UDC 336.717:004.78]:338.27
JEL classification: C38, C45, C58, G21, G51
DOI: https://doi.org/10.17721/1728-2667.2025/226-1/8

Danylo KRASOVYTSKYI, PhD Student
ORCID ID: 0009-0008-7017-0175
e-mail: daniel.krass2112@knu.ua
Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

DEVELOPMENT OF RETAIL LOAN SCORECARD USING MACHINE LEARNING

Background. ***Banks use credit scoring to track loan performance, manage provisions, and adjust lending policies. This method assigns points based on a borrower's loan history and unique characteristics, enabling lenders to predict default risk and improve credit conditions for low-risk borrowers. With increased data access and computing power, it is possible to do credit scoring via new methods with possibly better predictive power. This study aims to develop a scorecard for Ukrainian retail borrowers using Credit Registry data, exploring the effectiveness of logistic regression and Support Vector Machine (SVM) methods. Key research questions address the potential for binning data to improve model interpretability, the accuracy probability of default estimates, and differences in decision thresholds across scorecards built with logistic and SVM models.***

Methods. ***The study applies Weight of Evidence (WoE) binning, a technique that transforms variables to establish a monotonic relationship with default risk, thereby improving interpretability and model robustness. Using this binned data, the paper constructs scorecards with logistic regression and SVM. Each scorecard uses predictor variables such as Debt Service to Income (DSTI), age, interest rates, and days overdue to assess default likelihood. Scores are assigned based on each variable's impact on default probability.***

Results. ***Findings indicate that it is possible to develop a scorecard based on Credit Registry data. Logistic regression and SVM models yield similar score distributions, with high predictive accuracy and robustness as measured by accuracy and F1-score. The scorecard approach provides transparency and interpretability; for instance, borrowers with a DSTI exceeding 40% receive lower scores, indicating higher risk.***

Conclusions. ***Banks may use both logistic and SVM models for real-time credit assessments, leveraging accessible borrower characteristics to streamline decision-making. For regulators, the scorecards can support policy frameworks that restrict lending based on borrower risk bins, thus mitigating risks arising from specific retail lending segments.***

Keywords: ***machine learning, scorecard, default prediction, logistic regression, support vector machine.***

Background

Banks always keep track of how their loans perform. Based on this information they do provisioning and adjust their lending policies. One of the methods used to assess loan performance is credit scoring. It is utilized not only during the loan issuance but also during its lifetime. The idea behind credit scoring is simple: based on all loans the borrower has (their type, sum, maturity, etc.) and borrower-unique characteristics (like age) the person is awarded points. The more points are awarded the better credit conditions would be and the person will default less likely.

According to the World Bank report (The World Bank Group, 2019), credit scoring applications have become popular lately thanks to better data access and increased computing power. It now extends beyond approving credit applications, including but not limited to pricing financial services, setting credit limits, determining capital requirements by regulators, and supporting customer management. Weston and Hinson (2023) suggest that retail lenders should care about their credit rating since it directly affects their lifetime savings. For instance, an excellent credit score may save more than 80 thousand US dollars for a 30-year mortgage.

The paper aims *to develop a scorecard for Ukrainian retail loans using data from the Credit registry*. We employ logistic regression and the Support Vector Machine (SVM) method to understand whether modern machine learning (ML) methods can outperform traditional logit model.

Thus, our research questions are:

- Can we split Credit registry data into bins to enhance model interpretability and simplicity?
- Can we develop a scorecard that will show the client's credit quality in terms of default probability using logistic regression and SVM model?
- Will the threshold values for credit decisions differ between scorecards built upon the logistic and SVM model?

Literature review

Scorecard is a widely used method in credit risk assessment. For example, in the United States credit scores are important since they not only improve someone's chances of getting a loan, but also, they have a direct reflection in terms of these loans. Credit scores are also useful for banks since they show people's relative creditworthiness and can show the differences based on different characteristics.

Wu and Pan (2021) compared logistic regression, SVM, and random forest methods in predicting defaults. The logit model showed the best performance and thus was chosen for scorecard development. The authors used the weight of evidence (WoE) binning before building a model. In the end, they chose a score of 510 as a threshold since it divides data in half and made symmetrically three classes to the right and three classes to the left of this point. The calibration of parameters in the scoring function allowed receiving the range of scores 300–850, which is the range for the famous US FICO scoring system.

Lee et al. (2021) compared the results of binning using WoE and decision tree. The key feature of WoE is that for simplicity purposes it allows to do monotonic binning. A decision tree is a complex method that can cause a situation where in the last step the proportion of defaults will be the opposite of the bin logic. In WoE binning the decision boundaries can be adjusted manually which allows for having a monotone relationship between default proportion and variable of interest. The authors in the end suggest using WoE as a main binning algorithm.

Dong et al. (2012) opposed using ML methods because their outputs are difficult to interpret. Instead, the authors suggest using logistic regression with random coefficients to develop a scorecard. This complex approach led to an increase in the overall accuracy of prediction from 71 % to 74%.

Wang et al. (2020) prove that using ML methods such as XGBoost is the opposing way to building a scorecard. In their opinion, since scorecards have a straightforward explanation, the same is achieved for the XGBoost model by using SHAP values.

