Forecasting credit risk using multinomial logistic regression

Mariya Nazarkevych[1,†], Rostyslav Yurynets[1,*,†] and Zoryna Yurynets[2,†]

[1] Department of Information Systems and Networks, Lviv Polytechnic National University, Lviv, Ukraine
[2] Ivan Franko Lviv National University, Lviv, Ukraine

Abstract

The article evaluates and forecasts credit risk. The proposed multinomial logistic regression model allows to evaluate and forecast the credit risk indicator of financial institutions taking into account various factors. The following factors were used in the study: average monthly income, sum of the credit, term of the credit, interest rate, age, expert's assessment of a professional, economic and social status of the client. The use of regression analysis methods allows to identify implicit relationships between elements of credit transactions and to solve many problems related to the analysis and forecasting of credit risk. Multinomial logistic regression allows financial institutions to classify credits into multiple classes, providing a more granular view of credit risk that better reflects the complexity of real-world lending outcomes.

Keywords

Machine Learning, Credit Risk, Multinomial Logistic Regression Model, Factors, Forecasting

1. Introduction

Predicting credit risk is a major challenge for financial institutions worldwide. As essential parts of financial development, banks are exposed to significant credit risk when customers do not repay their loans. The long-term viability of the banking sector relies on strong credit risk management systems that can accurately forecast default likelihoods and inform lending choices. The assessment of credit risk is a fundamental element of a financial institution's risk management strategy, helping to predict borrower defaults and reduce potential financial losses. While traditional methods have long depended on standard statistical models, the field has been revolutionized in recent years by the adoption of machine learning techniques.

Among these new methods, Multinomial Logistic Regression (MLR) has proven to be an especially effective tool for modeling credit risk. MLR is an extension of binary logistic regression, designed to handle cases with multiple distinct outcomes. This feature is highly beneficial in the field of credit risk modeling, as it enables financial institutions to categorize credits into more detailed risk levels rather than just a simple default or non-default classification. The mathematical basis of MLR allows for the calculation of probabilities across several classes, offering a more detailed perspective on credit risk that captures the complexities of real-world financial outcomes. MLR functions by modeling the connections between a set of predictor variables and a categorical outcome that has multiple classes. The model determines the probability that a given observation falls into each potential outcome category, ensuring that these probabilities add up to one. This multi-class functionality gives MLR a distinct advantage in credit risk situations that require sorting credits into various risk tiers.

★MoDaST 2025: Modern Data Science Technologies Doctoral Consortium, June, 15, 2025, Lviv, Ukraine

* Corresponding author.
† These authors contributed equally.
1 mariia.a.nazarkevych@lpnu.ua (M. Nazarkevych); rostyslav.v.yurynets@lpnu.ua (R. Yurynets); zoryna_yur@ukr.net (Z. Yurynets)
0000-0002-6528-9867 (M. Nazarkevych); 0000-0003-3231-8059 (R. Yurynets); 0000-0001-9027-2349 (Z. Yurynets)

2. Literature review

Several important studies have investigated the use of MLR in the context of credit risk modeling. A study by Arutjothi and Senthamarai compared the performance of a standard Multinomial Logistic Regression model against a tuned version for credit risk assessment and default classification [5]. Their analysis was based on a financial dataset from the UCI repository, which included 30,000 records centered on default prediction indicators. The research concluded that Multinomial Logistic Regression performs significantly better than other classifier models in predicting credit risk. Their results showed that a KM-MLR model achieved the highest performance, with both precision and accuracy reaching 82%. It was also noted that their tuned MLR model outperformed the standard version, showing a performance increase of about 2%. The researchers benchmarked MLR against other classification methods, such as the K-Nearest Neighbour (K-NN) classifier, standard Logistic Regression (LR), and Decision Tree (DT) classifiers, finding MLR to have superior predictive power for credit risk assessment.

The strong performance of MLR is due to several key characteristics. These include its capacity to manage multiple outcome categories at once, its relative ease of interpretation compared to "black box" machine learning models, and its ability to effectively model the intricate relationships between predictor variables and credit outcomes. These features make MLR an excellent choice for credit risk modeling, where both predictive accuracy and model transparency are highly valued.

Further contributions come from Adha, Nurrohmah, and Abdullah, who used a multinomial logistic regression model to identify factors influencing default and attrition events in credit scenarios [35]. They used credit card datasets to pinpoint which factors affected different credit outcomes and also proposed spline regression with a truncated power basis as a complementary method to MLR. Similarly, Motedayen et al. conducted a study to identify and assess factors that impact credit risk management by applying the multinomial logistic regression model [34]. Their research helps clarify which variables are most influential on credit risk across various categories.

Studies using MLR have identified several variables with statistical significance in forecasting credit risk. Although not all sources provide extensive details, the literature indicates that variables such as credit value, age, past credit history, occupation, the term of credit repayment, number and amount of installments, history of credit extensions, collateral type, average account balance, facility interest rate, type of facility, and education level can all have a significant effect on the credit risk assessment for individual borrowers.

In addition to traditional financial data, there is a growing interest in using alternative data sources to improve predictive accuracy in credit risk modeling. Such sources might include social media activity, utility payment records, and other non-traditional indicators of a person's creditworthiness. The use of this alternative data is a developing area in credit risk modeling, and the flexibility of MLR makes it a suitable method for incorporating these varied data types.

