Performance Evaluation of Machine Learning Models for Credit Risk Prediction

Chief Assist. Prof. Dr. Yanka Aleksandrova
University of Economics - Varna, Varna, Bulgaria
yalexandrova@ue-varna.bg

Assoc. Prof. Dr. Silvia Parusheva
University of Economics - Varna, Varna, Bulgaria
parusheva@ue-varna.bg

Abstract

The purpose of this research paper is to propose an approach for calculating the optimal threshold for predictions generated by binomial classification models for credit risk prediction. Our approach is considering the cost matrix and cumulative profit chart for setting the threshold value. In the paper we examine the performance of several models trained with homogeneous (Random Forest, XGBoost, etc.) and heterogeneous (Stacked Ensemble) ensemble classifiers. Models are trained on data extracted from Lending Club website. Different evaluation measures are derived to compare and rank the fitted models. Further analysis reveals that application of trained models with the set according to the proposed approach threshold leads to significantly reduced default loans ratio and at the same time improves the credit portfolio structure of the Peer-to-Peer lending platform. We evaluate the models performance and demonstrate that with machine learning models Peer-to-Peer lending platform can decrease the default loan ratio by 8% and generate profit lift of 16%.

Keywords: machine learning, credit risk prediction, artificial intelligence, Peer-to-Peer lending, stacked ensemble classifiers

JEL Code: O33

Introduction

The market for alternative finance globally has grown remarkably. For the period starting from 2015 to 2018 its volume has almost doubled, reaching USD 305 billion (Dimitrov et al., 2020). Nearly 97% of this sector is made up of crowdlending platforms, with Peer-to-Peer consumer lending being the most common alternative finance business model with a share of around 64% of all models in this sector (Ziegler & Shneor, 2020).

Peer-to-Peer lending brings undeniable advantages for both investors and online platforms, businesses, and individual consumers. However, this business model also poses risks arising from the specifics of this sector and the dynamic environment. One of the main risks is related to an increase in the share of default loans. Bad loans are a serious threat to many investors entering the market as well as to online P2P lending platforms and borrowers. This also determines the crucial importance of the process of assessing borrowers and predicting the risk of non-repayment of the loan.

At the same time, the improvement of artificial intelligence and machine learning technologies in recent years has led to their ubiquitous application in all spheres of economy and public life (Turygina et al., 2019), (Petrov et al., 2021). Successful examples of the use of machine learning to predict and prevent serious threats and unintended consequences are numerous and continuously demonstrate the excellent capabilities of these credit risk predicting methods.

1. Research methodology

The model evaluation and comparison are essential to choose the best model for implementation. When assessing model performance and its predictive power several metrics are widely used depending on the model types. The most important evaluation measures for binomial classification models are accuracy, balanced accuracy, sensitivity, specificity, area under ROC curve, F1 score, Kappa coefficient, etc. It should be considered that the set threshold directly influences the measures derived from the classification matrix. Changing the threshold level results in a change in values of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) cases. To determine the threshold machine learning algorithms usually set it at such level as to achieve the best value of F1 score. However, when implementing machine learning models in the credit approval process it should be considered not only the mentioned measures but the results from the model implementation.

When choosing the best model, we propose to calculate and consider the cost matrix (Gutzwiller & Chaudhary, 2020). The costs in the matrix are marked with C (i, j) and are equal to the average costs associated with misclassification of class i cases in class j. For example, C (1,0) is the cost to the company of classifying bad credit as good. When lending, these costs equate to the average loss from a bad loan. Costs due to incorrect classification of good credit as bad - C (0,1) - are lost profits and equal to the average profit from the loans repaid.

To compare the different models in terms of the possible financial result of their implementation, we propose to use a profit growth metric (Lift) that shows the percentage change in the result with and without the application of machine learning models. We denote as positive cases loans fully paid on term. This metric is calculated as in (1).

$$Lift\ (\%) = \frac{TP * C(0,1) + FP * C(1,0)}{(TP+FN) * C(0,1) + (TN+FP) * C(1,0)} - 1 \qquad (1)$$

It can be concluded from (1) that the highest lift would give a model where there is a minimum number of false negative and a maximum number of true negative cases. If the product of the number of false negative and average profit is greater than the absolute value of the product of truly negative cases and their respective misclassification costs, it will have a positive increase in the financial result and therefore a profit from the application of machine learning. Otherwise, the model will result in a loss for the online lending company.

