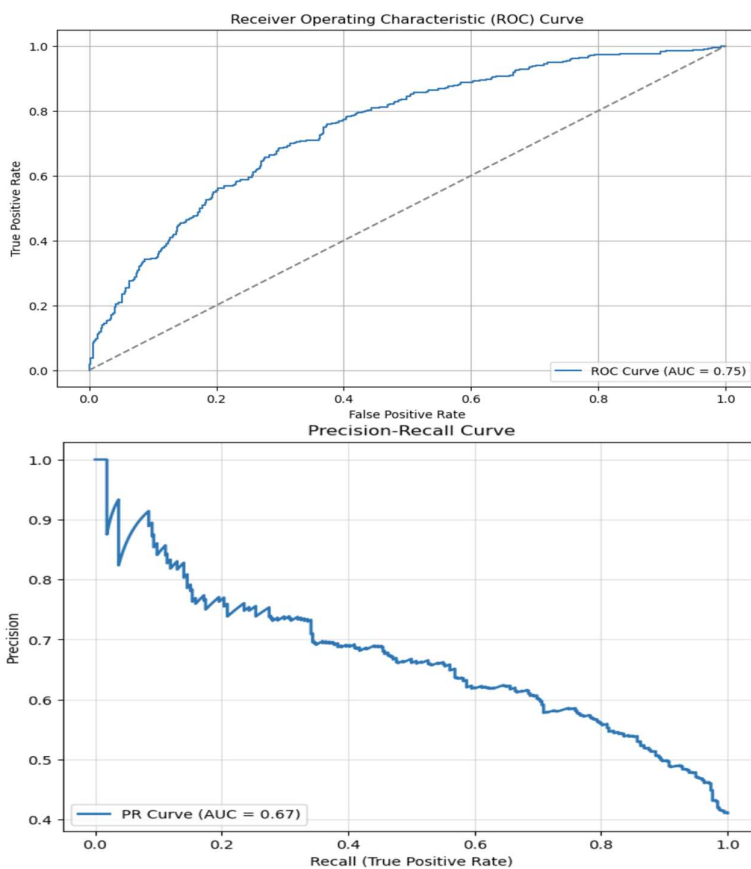


Figure 6. Curves ROC and PR.

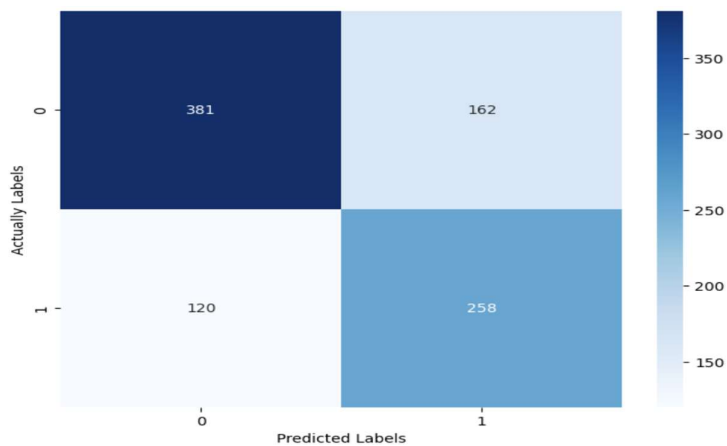


Figure 7. Confusion matrix.

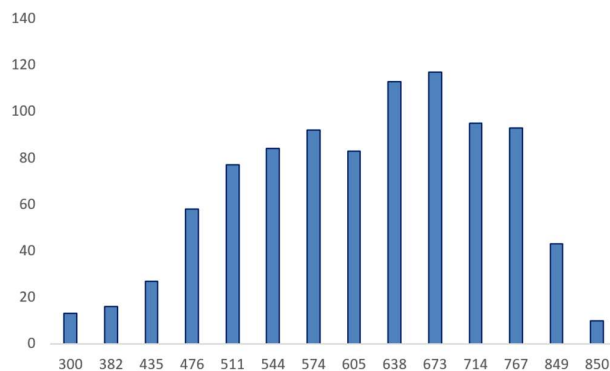


Figure 8. Distribution of scores.

## 5. Conclusions

It is important for every credit organization to identify and manage credit risks. In order to mitigate that risk, it is important to implement a reasonable decision-making system. The aim of this research was to develop a new machine learning model for binary classification. The results of this research work show, that as a result of a combination of ideas leaning in the foundation of the well-known classic models it is possible to get a new algorithm, which will show quite a big result. Within the framework of this topic the use of new algorithms of this type can mitigate credit risk encountered by the credit organizations and the losses resulting from it.