Despite its benefits, using MLR for credit risk modeling comes with certain challenges and limitations that must be managed. Common issues in credit risk modeling that also affect MLR include data imbalance, feature selection, model interpretability, and computational efficiency. Model interpretability can be seen as both a strength and a weakness of MLR. While it is more transparent than many "black box" methods, its application in credit decisions affecting consumers requires clear explanation. The coefficients in an MLR model can be understood to see the link between predictor variables and the probabilities of outcomes, but the multi-class nature of MLR adds a layer of complexity not present in binary logistic regression.

Furthermore, validation frameworks may need to be updated to manage the complexities of machine learning applications in credit risk modeling [13]. As MLR and similar techniques become more widespread, having strong validation processes is crucial for meeting regulatory standards and for effective risk management.

When measured against more advanced models like Random Forest and LightGBM, MLR remains a strong contender in certain situations. One comparative study showed that while advanced algorithms often deliver powerful predictions, MLR is still competitive, particularly when dealing with clearly defined categorical outcomes [11]. Research on "Buy Now Pay Later" customers has demonstrated MLR's flexibility in evaluating credit risk for various financial products. That study compared MLR's performance to both standard and advanced models, confirming its continued relevance in modern credit risk assessment.

Key strengths of MLR include:

- Multi-Class Classification: Unlike binary logistic regression, MLR can assign borrowers to several categories, such as defaulted, non-defaulted, or late payments, which is valuable for detailed credit assessments.
- Interpretability: A significant advantage of MLR is its interpretability. The model's coefficients are easily interpreted as odds ratios, showing how a change in a predictor variable impacts the likelihood of each outcome. This transparency is vital for financial institution stakeholders who rely on model outputs for lending decisions.
- Probability Estimates: MLR calculates probability estimates for each class, which helps lenders see not only the most probable category for a borrower but also the level of uncertainty in that prediction. This capability is essential for risk management, as it allows institutions to better evaluate potential risks.
- Computational Efficiency: MLR algorithms are generally less demanding on resources than more complex models like deep learning networks. This is an advantage in settings with limited or expensive computational power. For example, one implementation of MLR on a large, encrypted dataset took about 17 hours on a single machine, proving it can produce results without requiring excessive computing resources.

Practical applications of MLR include:

- Risk Segmentation: MLR is effective for dividing borrowers into different risk segments. A study by Anderson [4] used MLR to sort borrowers into low, medium, and high-risk groups based on their financial behavior and credit history, achieving high accuracy and enabling more sophisticated credit decisions.
- Default Prediction for SMEs: Smith and Johnson [38] applied MLR to forecast the probability of default for small and medium enterprises (SMEs), classifying them as "no default," "partial default," or "full default," which demonstrated MLR's ability to handle complex default behaviors.
- Portfolio Management: In portfolio management, MLR is used to evaluate risk distribution across various asset classes. A study by Brown et al. [8] utilized MLR to assess the risk profiles of different credit portfolios, which helped improve risk diversification and allocation strategies.
- Granular Risk Analysis: MLR allows for detailed risk segmentation, which is important for setting prices and determining capital reserves. Chen et al. [10] used MLR to forecast S&P ratings for companies by incorporating liquidity ratios, leverage, and industry volatility, reaching an accuracy of 78% and outperforming ordered logit models in cases where ratings did not have a natural order.
- Use of Alternative Data: Martinez & Ruiz categorized personal credit applicants into "low," "medium," and "high" risk groups using MLR combined with alternative data like rent payment history and utility bills.

Recent progress in MLR has involved integrating it with other machine learning techniques to boost predictive power. For instance, hybrid models that merge MLR with decision trees or neural networks have shown potential for increasing accuracy and robustness in credit risk modeling. The application of regularization techniques such as LASSO and Ridge regression has also been investigated to tackle multicollinearity and enhance model generalization.

3. Methodology

For credit risk assessment, a relationship should be established between factors and probability size of credit risk. We will take the credit risk indicator according to three values (satisfy, partially satisfy and not satisfy). Therefore, it is necessary to forecast the credit risk indicator a model of multinomial logistic regression. Multinomial logistic regression is used to forecast the probability of an event by the values of a set of features.

In order to predict the future state of credit risk, it is important to model it. This implies using several quantitative factors (behavioral factors) to estimate a qualitative variable (intellectual factors). You can use tools that help you choose the best numbers to measure things in a way that discriminates between different groups. Multinomial logistic regression enables the identification of the group of credit risk. Furthermore, multinomial logistic regression enables the consideration of the likelihood that a risk will be categorized as a specific group risk.

All multiple-choice options are assigned numbers in a random sequence ranging from 0, 1, 2, ..., K [39]. The likelihood of any given option occurring is determined using a polynomial logit model.

$$P(z_i = k) = \frac{e^{x_i q_k}}{\sum_{k=0}^{K} e^{x_i q_k}} \tag{1}$$

where q_k are unknown parameters; x_1 is individual creditworthiness ratio; x_2 is age; x_3 is expert assessment of the profession and the socio-economic status of the client.