Online P2P lending companies should strive to maximize profits but should at the same time maintain the default loans ratio within acceptable limits. As a result of the application of a machine learning model, the share of bad loans is calculated as in (2).

$$Default\ ratio = \frac{FP}{TP+FP} \qquad (2)$$

The next step is to determine the optimal threshold for generating the predictions. A threshold of 0.5 is usually set for classification models, but this value does not always lead to optimal performance metric values. A decrease of the threshold below 0.5 increases the number of positive predictions, but also the false positive count. Respectively, an increase of the threshold above 0.5 increases the number of false predictions which can lead to an increase in the number of false negatives cases. The appropriate threshold for the best false positive rate can be determined by the ROC curve as well. In this case, however, we are looking not only for a model that is accurate, but also one that allows optimal financial results to be achieved from its implementation, while naturally also observing an acceptable default credit ratio and false positive cases.

To determine the optimal threshold, we propose the following sequence of steps:

1. The trained binary classification model is applied to the data set and predictions are generated in the form of a probability of belonging to each of the two classes.
2. The data set is sorted in increasing order of the estimated probability of belonging to the positive class (p0).
3. The accumulated financial result for the entire data set is calculated. The revenue is the interest paid on the good loans and the costs – the outstanding part of the principal for bad loans.
4. A threshold should be determined where the financial result is optimal. For better performance, we recommend building a graph of the cumulative financial result at the different thresholds (see Figure1)