When the scorecard is built, borrowers with characteristics that are not that much different will be collected in the bins. Another way of doing so is using unsupervised ML methods such as clustering to find people the characteristics close to each other. For example, the application of K-nearest neighbors methods will result in using cluster numbers as independent variables in the model. Machado and Karray (2022) exploited this approach, which resulted in better performance characteristics.

Du Pisanie et al. (2023) suggested that scorecards might be very sensitive to changes in population characteristics. They proposed the simulation technique to test the stability of the scorecards' predictions. The Population stability index allowed authors to check how far the average probability of default shifts with a shift in the distribution of one variable.

Idbenjra et al. (2024) suggest that credit scorecards might be improved by using segmentation-based modelling. They propose estimating logit leaf model. It finds segments of borrowers, based on common features, estimates separate logistic regressions for them and suggests the score. The main advantage is that logit leaf model controls for more complex relations between variables.

Makhado (2023) points out that credit scoring models have numerous drawbacks. They are based on historical data and it can take much time to transform inner patterns. For example, people from less affluent areas have on average worse credit history. It leads to higher interest rates for them, which causes even more defaults among this group. Furthermore, scorecards often rely on proxies such as postal codes, which may appear harmless for default prediction but can reflect longstanding socioeconomic disparities (Consumer Financial Protection Bureau, 2012). In addition to, the credit history is usually an important factor for credit scoring. According to Hull (2015), people in developed countries may fall into a loop when they are refused credit since they do not have history, which they do not have because they are refused credit.

Ukrainian researchers also applied scorecards to credit risk analysis. Kolomiiets and Kochorba (2024) built a scorecard based on the logit model. They maintained monotonic binning even in categorical variables. The authors found that age and duration of loan contribute positively to default probability meaning that they will decrease overall score. This paper contributed to the existing literature by building a scorecard model using granular Ukrainian Credit registry data. We also compare results that are built using logistic regression and the SVM model.

Data. In this study, we used unique data from 2020 to 2024 from the Credit registry developed and supported by the National Bank of Ukraine (NBU). The frequency of data for this study is quarterly, which was transformed into annual because the probability of default according to NBU's Regulation № 351 "On credit risk estimation" is calculated on a one-year horizon (On Measuring Credit Risk …, 2016).

We subset only data on borrowers, which have at least one loan denominated in national currency (UAH) and issued after 01 January 2005. The loan types in the data are mortgage, consumer loan, and credit card loan. For this study, we aggregate data on the borrower level without subsetting specific loan types but may consider it for further research to develop separate scorecards for different loan types.

We define the dependent variable as a binary default indicator, in line with Dirma and Karmelavicus (2023). If the household was in default on a loan on the given date, such a borrower was excluded for the next year. If the borrower was not in default on the loan on the given date and then did not default during the year, the default indicator equals 0. Otherwise, if the borrower was not in default on the given date and then defaulted during the year, the dependent variable equals 1. We employ the rolling window approach. The schematic representation of the construction process of the dependent variable is shown in Fig. 1.

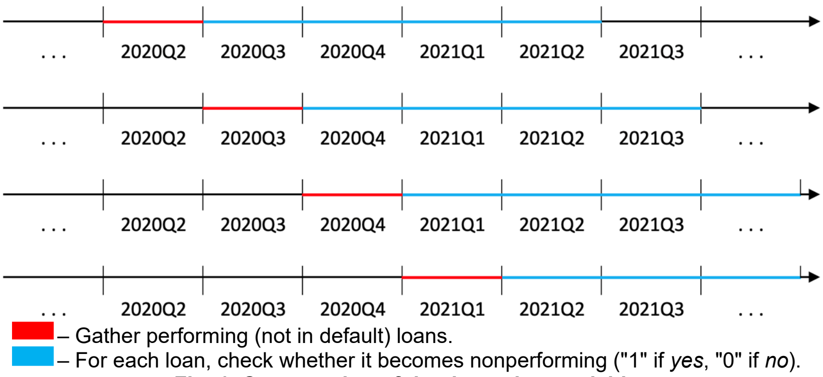


Fig. 1. Construction of the dependent variable

Source: Dirma and Karmelavicus (2023).

We incorporate the same variables as in Krasovytskyi and Stavytskyy (2024) namely income, DSTI, age, residual maturity, interest rate, credit risk, and GDP growth with the addition of new ones. The first new variable is a dummy variable whether the borrower has a mortgage loan. Since mortgages are long-term loans usually with variable interest rates (for loans granted before 2022), we expect that it will be difficult for borrowers to serve them as a result leading to defaults.

The second new variable is the current Loan-to-Value ratio (LTV) of the biggest loan the borrower has. When the bank issues a loan to minimize its loss given default it demands the collateral from the borrower. For mortgages and car loans, these are usually the objects that are bought. In case of the borrower's default bank may obtain the object and sell it via the market price. The goal for the bank is to keep the loan value lower than the price of the object so that all risks will be covered. In countries where LTV is a macroprudential limit values around 70–80 % are the maximum limit for mortgages. We limit the values of LTV with 100% in our data since it makes no sense for banks to provide loans for bigger sums than the property costs.

The third new variable is the days overdue of the biggest loan principal. The number of days overdue is the main indicator that defines credit class according to Resolution 351. More than 95 % of borrowers pay on time, but the rest don't. The more time passes since the borrower stopped paying, the less likely he will repay the debt. Another new variable is gender. Gati (2023) proves that female borrowers default less than male ones. On the contrary Costa et al. (2020) found that gender identity doesn't play a role in default prediction.