The following notations are presented here [12]:

$$\begin{gathered} \boldsymbol{q} = (q_0, q_1, \ldots, q_s)^T, \\ \boldsymbol{x}_i = (1, x_{i1}, \ldots, x_{is}), \\ \boldsymbol{x}_i\boldsymbol{q} = q_0 + q_1 x_{i1} + q_2 x_{i1} + \ldots + q_s x_{is} \end{gathered} \tag{2}$$

To identify model (1), it is common to rely on rationing. $q_K = 0$ [16]. Then

$$P(z_i = k) = \frac{e^{x_i q_k}}{1 + \sum_{k=0}^{K-1} e^{x_i q_k}} \quad j = 0, 2, \ldots, K-1, \tag{3}$$

$$P(z_i = K) = \frac{1}{1 + \sum_{k=0}^{K-1} e^{x_i q_k}} \tag{4}$$

One aspect of constructing a polynomial logit model is this specific feature, where only the coefficients for the first K dependencies, denoted as q_0, q_1, ... , q_{K-1}, are computed. This process aligns with formula (3), corresponding to the initial K probabilities: $P(z_i = 0)$, $P(z_i = 1)$, ... , $P(z_i = K-1)$. The probability of selecting the final option, $P(z_i = K)$, is not directly calculated but is instead determined separately based on formula (4).

Based on the entered variable u_{ij}, we formulate the logarithmic likelihood function[18, 35]

$$\ln G = \sum_{i=1}^{n} \sum_{j=0}^{K} u_{ij} \ln \left(\frac{e^{x_i q_k}}{\sum_{k=0}^{J} e^{x_i q_k}} \right) \tag{5}$$

By differentiating expression (5) with respect to $\boldsymbol{q}_j$, we derive a set of equations that characterize the maximum likelihood estimation system

$$\frac{\partial \ln G}{\partial q_j} = \sum_{i=1}^{n} \sum_{j=0}^{K} u_{ij} \frac{\partial}{\partial q_j} \ln \left(\frac{e^{x_i q_k}}{\sum_{k=0}^{J} e^{x_i q_k}} \right) =$$

$$=\sum_{i=1}^{n}\sum_{k=0}^{K}u_{ij}\frac{\sum_{s=0}^{K}e^{x_i q_s}}{e^{x_i q_k}}\left(\frac{e^{x_i q_k}\sum_{s=0}^{J}e^{x_i q_s}-e^{x_i q_k}e^{x_i q_k}}{\left(\sum_{s=0}^{J}e^{x_i q_s}\right)^2}\right)x_i^{'}= \qquad (6)$$

$$=\sum_{i=1}^{n}u_{ik}\left(1-\frac{e^{x_i q_k}}{\sum_{s=0}^{K}e^{x_i q_s}}\right)x_i^{'}=0, \quad k = 0, 1, 2, \ldots, K-1.$$

The solution for this system, assuming $\boldsymbol{q}_k = 0$, is determined numerically using the Newton-Raphson method [31]. The computational procedure is organized in such a way that the model values corresponding to the final alternative are set to zero. In practical terms, if $\boldsymbol{q}_0 = 0$ is required instead of $\boldsymbol{q}_k$, the data related to the alternative with $k = 0$ must be input last.

The Newton-Raphson method typically necessitates multiple iterative steps to achieve convergence [37].

$$q_{t+1}=q_t-\frac{\partial \ln G(q_t)}{\partial q}\left[\frac{\partial^2 \ln G(q_t)}{\partial q \partial q'}\right]^{-1} \qquad (7)$$

The implementation of the Newton-Raphson method necessitates the computation of a matrix consisting of second-order partial derivatives:

$$\frac{\partial^2 \ln G}{\partial q_j \partial q_i^{'}}=\frac{\partial}{\partial q_i^{'}}\sum_{i=1}^{n}\left(u_{ij}-\frac{e^{x_i q_k}}{\sum_{s=0}^{K}e^{x_i q_s}}\right)x_i^{'}=$$

$$=-\sum_{i=0}^{n}\left(\frac{e^{x_i q_k}\sum_{s=0}^{K}e^{x_i q_s}-e^{x_i q_k}e^{x_i q_k}}{\left(\sum_{k=0}^{K}e^{x_i b_k}\right)^2}\right)x_i^{'}x_i= \qquad (8)$$

$$=-\sum_{i=0}^{n}\left[\frac{e^{x_i q_k}}{\sum_{s=0}^{K}e^{x_i q_s}}\left(E(k=l)-\frac{e^{x_i q_k}}{\sum_{s=0}^{K}e^{x_i q_s}}\right)\right]x_i^{'}x_i$$

In the resulting expression, $\boldsymbol{E}(k = l)$ equals 1 when k is equal to l and 0 otherwise.

4. Empirical results

The evaluation of client solvency by the model of multinomial logistic regression in the presented paper required the statement of correlation between risk factors and probability size of credit risk. Table 1 presents a matrix fragment of values of nine indexes which were selected during the analysis in which a binary variable y describes the following situations: 0 is no credit provided, 1 is credit provided, 2 is partially credit provided.

Table 1 introduces designations: d – average monthly income, s – the sum of the credit, t – the term of the credit, r – an interest rate, B – the age, O – expert’s assessment of a professional, economic and social status of the client.

Table 1.
Matrix fragment of values indexes for the evaluation of client solvency

No.	*z*	*d*, UAH	*s*, UAH	*t*, months	*r*	*B*, years	*O*
11	1	9900	25000	12	10	48	4
12	1	7000	21000	28	12	39	4
13	0	10200	37500	18	11	63	4
14	1	17600	60000	42	14	55	5
15	0	15000	50000	14	11	68	5
16	0	10600	35000	24	12	56	2
17	2	14300	45000	18	11	37	2
18	2	11800	33000	24	12	44	3

The calculation of solvency coefficient of the ownership is performed by a formula:

$$x_1 = \frac{d}{s\left(\frac{1}{t} + \frac{r}{12 \cdot 100}\right)}$$

The prepared dataset, partially outlined in Table 2, is utilized to estimate the unknown parameters of the multinomial logistic regression model (1). In this context, x_1 represents the individual's creditworthiness ratio, x_2 corresponds to the age of the individual, and x_3 denotes the expert evaluation of the client’s profession and socio-economic status.