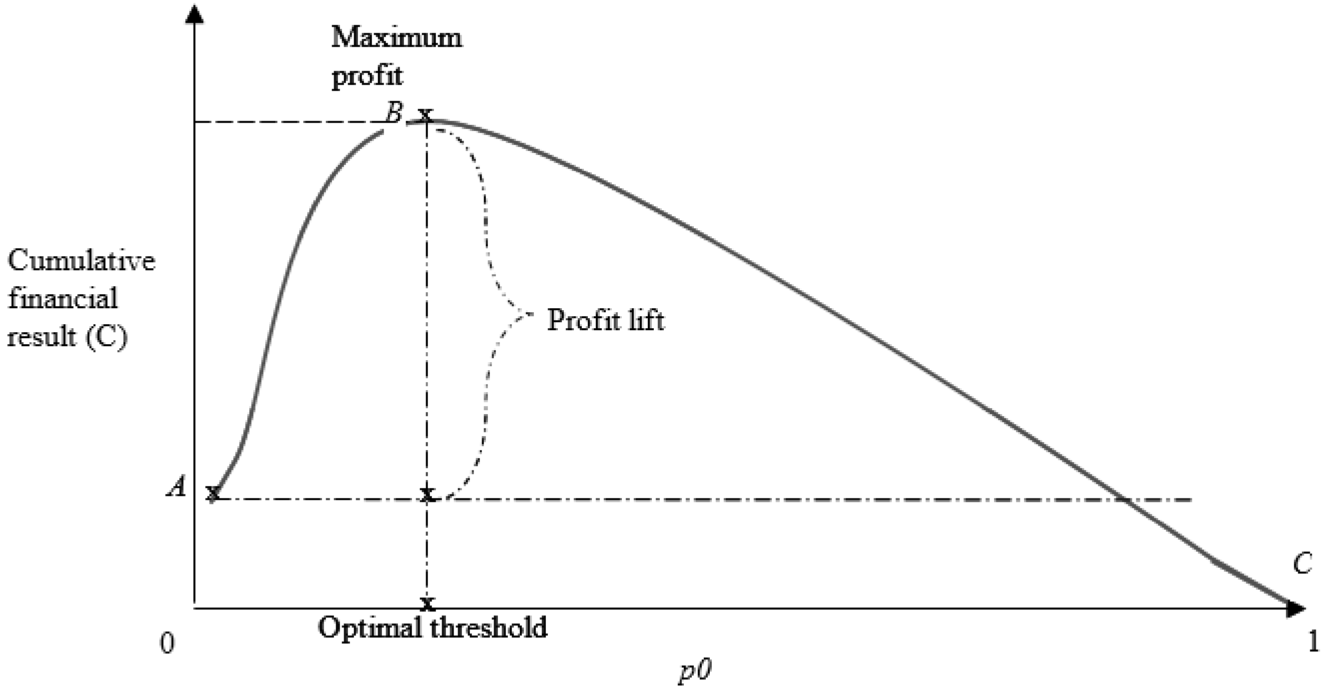

Figure 1. Cumulative financial result at different thresholds (p0)

The chart presented on Figure 1 provides an opportunity to visually explore the advantages of applying the trained model for predicting the status of loans. The cumulative result for probability threshold $p0_j$ is calculated using the following equation (3):

$$CumResult_j = \sum_{i=1}^{j} Result_i \ \left| p0_i \geq p0_j \right| \qquad (3)$$

$Result_i$ is the financial result, profit or loss, for the ith case and cases from 1 to i belong to the positive class with probability greater than or equal to the probability at threshold j – $p0_j$. The financial result in point A (see Figure 1) is equal to the sum of all profits and costs associated with loans granted and shows what is the financial result without machine learning implementation for credit status prediction. Since the data set is pre-sorted by increasing p0 values, then if the trained model had any predictive power at the beginning, negative cases (default loans) would prevail. The cumulative financial result sums up the profits and losses of all loans for which p0 is greater than or equal to the current threshold, and therefore at the beginning point A to point B there is an increase in profit due to the elimination of default loans. The increase in profit in the part of the graph from point A to point B shows the result of applying machine learning methods to predict default loans. However, with the p0 threshold increasing, more loans are classified as bad and fewer are approved, resulting in a reduction in profits due to lost income from solvent borrowers. At the end point C p0 is approximately equal to 1, which means that all loans are classified as bad and unapproved for financing and therefore the financial result is 0 – no loss, but no profit.

After determining the value of the p0 threshold at which there is an optimal profit, the different performance metrics such as accuracy, specificity, sensitivity, F1 score, Kappa coefficient, etc. should be calculated. It is also important to analyze the default credit ratio after applying the machine learning model at the selected threshold.

2. Data cleaning and processing

The dataset is downloaded from the Lending Club platform's website. This online P2P lending platform provides data in .csv for the loans granted since 2007. By 2016, the data are provided on an annual or several-year basis, and due to the increased number of loans from 2016 data is published each quarter. The last set of data provided by Lending Club at the time of the survey is for the second quarter of 2020, updated on 28.07.2020.

When evaluating and predicting the probability of default of the loan applicant many predictors can be used. Polena and Regner (Polena & Regner, 2016) use demographics characteristics (gender, age, marital status), education, employment length, income. Zhou, Zhang and Luo (Zhou et al., 2018) consider loan purpose, interest rate. Wang (Wang et al., 2018) take into account debt to income ratio, loan term. Total borrower's financial assets are also included in research by Serrano-Cinca and Gutierres-Nieto (Serrano-Cinca et al., 2015) and Carmichael (Carmichael, 2014) use customer behavior in the set of the predictors of credit risk.

Data provided by Lending Club is a subject of numerous studies. Serrano-Cinca (Serrano-Cinca et al., 2015) select 18 factor variables classified into five groups: borrower assessment by the lending organization, credit characteristics, credit history and indebtedness. Similar are the variables selected by Carmichael (Carmichael, 2014). Ariza-Garzon (Ariza-Garzon et al., 2020) also include the employment length at the current work, previous experience with the P2P lending platform, state code of the borrower, FICO score.

When selecting variables involved in the structure of the data set subject to subsequent analysis, the following criteria are considered:

- low percentage missing values. All selected variables have a percentage of missing values below 3%.
- lack of constant values for all observations.
- Information value based on theoretical research in the field of creditworthiness assessment.