The final list of indicators for modelling and their descriptive statistics are presented in Table 1. Since the number of defaults in the original data is only close to 6%, we do the downsampling procedure. The downsampling was done to preserve the distribution of characteristics of non-defaulted borrowers. It will guarantee that we capture correctly the characteristics of defaulted borrowers and at the same time don't overfit the model. The number of observations in the final sample is 1293974, with a unique number of borrowers equal to 685073. Annex A presents additional statistics for each variable.

Table 1

Descriptive statistics of the dataset

Variable	Level	Unit of measurement	Min	Max	Median	Mean	Proportion of zeros, %
DSTI	Borrower	%	0	300	14	29	0
Age	Borrower	Years	18	72	42	43	0
Maximum loan interest rate	Loan	%	0	75	37	28	22
Annual income	Borrower	Thousand UAH	0	15407	193	470	0
Aggregate borrower credit risk	Borrower	Thousand UAH	0	24657	1,9	12,2	0
Residual maturity of a maximum loan	Loan	Years	0	24	7,7	7,9	0
LTV	Loan	%	0	100	0	2	96
Days overdue	Loan	Days	0	365	0	5	88
GDP growth	Macroeconomy	%	-35,9	19,4	2,3	-3,7	0

Source: own calculations based on the NBU's Credit registry.

Methods

This study employs a multistep analysis of credit risk as in (Dung, 2018). As a first step, we do the Weight of Evidence binning as in Wu and Pan (2021). Based on the binning results we estimate the logistic regression and the SVM machine-learning model. These models are used to build a credit scorecard. Based on the scores we find the optimal cutoff points to classify borrowers into defaulted and non-defaulted.

Weight of Evidence binning. The Weight of Evidence (WoE) is a statistical measure used in predictive modeling, particularly in credit risk scorecards, to transform categorical or continuous variables into a format that is more suitable for understanding

$$WoE_i = \ln\left(\frac{\frac{Good_i}{Total\ good}}{\frac{Bad_i}{Total\ bad}}\right), \quad (1)$$

where WoE_i – the weight of evidence of bin *i;* $Good_i$ and Bad_i – number of "good" and "bad" outcomes in bin *i;* $Total\ good\ and\ Total\ bad$ – total number of "good" and "bad" outcomes in the data.

Using WoE for binning is justified on the following basis. Firstly, WoE ensures that the relationship between the predictor and the dependent variable (default or no default) is monotonic, which is necessary for credit risk modeling. It makes decision-making simpler and more transparent. Secondly, Binning helps to smooth out the effect of outliers by grouping values into bins. This reduces the impact of extreme values, leading to more stable models. Thirdly, the WoE binning transformation makes the model more interpretable. Each bin's contribution to the risk score can be easily understood, making it easier to explain the model to stakeholders. The main disadvantage of binning is that it is partially subjective meaning that the number of bins and the cutoff points are not chosen based on some mathematical function. In this paper, we do not select more than five bins per variable to keep the model as simple as possible. An example of bins is given in Fig. 2. Here we developed four bins with interpretable cutoff points of 10 %, 20 %, and 40 % and the number of borrowers in each bin is relatively close to each other. The share of defaults per bin also increases monotonically.

We calculate the weight of evidence of each bin of each variable. The data is then substituted with the WoE of those bins to which the point belongs. This newly updated sample is used for modelling.

Modelling and calculating scores. We estimate two models namely logistic regression traditionally used in credit risk analysis and SVM, which is an ML method for comparison. We use SVM since it is not a tree-based method and it provides weights as an estimation result, which are comparable with coefficients in the logistic regression. Thus, we can directly use the SVM weights to calculate scores. The scores per bin of each variable are calculated using the following formula in accordance with Siddiqi (2012):

$$Score_{i,j} = \beta * \gamma_i * WoE_{i,j}, \quad (2)$$

where $Score_{i,j}$ – score awarded if the variable *i* is in the bin *j; β* – coefficient that defines score distribution; γ_i – model coefficient of the variable *I;* $WoE_{i,j}$ – the weight of evidence of variable *i* in the bin *j.*

The final score that is awarded to a person is calculated using the following formula:

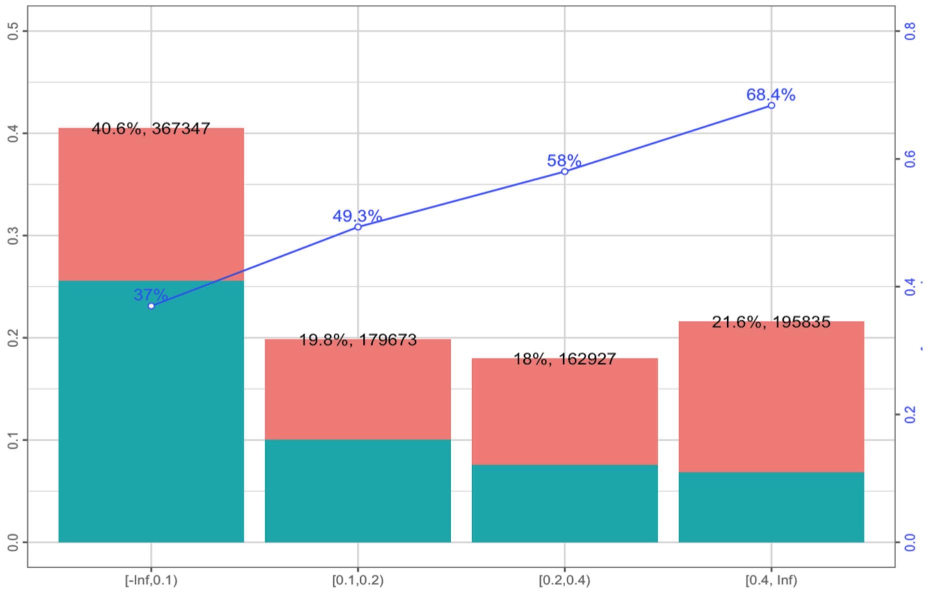
$$S_k = \alpha + \sum Score_{i,k}, \quad (3)$$