We explore the relationship between the dependent variable z and the factors x_1, x_2, and x_3 by developing a multinomial logistic regression model. The parameters of this model are detailed in table 3.

Table 2
Matrix fragment of the value indicators

z	x_1	x_2	x_3
1	4,32	48	4
1	7,29	39	4
0	4,20	63	4
1	8,27	55	5
0	3,72	68	5
0	5,86	56	2
2	4,91	37	2
2	6,92	44	3

Table 3
Parameters of the multinomial logit model

	Level of - Response	Column	Estimate	Standard - Error	Wald - Stat.	p
Intercept 1	0	1	5,5911	4,65130	1,706624	0,180567
x_1	0	2	-7,8112	6,15515	1,791649	0,171782
x_2	0	3	1,6523	1,37282	1,684877	0,186222
x_3	0	4	-17,6863	14,74786	1,868580	0,190125
Intercept 2	1	5	-26,0048	18,05363	1,874813	0,149748
x_1	1	6	1,9130	1,52847	1,646393	0,200732
x_2	1	7	-0,1714	0,12857	1,777891	0,182409
x_3	1	8	6,0582	4,06256	2,223729	0,135904
Scale			1,0000	0,00000		

The obtained coefficient estimates can be considered statistically significant, as the standard errors are all smaller than their corresponding estimates, the values of Wald's statistics exceed the critical threshold, and all error probabilities remain below 0.2.

Now, let us formulate the analytical expression for the developed multinomial logit model:

$$P(z=0)=\frac{e^{5,59+(-7,8)x_1+(1,6)x_2+(-17,69)x_3}}{1+e^{5,59+(-7,8)x_1+(1,6)x_2+(-17,69)x_3}+e^{-26+(1,9)x_1+(-0,17)x_2+(6,06)x_3}}$$

$$P(z=1)=\frac{e^{-26+(1,9)x_1+(-0,17)x_2+(6,06)x_3}}{1+e^{5,59+(-7,8)x_1+(1,6)x_2+(-17,69)x_3}+e^{-26+(1,9)x_1+(-0,17)x_2+(6,06)x_3}}$$

$$P(z=2)=\frac{1}{1+e^{5,59+(-7,8)x_1+(1,6)x_2+(-17,69)x_3}+e^{-26+(1,9)x_1+(-0,17)x_2+(6,06)x_3}}$$

The derived expression can be utilized to estimate the likelihood of credit risk based on varying factor values.

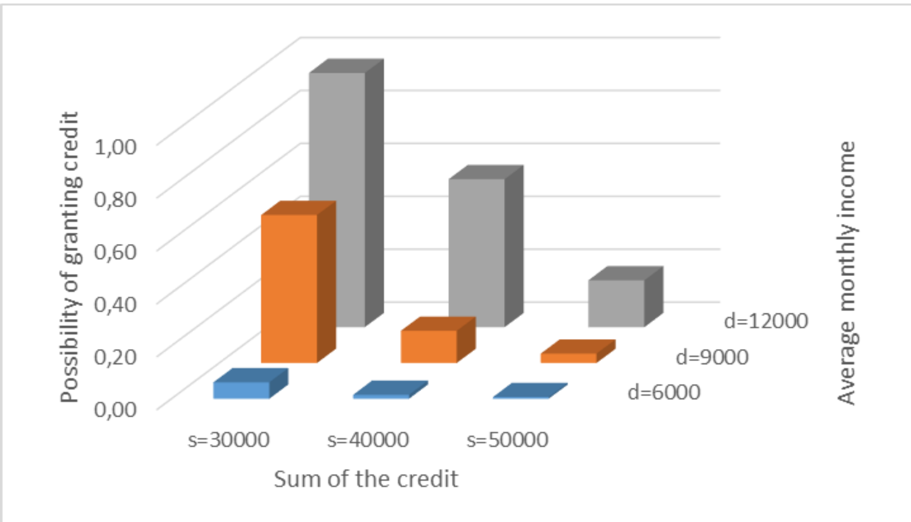

Figure 1: Possibility of granting credit depending on the average monthly income and sum of the credit

In particular, Fig. 1 and 2 show the dependence of the credit risk indicator (probability of granting credit to customers) on the change in the factor d (average monthly income) for certain values of the factor s (sum of the credit) and fixed values of the factors t, r, B, O (t=18, r=11, B=40, O=4).

Figures 1 and 2 show a decrease in the probability of granting credit and an increase in the probability of partial granting credit provided that the amount of the credit increases.

When a larger credit amount is requested, the credit risk for a bank increases because of the greater potential loss from a borrower default. This results in a lower probability that the bank will grant the full requested amount. To mitigate this risk, banks are more likely to approve a partial credit amount instead. This strategy allows them to limit their financial exposure while still partially meeting the borrower's needs.