In order to examine the impact of important macroeconomic indicators, data about the effective funding rate (Federal Reserve Bank of St.Louis, 2021) and unemployment rate (Federal Reserve Bank of St.Louis, 2021) have been added to the set. The data is synchronized with the issue date of the loans. After elimination of current loans, the set contains data on 1 914 456 loans, of which approximately 80% have been fully repaid and 20% default.

The structure of the data set with selected variables is shown in Table I. Variables marked with "*" are added at the feature engineering phase.

Appropriate techniques for outlier treatment were applied as outlier removal, capping and variable discretization. A significant part of variables contained missing values. Several R-packages for imputing missing values have been evaluated, such as MICE, Amelia, missForest, kNN. The conducted experiments showed that the best performing method with the smallest errors (root mean squared error, mean squared error, mean absolute error) is MICE which has been used for missing values imputation.

The final data set contained 1 467 296 cases describing credits from 2012 to 2017. The chosen independent variables can be classified into five categories (see Table 1):

A. General characteristics of the loan applicant.
B. Financial profile of the loan applicant.
C. Characteristics of the loan.
D. Indebtedness indicators.
E. Credit history of the loan applicant.

3. Model fitting

The fitting of machine learning models is implemented in H2O environment. H2O is an open source, scalable, distributed, fast, memory-based machine learning platform. It enables the building of machine learning models on big data and provides easy implementation of models in a working environment. The platform is built and provided by the H2O.ai company, whose corporate mission is the democratization of artificial intelligence. In the latest research for 2020 by the consulting company Gartner, H2O.ai is listed as a visionary in the field of data science and machine learning platforms (Krensky et al., 2020) and cloud services for artificial intelligence (Baker et al., 2021).

Table 1. Dataset structure

Variable	Description
A. General profile	
emp_length_n*	Employment length (numeric)
home_ownership	The home ownership status provided by the borrower
B. Financial profile	
annual_inc	The self-reported annual income
fico*	Average FICO score of the borrower
mort_acc	Number of mortgage accounts.
num_bc_tl	Number of bankcard accounts
num_il_tl	Number of installment accounts
num_sats	Number of satisfactory accounts
pub_rec	Number of derogatory public records
tot_cur_bal	Total current balance of all accounts
total_bal_ex_mort	Total credit balance excluding mortgage
total_bc_limit	Total bankcard credit limit
verification_status	Indicates if income was verified by LC, not verified, or if the income source was verified
C. Loan characteristics	
loan_amnt	The listed amount of the loan applied for by the borrower
purpose	A category provided by the borrower for the loan request.
term	The number of payments on the loan - 36 or 60.
title_words*	Number of words used by the borrower to describe the loan
Indebtedness	
bc_util	Total current balance to credit limit for all bankcard accounts.
dti	Total monthly debt payments excl. mortgage and the requested loan, divided by the monthly income.
dti_loan*	Principal payment of the requested loan to monthly income
num_rev_tl_bal_gt_0	Number of revolving trades with balance >0
revol_bal	Total credit revolving balance
revol_util	Revolving line utilization rate
D. Credit history	
mnths_since_first_crl	Months since first credit line opened
mo_sin_old_il_acct	Months since oldest bank installment account opened
mo_sin_old_rev_tl_op	Months since oldest revolving account opened
open_acc	The number of open credit lines in the borrower's credit file.
total_acc	The total number of credit lines in the borrower's credit file

For model training we consider homogeneous ensembles based on decision trees (XGBoost, extreme gradient boosted decision trees, gradient boosting machine, distributed random forest), deep learning networks and stacked ensembles, supported by H2O. Stacked ensemble is a heterogeneous ensemble algorithm that finds the optimal combination of a set of prognostic models using a process called “stacking” (H2O.ai, 2020). These ensembles support regression, binary and multinomial classification. The concept of combining models and stacking them into an ensemble model was published in 2007 (Van der Laan et al., 2007) and further developed in 2010 Polley &

Van der Laan, 2010). These two publications use the term "super learner" to mean heterogeneous ensemble models with the arrangement of models based on different algorithms and the use of cross-validation to build the combining algorithm, the so-called "metalearner".