where S_k – final score of person *k;* $Score_{i,k}$ – score of variable *i* of person *k; α* – coefficients that defines the distribution of scores. We take α as 600 and β as approximately 72 (50/ln(2)). Practically, the α determines the shift of the whole score distribution to the left or right. β determines the width of the distribution. These values are standard used in the scorecard as in Siddiqi (2012) and Dung (2018).

Results

Results comparison. To compare models, we find the optimal cutoff points based on metrics such as accuracy, precision, and F1 score. In the end, we offer a specific score cutoff which can be used for banks to decide whether to reject or accept the loan or to recognize default.

Estimation results. In this section, we presented the results for scorecards. We start by analyzing logistic regression results. Table 2 presents scores that are awarded for the key variables. Higher scores are associated with a lower probability of default. High values of DSTI, age, and days overdue are positively associated with lower scores. Moreover, scores for days overdue fall very quickly indicating the high importance of this variable for default. Credit risk significantly increases even after one week overdue. Younger individuals have higher credit scores, which is in line with Kolomiiets and Kochorba (2024), and Krasovytskyi and Stavytskyy (2024). Being a female leads to higher creditworthiness as well. Economic growth below 2 % is not enough to have positive scores. It means that for all borrowers under negative economic conditions scores will be lowered, which can lead to default recognition or rejecting new loans.



– Number of defaulted borrowers. – Number of nondefaulted borrowers. – Share of defaulted borrowers per bin

Fig. 2. DSTI bins

Source: author's elaboration.

Table 2

Scores based on logistic regression

Variable	Bin	Score	WoE
DSTI	<10%	15	-0.53
	10-19%	1	-0.03
	20-39%	-10	0.33
	>40%	-22	0.77
Days overdue	0	11	-0.21
	1-6	-86	1.64
	7-29	-117	2.23
	30-59	-132	2.52
	>60	-143	2.73
Age	<35	19	-0.27
	35-44	2	-0.02
	45-54	-9	0.13
	>55	-15	0.23
Gender	Male	-7	0.07
	Female	8	-0.09
GDP growth	<2%	-19	0.17
	>=2%	15	-0.13
Income	<200	-2	0.03
	>=200	2	-0.03
LTV	0	-1	0
	>0	19	-0.1
Dummy for mortgage	0	-1	0
	1	27	-0.3
Interest rate	<40	8	-0.04
	>=40	-15	0.08
Credit risk	<0.5	47	-0.9
	0.5-2	21	-0.4
	2-10	-5	0.09
	>10	-81	1.53
Dummy for consumer loan	0	8	-0.39
	1	-7	0.37
Residual maturity	<1	-19	-0.02
	>1	15	0.01

Source: calculated by authors.

In Fig. 3, we can see that non-defaulted borrowers have higher scores than defaulted ones. There is no clear separation boundary since there are numerous defaults around 600. However, the left tail of defaulted borrowers is much longer and there is a smaller number of defaults after the score of 600.

To find the optimal threshold, we calculate performance metrics such as Accuracy, precision, F1-score, etc. We set possible threshold values to vary from 400 to 700 points with a step of 20. This ensures a balance between the optimal threshold and its easy interpretability.

Figure 4 presents the model's performance metrics on different cutoff thresholds. Accuracy is maximized at 600, while F1-score is maximized at 620. We choose 600 as a threshold for default recognition because 620 is too close to 628 which is the average value for non-defaulted borrowers. Now let's compare logistic regression results with the SVM model. Since the WoE values were calculated before the modelling stage, they are the same for both models. The only difference in scores is caused by the model coefficients. Although the results of each individual bin of variables are close to those produced by logistic regression, the final distribution of credit scores may differ a lot, because we have 12 variables in the model.