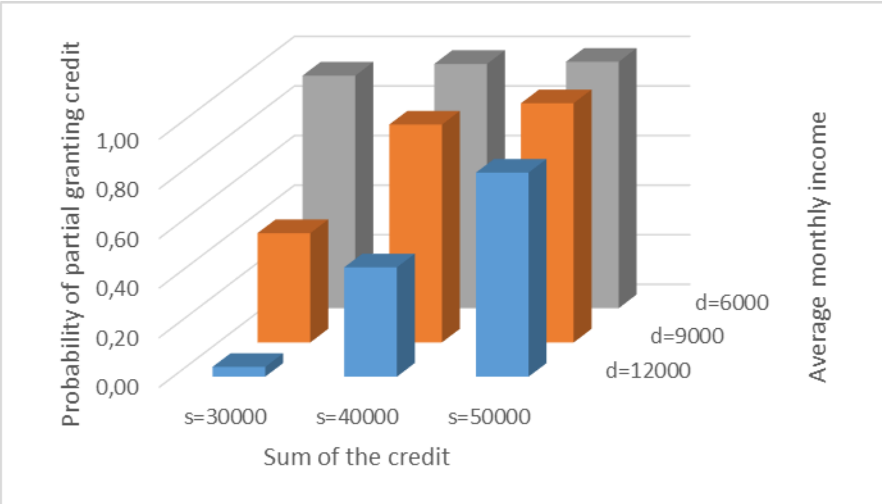

Figure 2: Probability of partial granting credit depending on the average monthly income and sum of the credit

Figures 3 and 4 show the dependence of the credit risk indicator on the change in the factor s (sum of the credit) for certain values of the factor r (interest rate) and a fixed value of the factors d, t, B, O (t=18, d=9000, B=52, O=3).

Figures 3 and 4 show an increase in the probability of granting a credit and a decrease in the probability of partial granting a credit due to an increase in the interest rate.

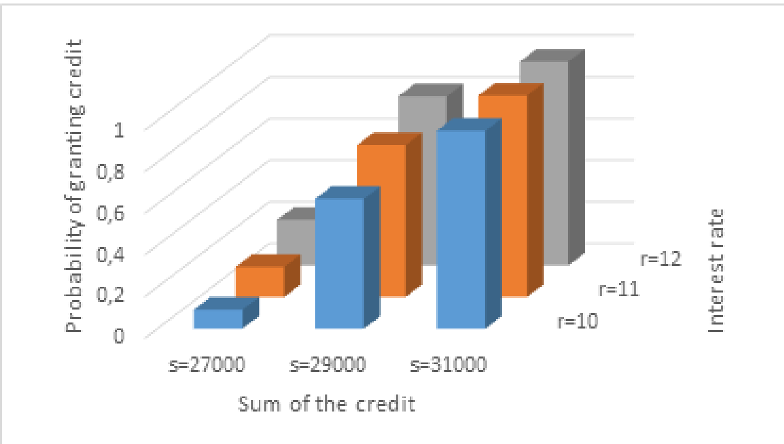

Figure 3: Possibility of granting credit depending on the sum of the credit and interest rate

A higher interest rate on a loan increases the cost for the borrower, making the credit less appealing and more expensive to service. However, from the bank's perspective, a higher rate can increase the likelihood of the credit being granted, as the increased rate compensates for the elevated risk and generates more profit for the institution. Conversely, a higher interest rate tends to lower the probability of a partial credit grant. This is because the bank aims to avoid the extra operational costs and complexities of partial lending while maximizing profit from the full loan amount. A higher rate enables the bank to earn more income even from a fully granted credit, which reduces the incentive for partial disbursements.

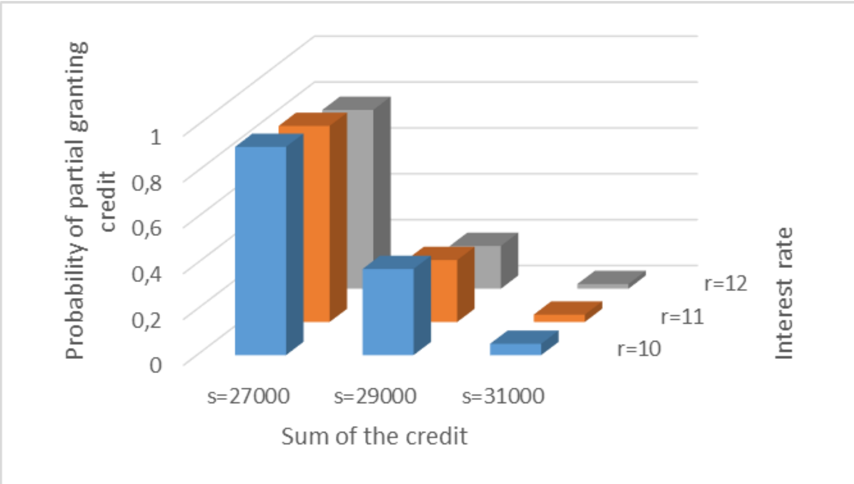

Figure 4: Probability of partial granting credit depending on the sum of the credit and interest rate