The models had been training with h2o.automl() function with set parameters for class balance, maximum number of models 20, excluding the two stacked ensembles, 5-fold cross validation, stopping metric logloss and time limit of 10800 seconds. As a result, 2 heterogeneous ensembles and 20 base classifiers had been trained, broken down by algorithm type, as follows: 7 with XGBoost, 7 – Gradient Boosting Machine, 3 – Deep Learning, 2 – Distributed Random Forest and 1 – Generalized Linear Model.

4. Model comparison

When assessing the performance of models, the cost matrix was considered. Costs due to misclassification of good credit as bad are equal to lost profits due to refusal of the loan request. These costs amount to USD 2616, which is the average profit of a loan repaid. The costs incurred by the incorrect classification of bad credit as good are equal to USD 5952, as is the average loss on a default loan.

The performance of the best ten models on the test set is shown in Table II. All models were applied to a test set of 406 368 cases where the total actual profit was USD 400 814 860. Using the values of an average profit and loss, an estimated profit that Lending Club would have received as a result of applying the relevant model was calculated. It is assumed that when using a trained model, all loans for which there is a negative prediction are canceled and all positively predicted loans are approved. The estimated profit shall be equal to the profit from the correctly classified actual positive cases plus losses from the incorrectly classified actual negative cases, with losses recorded with a negative sign. The "Lift" column calculates the percentage change in estimated profit relative to the actual profit Lending Club received from these loans. In the last column "Default ratio" the share of bad loans is calculated if the P2P lending platform applies the relevant model and decides to approve or refuse the loan request based on the predictions generated by the model.

Table 2. Performance of the top ten models

Model	Sensitivity	Specificity	Lift	Default ratio
StackedEnsemble- All_Models	0.7561	0.5312	9%	13%
StackedEnsemble-BestOfFamily	0.7474	0.5408	8%	13%
XGB_3	0.7008	0.6046	5%	12%
DRF_1	0.6964	0.6112	5%	12%
XGB_1	0.7424	0.5414	7%	13%
GBM_3	0.7297	0.5596	6%	12%
XGB_4	0.7073	0.5915	5%	12%
XGB_GR_3	0.7054	0.5936	5%	12%
XGB_GR_1	0.7044	0.5950	5%	12%
XRT_1	0.7226	0.5655	5%	12%

The ratio between costs related to misclassification is approximately 1:3 in favor of the negative class, i.e., losses from misclassifying a negative case as positive are three times higher than lost profits in misclassifying a positive case as negative. At the same time, the expected distribution between the two classes is 1:4 in favor of the positive class. The evaluation measures showed that models with better accuracy and Kappa values were those with better recognition of the positive cases.

In all models, the default ratio after model application is significantly lower than the default ratio in the original set. For the ten models presented in Table II, this ratio ranges from 12% to 13%, while in the data set this ratio is about 20%. This strongly supports the advantages of machine learning models for credit risk prediction and reduction of the share of bad credits.

The highest profit lift (9%) was observed in the heterogeneous ensemble model StackedEnsemble AllModels. It had the highest Kappa values (0.2356) and sensitivity (0.7561). Therefore, this model was chosen to demonstrate profit maximization by setting the optimal threshold.

5. Calculating the best threshold value

The probability threshold for positive class predictions (p0) is set by the algorithms to achieve a maximum value of F1 score. In the selected best model StackedEnsemble-AllModels this threshold is set at 0.815, achieving F1 score of 0.811. This means that all cases where p0 is above 0.815 are classified as positive and the rest as negative.

In determining the level of the optimal threshold for maximizing profit, we apply the proposed approach of calculation. For the heterogeneous ensemble method according to the proposed methodology its value is 0.625. With threshold of 0.625 the profit lift is 16%, i.e., 7% higher than the lift with the 0.811 threshold calculated by the algorithm. Lowering the p0 threshold in general leads to an increase in the number of positive predictions, and hence to better sensitivity of the model. Models that achieve higher levels of profit are not necessarily the most accurate, but those where there is a sufficiently good ability to correctly classify positive cases. Of course, too low p0 values would lead to an increase in misclassified bad loans, and hence to an increase in the default ratio.