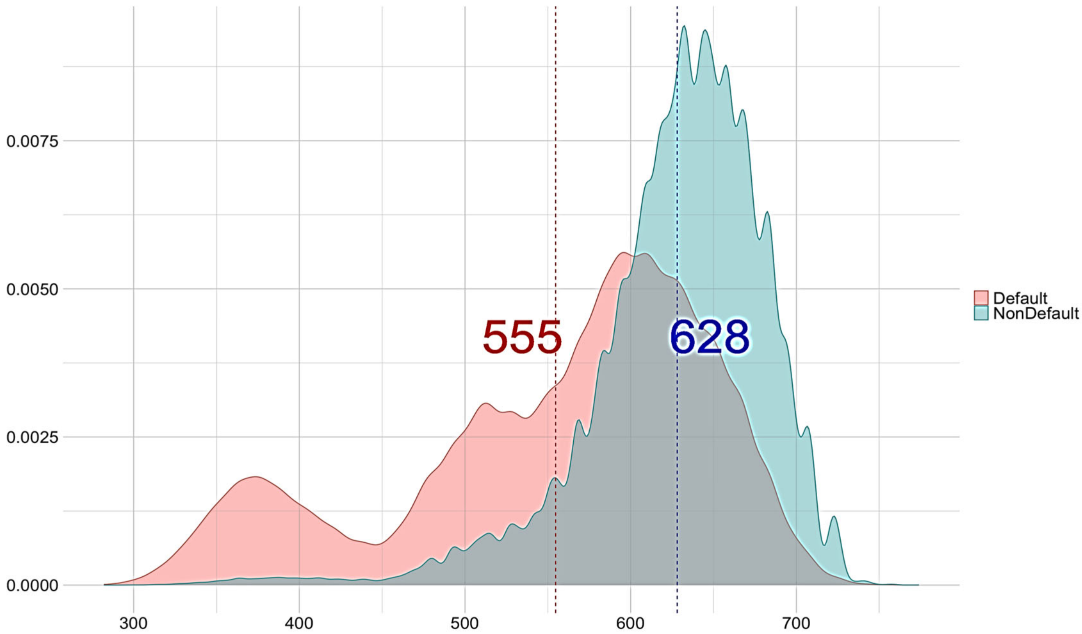


Fig. 3. Distribution of scores based on logistic model across default groups

Source: calculated by authors.

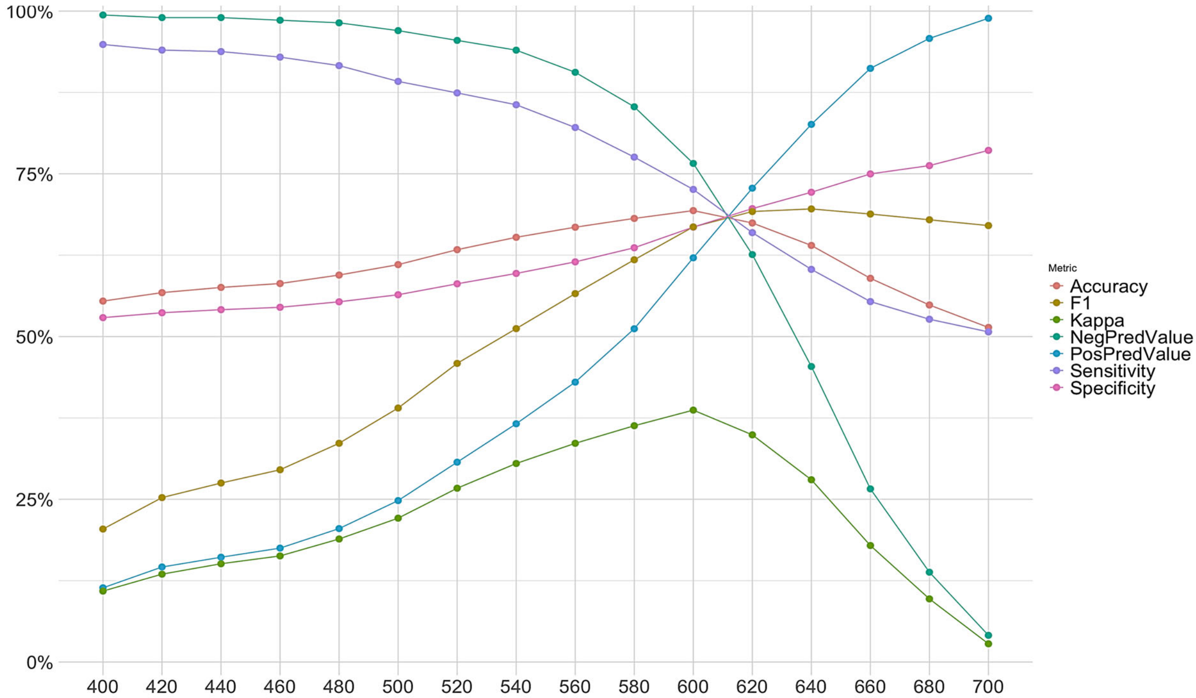


Fig. 4. Performance metrics of the logistic model

Source: calculated by authors.

Table 3

Scores based on SVM model

Variable	Bin	Score	WoE
DSTI	<10%	13	-0.53
	10-19%	0	-0.03
	20-39%	-8	0.33
	>40%	-19	0.77
Days overdue	0	10	-0.21
	1-6	-73	1.64
	7-29	-99	2.23
	30-59	-112	2.52
	>60	-121	2.73
Age	<35	15	-0.27
	35-44	1	-0.02
	45-54	-7	0.13
	>55	-13	0.23
Gender	Male	-5	0.07
	Female	6	-0.09
GDP growth	<2%	-17	0.17
	>=2%	13	-0.13
Income	<200	-2	0.03
	>=200	2	-0.03
LTV	0	-1	0
	>0	15	-0.1
Dummy for mortgage	0	0	0
	1	15	-0.3
Interest rate	<40	6	-0.04
	>=40	-12	0.08
Credit risk	<0.5	50	-0.9
	0.5-2	22	-0.4
	2-10	-5	0.09
	>10	-86	1.53
Dummy for consumer loan	0	10	-0.39
	1	-9	0.37
Residual maturity	<1	-2	-0.02
	>1	7	0.01

Source: calculated by authors.