**Author Contributions:** Conceptualization, G.A. and A.G.; methodology, G.A. and A.G.; software, G.A.; validation, G.A.; formal analysis, G.A.; investigation, G.A.; writing—original draft preparation, G.A. and A.G.; writing—review and editing, G.A. and A.G.; visualization, G.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author due to privacy reasons.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. *Expert Systems with Applications*, 134, 26-39. <https://doi.org/10.1016/j.eswa.2015.12.013>
2. Alagic, A., Zivic, N., Kadusic, E., Hamzic, D., Hadzajlic, N., Dizdarevic, M., & Selmanovic, E. (2024). Machine learning for an enhanced credit risk analysis: A comparative study of loan approval prediction models integrating mental health data. *Machine Learning and Knowledge Extraction*, 6(1), 53-77 <https://doi.org/10.3390/make6010004>
3. Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7), 1145-1159. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2)
4. Bulut, C., & Arslan, E. (2024). Comparison of the impact of dimensionality reduction and data splitting on classification performance in credit risk assessment. *Artificial Intelligence Review*, 57(9), 252. <https://doi.org/10.1007/s10462-024-10904-1>
5. Chang, V., Sivakulasingam, S., Wang, H., Wong, S. T., Ganatra, M. A., & Luo, J. (2024). Credit Risk Prediction Using Machine Learning and Deep Learning: A Study on Credit Card Customers. *Risks*, 12(11), 174. <https://doi.org/10.3390/risks12110174>
6. Curcio, D., & Hasan, I. (2015). Earnings management and bank risk during the financial crisis. *International Review of Financial Analysis*, 39, 237-249. <https://doi.org/10.1016/j.irfa.2015.04.017>

7. Dyrland, K., Lundervold, A. S., & Porta Mana, P. G. L. (2023). Does the evaluation stand up to evaluation? A first-principle approach to the evaluation of classifiers. arXiv preprint arXiv:2302.12006. <https://doi.org/10.48550/arXiv.2302.12006>
8. Gallet, B., & Gowanlock, M. (2022). Leveraging GPU tensor cores for double precision Euclidean distance calculations. arXiv preprint arXiv:2209.11287. <https://doi.org/10.48550/arXiv.2209.11287>
9. Guan, C., Suryanto, H., Mahidadia, A., Bain, M., & Compton, P. (2023). Responsible credit risk assessment with machine learning and knowledge acquisition. *Human-Centric Intelligent Systems*, 3(3), 232-243. <https://doi.org/10.1007/s44230-023-00035-1>
10. Klein, M., & Thompson, R. (2021). Advances in credit scoring models and their application in risk management. *Journal of Financial Risk Management*, 28(3), 135-149. <https://doi.org/10.1016/j.jfrm.2021.01.008>
11. Mhlanga, D. (2021). Financial inclusion in emerging economies: The application of machine learning and artificial intelligence in credit risk assessment. *International journal of financial studies*, 9(3), 39. <https://doi.org/10.3390/ijfs9030039>
12. Montevechi, A. A., de Carvalho Miranda, R., Medeiros, A. L., & Montevechi, J. A. B. (2024). Advancing credit risk modelling with Machine Learning: A comprehensive review of the state-of-the-art. *Engineering Applications of Artificial Intelligence*, 137, 109082. <https://doi.org/10.1016/j.engappai.2024.109082>
13. Niu, Z., & Zheng, Y. (2020). Machine learning for credit risk prediction: A systematic literature review. *Journal of Risk and Financial Management*, 8(11), 169. <https://doi.org/10.3390/jrfm8110169>
14. Pandas Development Team. (2023). pandas.get\_dummies — pandas 2.2.3 documentation. Retrieved from [https://pandas.pydata.org/docs/reference/api/pandas.get\\_dummies.html](https://pandas.pydata.org/docs/reference/api/pandas.get_dummies.html)
15. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830. <https://doi.org/10.5555/1953048.2078195>
16. Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: a systemic review. *Neural Computing and Applications*, 34(17), 14327-14339. <https://doi.org/10.1007/s00521-022-07472-2>
17. Siddique, A., Hasan, M. M., Mollah, M. B., & Rashid, M. (2022). Credit risk management and financial stability: Evidence from the global financial crisis. *Journal of Financial Stability*, 60, 100967. <https://doi.org/10.1016/j.jfs.2022.100967>
18. Smith, J. A., & Brown, L. R. (2020). A comprehensive guide to correlation coefficients: Pearson, Kendall, and Spearman. *Journal of Statistical Methods*, 34(2), 45-58. <https://doi.org/10.1016/j.jstatmeth.2020.01.005>
19. Suhadolnik, N., Ueyama, J., & Da Silva, S. (2023). Machine learning for enhanced credit risk assessment: An empirical approach. *Journal of Risk and Financial Management*, 16(12), 496. <https://doi.org/10.3390/jrfm16120496>
20. Thompson, E. R., & Lee, M. S. (2022). Advancements in descriptive statistics and graphical methods. *Journal of Statistical Analysis*, 58(2), 123-135. <https://doi.org/10.1016/j.jstatmeth.2022.03.005>
21. Trivedi, S. K. (2020). A study on credit scoring modeling with different feature selection and machine learning approaches. *Technology in Society*, 63, 101413. <https://doi.org/10.1016/j.techsoc.2020.101413>
22. Zhang, X., & Wang, Y. (2020). Feature scaling and its impact on machine learning performance: A study of standardization techniques. *Journal of Machine Learning Research*, 21(1), 122-135. <https://doi.org/10.1016/j.jmlr.2020.01.010>
23. Zhang, Y., Zhang, H., Wang, S., Zhang, X., & Zhao, H. (2020). Credit scoring using machine learning techniques. *IEEE Access*, 8, 23257–23272. <https://doi.org/10.1109/ACCESS.2020.3045685>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.