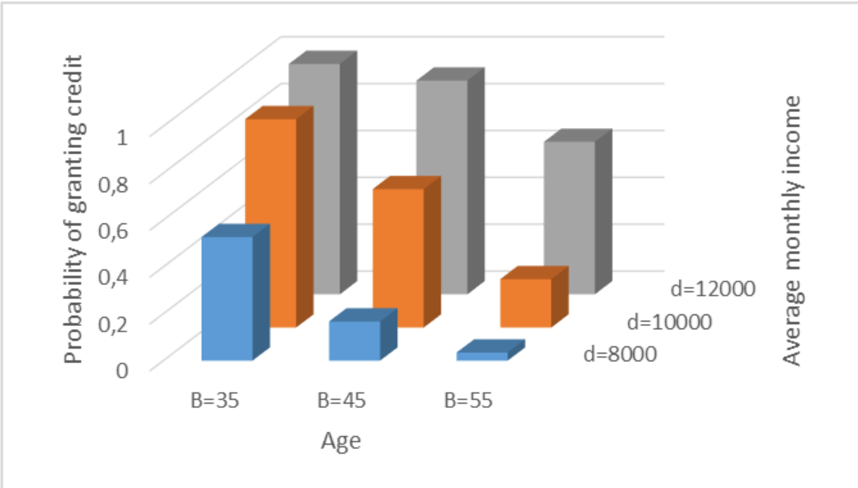

Figure 5: Possibility of granting credit depending on the age and average monthly income

Figures 5 and 6 show the dependence of the credit risk indicator on the change in the factor d (average monthly income) for certain values of the factor B (age) and a fixed value of the factors s, t, r, O (t=18, s=30000, r=11, O=4).

Figures 5 and 6 show a decrease in the probability of granting a credit and an increase in the probability of partial granting a credit due to increasing age. This is due to the fact that lenders consider age as a risk factor, especially when the candidate's age on the date of full fulfillment of obligations under the credit agreement exceeds a certain threshold.

Figures 7 and 8 show the dependence of the credit risk indicator upon changing the factor s (sum of the credit) for certain values of the factor t (term of the credit) and fixed values of the factors d, r, B, O (O=4, d=10000, r=11, B=40).

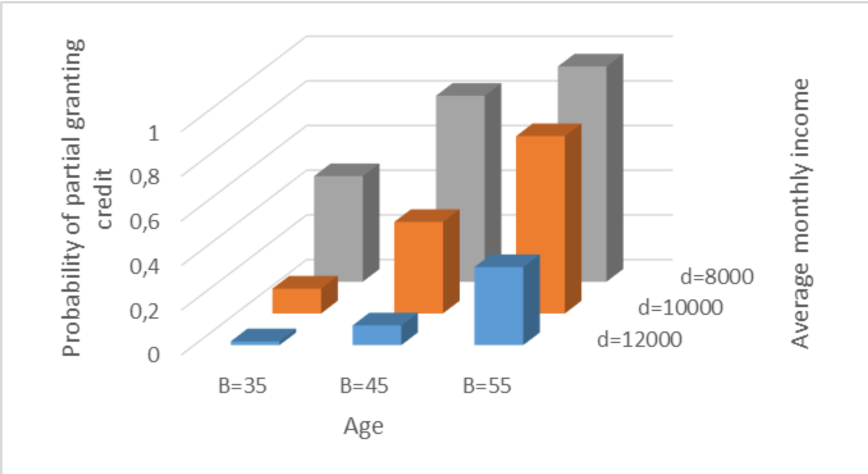

Figure 6: Probability of partial granting credit depending on the age and average monthly income

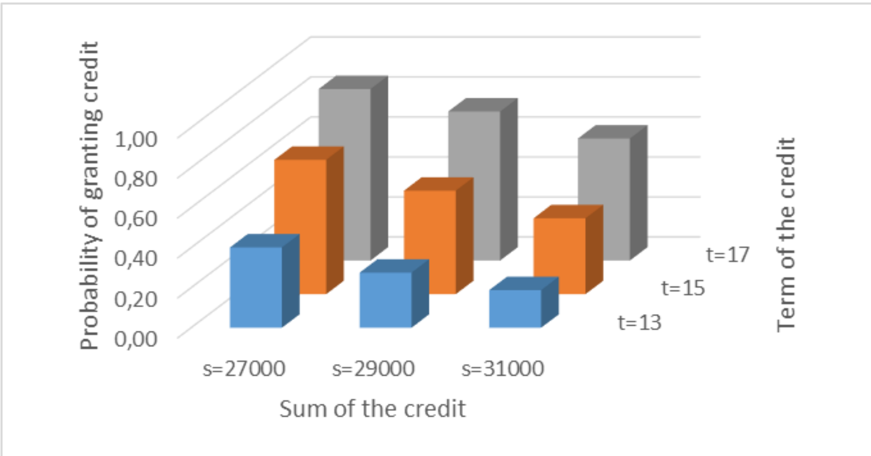

Figure 7: Possibility of granting credit depending on the sum of the credit and term of the credit

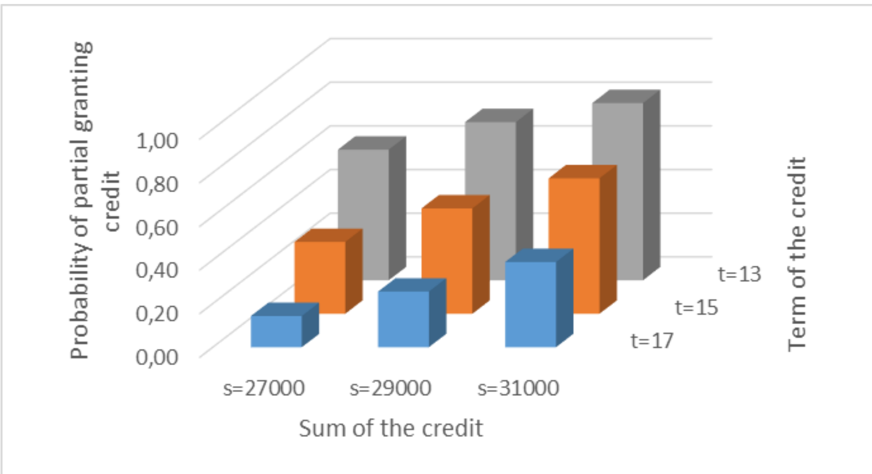

Figure 8: Probability of partial granting credit depending on the sum of the credit and term of the credit