Figure 5 shows that the difference between average defaulted and non-defaulted borrowers is close to that one from logistic regression (71 versus 73 before). The shapes of the distribution of both classes are also close to that depicted earlier by the logit model. We can conclude that despite a more complicated estimation process SVM model provides nearly the same results as logistic regression.

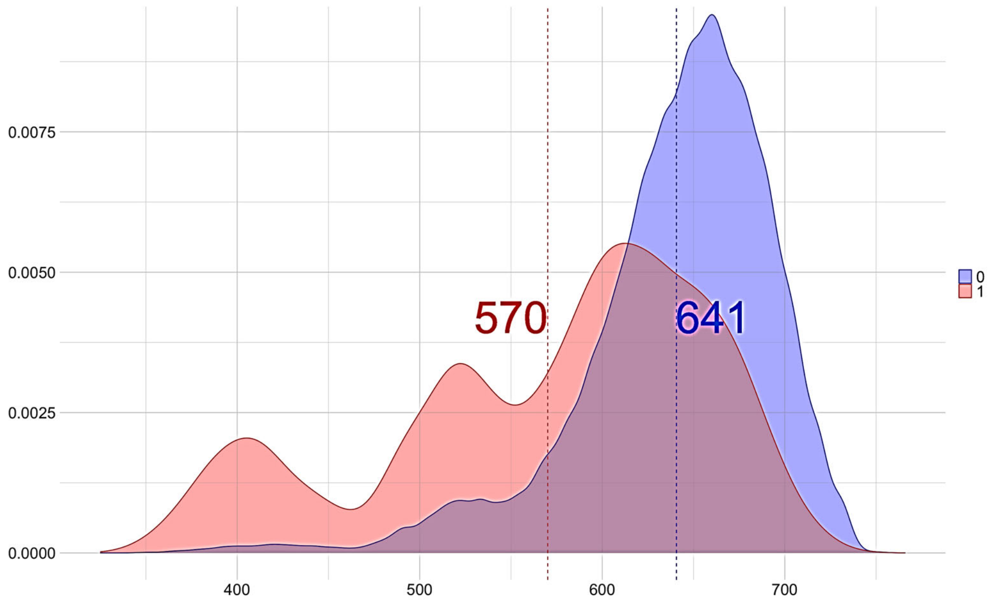


Fig. 5. Distribution of scores based on the SVM model across default groups

Source: calculated by authors.

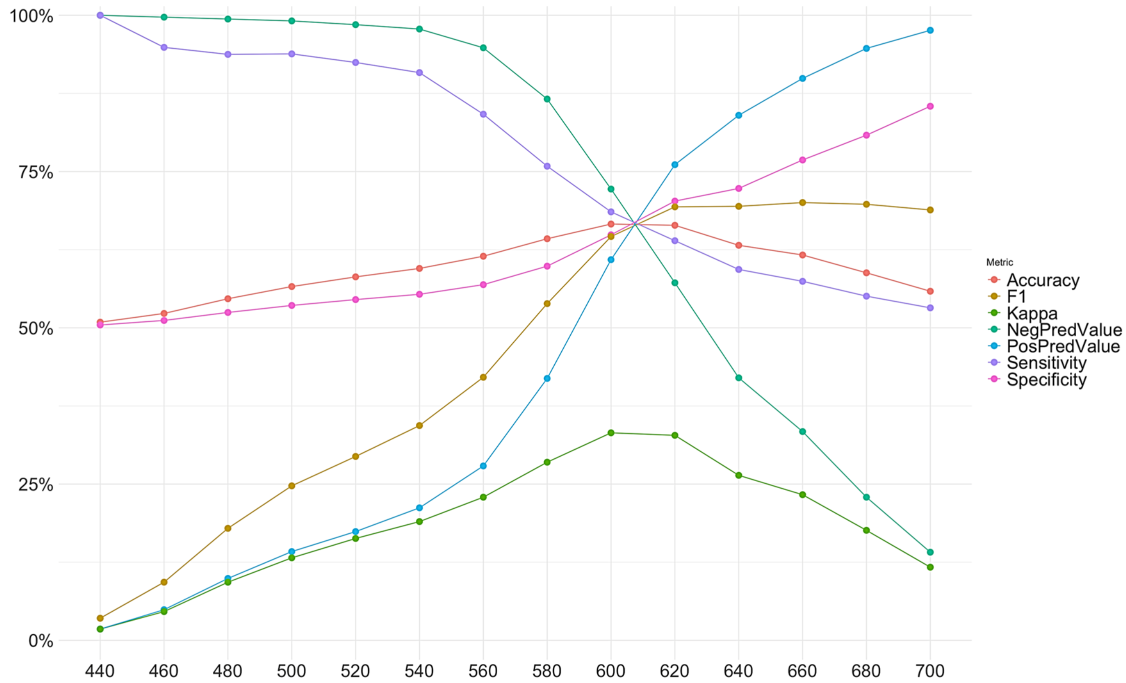


Fig. 6. Performance metrics of the SVM model

Source: calculated by authors.

Performance metrics of the SVM model's thresholds are also close to the logistic ones. The only difference is that the F1 score does not fall as fast after 620 points. The accuracy is still maximized around 600–620. We may choose the optimal threshold of 620 points since both accuracy and F1 score are maximized there, and it is far from 641, which is a non-default average. However, from a practical point of view, both 600 obtained from logistic regression and 620 obtained from the SVM model are close to each other.

Discussion and conclusions

In this study, we compared scorecards built with the help of logistic regression and SVM model based on Ukrainian Credit registry data. *Firstly*, we showed that it is possible to divide Credit registry data into bins based on the weight of evidence maintaining monotonicity and general economic logic. This binning made the scorecard interpretable and easy to use. *Secondly*, the results of both models are very close despite completely different estimation strategies. Although, some bins provided different scores across models the total scores earned by borrowers are similar to each other. The average score for non-defaulted borrowers was 628 by logit and 641 by SVM. *Thirdly*, the optimal thresholds for determining default were chosen to be 600 for logit and 620 for the SVM model. The results are close in both absolute values and performance metrics such as accuracy and F1 score.

As for policy recommendations, the results presented in the paper are useful for both commercial banks and regulators. Scorecards can be used by banks to assess loans both at origination as well as during their lifetime. Using straightforward credit scoring based on fully computable characteristics such as interest rate, age, DSTI and days overdue can simplify credit decisions made by financial institutions.

Financial regulation institutions, such as National Bank of Ukraine, may use the results for developing different policies for different bins of borrowers based on their characteristics. By reviewing the scorecards of leading retail banks, the central bank can monitor lending practices and identify potentially risky segments. Moreover, scorecards can be used in stress testing exercises to simulate how financial institutions perform under adverse scenarios.

Despite the promising results, this research has several limitations. The focus on logistic regression and SVM models excludes other potentially more powerful machine learning techniques, such as Random Forest or XGBoost, which could improve predictive performance. Furthermore, scorecard models are sensitive to changes in population characteristics, which could impact the stability of the model's predictions over time.

As for further studies, we may compare scorecards with other ML methods namely tree-based ones and XGBoost. We also might consider some feature engineering techniques, because maybe some interactions or polynomials may improve the distinction between non-defaulted and defaulted borrowers. Moreover, in future we will develop are multi-class scorecard which will make bank decisions more flexible and allowing for more tailored lending decisions, pricing strategies, and risk management practices.

References

Consumer Financial Protection Bureau. (2012). *Analysis of differences between consumer- and creditor-purchased credit scores.* https://files.consumerfinance.gov/f/201209_Analysis_Differences_Consumer_Credit.pdf

Costa e Silva, E., Lopes, I. C., Correia, A., & Faria, S. (2020). A logistic regression model for consumer default risk. *Journal of applied statistics*, 47(13–15), 2879–2894. https://doi.org/10.1080/02664763.2020.1759030

Dirma, M., & Karmelavičius. J. (2023). Micro-assessment of macroprudential borrower-based measures in Lithuania. *Bank of Lithuania Occasional Paper Series*, 46.

Dong, G., Lai, K.K., & Yen, J. (2012). Credit scorecard based on logistic regression with random coefficients. *International conference on computational science, procedia computer science*, 1, 2463–2468.

Dung, N.C. (2018). *An application of credit scoring: developing scorecard model for a Vietnam commercial bank.* Rpubs. https://rpubs.com/chidungkt/442168

Du, Pisanie J., Allison, J.S., & Visagie J. (2023). A Proposed Simulation Technique for Population Stability Testing in Credit Risk Scorecards. *Mathematics*, 11(2), 492. https://doi.org/10.3390/math11020492

Gati, N.J. (2023). *Socio-Demographic Determinants of Default Rate among Digital Lending Platform Borrowers in Nairobi County, Kenya* [Doctoral dissertation, Moshi Co-operative University]. MoCU repository http://repository.mocu.ac.tz/bitstream/handle/123456789/1185/Nyakeri%20Gati.pdf?sequence=1&isAllowed=y

Hull, K.G. (2015). Hispanics face hurdles in access to credit, mortgages. *The Charlotte Observer.* https://www.charlotteobserver.com/news/local/article28765174.html

Idbenjra, K., Coussement, K., & De Caigny, A. (2024). Investigating the beneficial impact of segmentation-based modelling for credit scoring. *Decision Support Systems*, 179, 114170. https://doi.org/10.1016/j.dss.2024.114170

Kolomiiets, Y., & Kochorba, V. (2024). Assessing the Credit Risks in the Risk Management System of Banking Structures, *Business Inform*, 1, 320–332. https://doi.org/10.32983/2222-4459-2024-1-320-332 [in Ukrainian]. [Коломієць, Ю.Ю., & Кочорба В.Ю. (2024). Оцінювання кредитних ризиків у системі ризик-менеджменту банківських структур. *Бізнес Інформ*, 1, 320–332]. https://doi.org/10.32983/2222-4459-2024-1-320-332

Krasovytskyi, D., & Stavytskyy, A. (2024). Predicting Mortgage Loan Defaults Using Machine Learning Techniques, *Economy*, 103(2), 140–160. https://doi.org/10.15388/Ekon.2024.103.2.8

Lee, C.Y., Koh, S.K., Lee, M.C., & Pan, W.Y. (2021). Application of Machine Learning in Credit Risk Scorecard. *Soft Computing in Data Science. SCDS 2021. Communications in Computer and Information Science,* 1489, 321–334. Springer. https://doi.org/10.1007/978-981-16-7334-4_29