Now let's estimate the probabilities of granting credits to new customers using the developed model. The information for estimating the granting of credits is presented in Table 4.

Table 4
Value of indexes for assessing the solvency of new clients

No.	*d*, UAH	*s*, UAH	*t*, months	*r*	*B*, years	*O*	*P*(*z*=0)	*P*(*z*=1)	*P*(*z*=2)
1	11000	38000	18	11	61	4	0,66	0,01	0,33
2	12000	23000	16	10	35	3	0,00	0,56	0,44
3	14000	62000	24	12	52	4	0,00	0,09	0,91

The calculations show that it is possible to provide credit only to the second and third clients.

5. Conclusions

Multinomial Logistic Regression offers significant advantages for credit risk modeling by providing a strong balance of predictive accuracy, transparency, and the ability to classify outcomes into multiple categories. The reviewed literature shows that MLR performs better than several other methods in predicting credit risk. These performance benefits, along with its capacity for multi-class classification, make it an extremely useful tool for financial institutions looking to improve their credit risk assessment processes.

To successfully apply MLR for credit risk modeling, it is essential to pay close attention to data preprocessing, parameter tuning, and the selection of variables. As the field of credit risk assessment continues to advance through the use of alternative data and hybrid models, MLR will likely continue to be a key part of the credit risk modeling toolkit. It can be used either as a standalone model or as a component in ensemble methods that utilize its strengths while compensating for its weaknesses.

Future research should focus on combining MLR with other complementary techniques, applying it to new alternative data sources, and developing methods to overcome challenges like data imbalance and complex validation needs. Such advancements will make MLR even more useful in credit risk modeling and will help establish more effective risk management practices in financial institutions across the globe. In summary, multinomial logistic regression is a highly relevant tool for credit risk modeling due to its flexibility in managing multi-class situations, its interpretability, and its ability to adapt to ordinal ratings. Research consistently confirms its value in real-world banking risk assessments, positioning it as a core component of modern credit risk strategies.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

Referenses

[1] A Data Science Case Study with R and mlr. Alpha-epsilon, 2018. http://www.alpha-epsilon.de/r/2018/07/13/a-data-science-case-study-with-r-and-mlr.

[2] Agarwal A., et al. Fairness in Credit Scoring: A Multinomial Logistic Approach. Journal of Responsible AI, 2023.

[3] Alam T., & Hao Y. Review of the literature on credit risk modeling: Emphasis on machine learning techniques. Semantic Scholar, 2019. https://www.semanticscholar.org/paper/Review-of-the-literature-on-credit-risk-modeling:-Alam-Hao/5a2ebab358547218215b3e47763ab6b5ff3b6794