Machado, M.R., & Karray, S. (2022). Assessing credit risk of commercial customers using hybrid machine learning algorithms, *Expert Systems with Applications,* 200, 116889. https://doi.org/10.1016/j.eswa.2022.116889

Makhado, P. (2023). *The Limitations of Traditional Credit Scoring Systems*. Medium. https://medium.com/@phindulo60/the-limitations-of-traditional-credit-scoring-systems-e92833fdfa8a

On Measuring Credit Risk Arising from Banks' Exposures. (2016). NBU Regulation № 351 dated 30.06.2016 [in Ukrainian]. [*Положення про визначення банками України розміру кредитного ризику за активними банківськими операціями.* (2016). Постанова НБУ № 351 від 30.06.2016]. https://zakon.rada.gov.ua/laws/show/v0351500-16#Text

Siddiqi, N. (2012). *Credit risk scorecards: developing and implementing intelligent credit scoring*. Wiley and SAS Business Series.

The World Bank Group. (2019). *Credit scoring approaches guideline*. https://thedocs.worldbank.org/en/doc/935891585869698451-0130022020/original/CREDITSCORINGAPPROACHESGUIDELINESFINALWEB.pdf

Wang, W., Lesner, C., Ran, A., Rukonic, M., Xue, J., & Shiu, E. (2020). Using Small Business Banking Data for Explainable Credit Risk Scoring. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(08), 13396–13401. https://doi.org/10.1609/aaai.v34i08.7055

Weston, L., & Hinson, K. (2023). *Why Your Credit Score Is Important?* https://www.nerdwallet.com/article/finance/great-credit-powerful-tool

Wu, Y., & Pan, Y. (2021). Application analysis of credit scoring of financial institutions based on machine learning model. *Complexity*, 2021(1), 9222617.

Отримано редакцією журналу / Received: 15.09.24
Прорецензовано / Revised: 20.10.24
Схвалено до друку / Accepted: 17.11.24

Данило КРАСОВИЦЬКИЙ, асп.
ORCID ID: 0009-0008-7017-0175
e-mail: daniel.krass2112@knu.ua
Київський національний університет імені Тараса Шевченка, Київ, Україна

РОЗРОБКА РОЗДРІБНОЇ СКОРИНГОВОЇ КАРТИ ЗА ДОПОМОГОЮ МАШИННОГО НАВЧАННЯ

В с т у п . ***Банки використовують кредитний скоринг, щоб відстежувати обслуговування кредитів, формувати резерви та коригувати політику кредитування. В основі методу лежить призначення балів на основі кредитної історії та унікальних характеристик позичальника, що дозволяє кредиторам передбачити ризик дефолту та покращити кредитні умови для позичальників з низьким рівнем ризику. Із розширеним доступом до даних і обчислювальною потужністю стало можливим проводити оцінювання кредитоспроможності за допомогою нових методів з кращою прогностичною потужністю. Це дослідження спрямоване на розробку скорингової карти для українських роздрібних позичальників з використанням даних Кредитного реєстру, порівняння ефективності логістичної регресії та методів методу опорних векторів. Ключові питання дослідження стосуються потенціалу групування даних для покращення інтерпретації моделі, покращенням точності оцінок і відмінностей у порогових значеннях для прийняття рішень у системах, створених за допомогою логістичної моделі та методу опорних векторів.***

М е т о д и . ***У дослідженні застосовано аналіз ваги доказів, який перетворює змінні для встановлення монотонного зв'язку з ризиком дефолту, тим самим покращуючи інтерпретацію та стійкість моделі. Використовуючи ці згруповані дані, у статті будуються скорингові карти за допомогою логістичної регресії та методу опорних векторів. Для оцінювання ймовірності дефолту кожна скорингова карта використовує передбачувані змінні, такі як співвідношення обслуговування боргу до доходу (DSTI), вік, процентні ставки та кількість днів прострочення. Оцінки призначаються на основі впливу кожної змінної на ймовірність дефолту.***

Р е з у л ь т а т и . ***Висновки свідчать про те, що можна розробити скорингову карту на основі даних Кредитного реєстру. Логістична регресія та метод опорних векторів дають схожі розподіли балів з високою точністю прогнозування та надійністю за показником F1. Скорингова карта забезпечує прозорість та можливість інтерпретації, наприклад, позичальники з DSTI понад 40 % отримують нижчі бали, що вказує на вищий ризик.***

В и с н о в к и . ***Банки можуть використовувати як логістичні моделі, так і моделі методу опорних векторів для оцінювання кредитоспроможності в реальному часі, використовуючи доступні характеристики позичальника для спрощення прийняття рішень. Для регуляторів скорингова карта може допомогти формувати політики, які обмежують кредитування, таким чином пом'якшуючи ризики, пов'язані з конкретними сегментами роздрібного кредитування.***

К л ю ч о в і с л о в а : ***машинне навчання, скорингова карта, прогноз дефолту, логістична регресія, метод опорних векторів.***

Автор заявляє про відсутність конфлікту інтересів. Спонсори не брали участі в розробленні дослідження; у зборі, аналізі чи інтерпретації даних; у написанні рукопису; в рішенні про публікацію результатів.

The author declares no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; in the decision to publish the results.