[4] Anderson J. Multinomial Logistic Regression in Credit Risk Assessment, 2018. Journal of Financial Risk Management.
[5] Arutjothi G., & Senthamarai C. Efficient Analysis of Financial Risks Using Multinomial Logistic Regression. Iconic Research and Engineering Journals, 8(6), 893-899, 2024.
[6] Bae S., et al. Survival MLR for Default Timing Prediction. Journal of Banking & Finance, 2023.
[7] Brown T., et al. Rating Transitions in Mortgage Portfolios. Financial Analysts Journal, 2021.
[8] Brown T., et al. (2019). "Risk Diversification in Loan Portfolios: A Multinomial Logistic Approach." Financial Analysts Journal.
[9] Credit Rate by Ordinal Multinormal Regression. MathWorks. Retrieved from https://www.mathworks.com/help/risk/credit-rate-by-ordinal-multinormal-regression.html
[10] Chen Y., et al. Predicting Corporate Ratings with MLR. Journal of Credit Risk, 2019.
[11] Caner Tas, et al. Comparison of Machine Learning and Standard Credit Risk Models Performances in Credit Risk Scoring of Buy Now Pay Later Customers. Middle East Technical University, 2023. https://open.metu.edu.tr/bitstream/handle/11511/104525/Comparison_of_Machine_Learning_and_Standard_Credit_Risk_Models__Performances_in_Credit_Risk_Scoring_of_Buy_Now_Pay_Later_Customers.pdf
[12] C. Ryan, Categorical Data Analysis. In: Data Science with R for Psychologists and Healthcare Professionals, CRC Press, 2021, available at: https://doi.org/10.1201/9781003106845-22 (accessed 21 January 2022)
[13] Deepa Shukla, Sunil Gupta. Machine Learning in Credit Risk Assessment: A Comprehensive Review. Journal of Economics and Statistics, 4(3), 1-7, 2024. https://journal.esrgroups.org/jes/article/view/6788
[14] Kyoohyung Han, et al. Efficient Logistic Regression on Large Encrypted Data. Cryptology ePrint Archive, 2018. https://eprint.iacr.org/2018/662.pdf
[15] Feng Y., et al. Dynamic MLR for Economic Cycles. International Journal of Forecasting, 2022.
[16] G. James, D. Witten, T. Hastie, R. Tibshirani. Statistical Learning. In: An Introduction to Statistical Learning, Springer Science+Business Media, LLC, 2021, pp. 15–57, available at: https://doi.org/10.1007/978-1-0716-1418-1_2 (accessed 21 January 2022)
[17] Garcia M., et al. Hybrid Models in Peer-to-Peer Lending. FinTech Journal, 2022.
[18] J. H. Kim, Multicollinearity and misleading statistical results, Korean journal of anesthesiology, 72(6), 2019, pp. 558-569.
[19] Joseph L. Breeden. A survey of machine learning in credit risk. Journal of Credit Risk, 17(3), 1-62, 2021. https://www.risk.net/journal-of-credit-risk/7878976/a-survey-of-machine-learning-in-credit-risk
[20] K. B. Salem, A. B. Abdelaziz. Conduct a multivariate analysis by logistic regression, La Tunisie médicale, 98(7), 2020, pp. 537-542
[21] Kim S., et al. Regularized MLR for Transaction Data. Machine Learning in Finance, 2023.
[22] Lee H., et al. (2021). "Hybrid Models for Credit Risk Prediction: Integrating MLR with Machine Learning." Computational Finance and Economics.
[23] Luka Vidovic, Lei Yue. Machine Learning and Credit Risk Modelling. S&P Global Market Intelligence, 2020. https://www.spglobal.com/marketintelligence/en/documents/machine_learning_and_credit_risk_modelling_november_2020.pdf
[24] Logistic Regression: Overview and Applications. Keylabs AI Blog. Keylabs AI, 2024 https://keylabs.ai/blog/logistic-regression-overview-and-applications/
[25] Lee H., et al. Modeling Credit Card Delinquency. Computational Finance and Economics, 2022.
[26] Lofstrom T. Credit Scoring with Machine Learning, 2022. http://kth.diva-portal.org/smash/get/diva2:1833653/FULLTEXT01.pdf
[27] Luana Cristina Rogojan, Andreea Elena Croicu and Laura Andreea Iancu. Modern Approaches in Credit Risk Modeling: A Literature Review. Proceedings of the International Conference on Business Excellence, 17(1), 676-685, 2023. https://sciendo.com/article/10.2478/picbe-2023-0145

[28] Malan R. J. The prediction of credit risk using data mining techniques (Doctoral dissertation). University of Pretoria, 2015. https://repository.up.ac.za/bitstream/2263/27333/1/dissertation.pdf
[29] M. Pagano, K. Gauvreau, H. Mattie. Logistic Regression. In: Principles of Biostatistics, Chapman and Hall/CRC, 2022
[30] M. Yilmaz, Multinomial logit model, 2018, available at: https://www.researchgate.net/publication/335336335_MULTINOMIAL_LOGIT_MODEL (accessed 21 January 2022)
[31] M. Abonazel, R. A. Farghali, Liu-Type Multinomial Logistic Estimator, Sankhya B: The Indian Journal of Statistics, 81(2/2), 2019, pp. 203-225.
[32] M R Adha, S Nurrohmah and S Abdullah. Multinomial Logistic Regression and Spline Regression for Credit Risk Modelling. Journal of Physics: Conference Series, 2018. https://www.researchgate.net/publication/329379730_Multinomial_Logistic_Regression_and_Spline_Regression_for_Credit_Risk_Modelling
[33] Michiko Miyamoto. The Comparisons of Data Mining Techniques for Predicting Credit Scoring. International Journal of Finance and Accounting, 3(5), 284-292, 2014. http://article.sapub.org/10.5923.j.ijfa.20140305.07.html
[34] Motedayen Nasrin, et al. Identification and Evaluation of the Factors Affecting Credit Risk Management Using the Multinomial Logistic Regression Model. Iranian Economic Review 2022, 26(3). https://doi.org/10.22059/ier.2022.89098
[35] M. R. Abonazel, R. A. Farghali, Liu-type multinomial logistic estimator, Sankhya B, 81(2), 2019, pp. 203-225.
[36] Nurrohmah A. Multinomial Logistic Regression and Spline for Risk. Semantic Scholar, 2023. https://www.semanticscholar.org/paper/Multinomial-Logistic-Regression-and-Spline-for-Risk-Adha-Nurrohmah/99d9c0ed9b3960ec43a18724b546ca469861b38e
[37] Yurynets R. Risk Assessment Technology of Crediting with the Use of Logistic Regression Model / Rostyslav Yurynets, Zoryna Yurynets, Dmytro Dosyn, Yaroslav Kis // Computational Linguistics and Intelligent Systems. Proceedings of the 3rd International Conference on Computational Linguistics and Intelligent Systems (COLINS-2019). Volume I. – Kharkiv, Ukraine, April 18-19, 2019. – P. 153-162.
[38] Smith R., & Johnson L. (2020). "Predicting SME Defaults Using Multinomial Logistic Regression." International Journal of Banking and Finance.
[39] T. Q. Asenso, H. Zhang, Y. Liang, Pliable lasso for the multinomial logistic regression, Communication in Statistics - Theory and Methods, 2020, available at: https://doi.org/10.1080/03610926.2020.1800041 (accessed 21 January 2022)