At the same time, we should recognize that this optimal threshold is as adapted as possible to the cases of the test set. To see if its value is applicable in other cases, another method of calculating cumulative profit is also applied. As established at the preliminary analysis, there are large variations in profits and losses in different categories of loans. Further research on these indicators shows that differences are also observed between loans from different subcategories, but the average levels of profit and loss in the training and test set for loans of the same subcategory and status are approximately equal. Therefore, to examine whether the optimal level of p0 does not depend on the specific cases, the average earnings of a fully repaid loan and the average loss of one bad loan for all loans of a subcategory were used (see Figure 2). When considering the real financial results for the subcategory and loan status, for the test data set, the optimal threshold is 0.625 which is relatively close to the 0.651 value calculated with average results by subcategory.

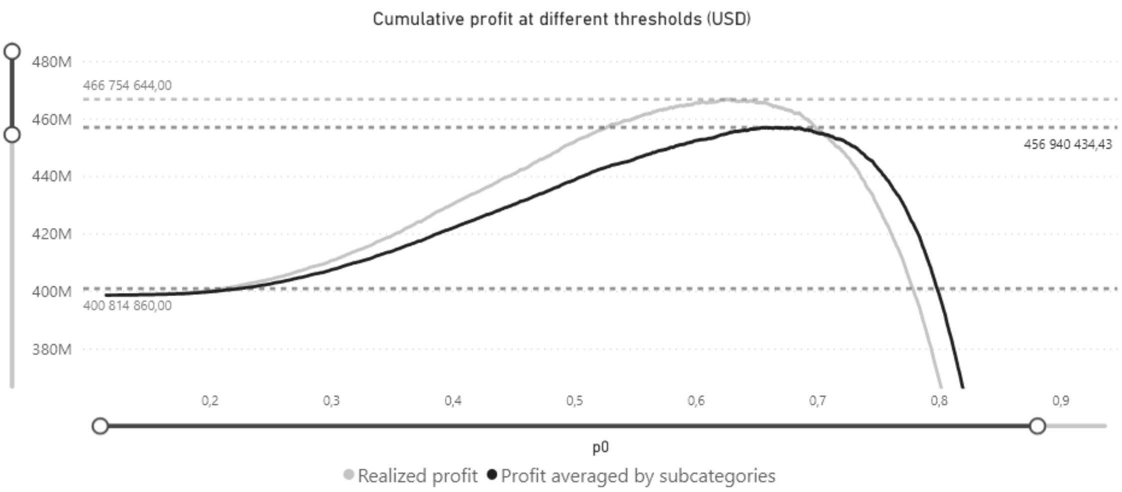

Figure 2. Real and averaged cumulative profit

Experiments were also carried out with other samples to determine the optimal threshold in both ways – using real and average profits and losses. The results show stability of the levels of the optimal threshold, and we can therefore conclude that to maximize the profit of applying the trained model to new, unknown data, p0 levels in the range of 0.62 to 0.65 should be considered, not those above 0.8 that determine an optimal F1 score.

6. Credit portfolio structure before and after machine learning application

Maintaining a diversified loan portfolio is an important element of the overall risk management strategy for the online P2P lending platform. Therefore, an examination of the structure of the portfolio of loans from the different categories of before and after application of machine learning models has been carried out. The purpose of such research is to determine whether there is a change in the structure of loans by the different categories. To analyze the effect of applying trained machine learning models, a comparison was also made regarding the default ratio in the test set before and after classifying cases with the best heterogeneous ensemble method – StackedEnsemble_AllModels. The results of the comparative calculations are given in Figure 3. For each category, the values of the indicators before and after application of the model are presented. Metrics measuring the structure and default ratio after applying the model are calculated based on the positive predictions generated by the model at the selected optimal threshold p0 for maximum profit. The aim is to simulate a scenario in which Lending Club applies the trained model during credit request process and approves only those where there is a positive prediction with a probability level above the selected threshold. Falsely positive cases are loans that are incorrectly defined as good and funded. Their number is used to calculate the default after applying the model.

As evident from Figure 3 the most significant decrease in default ratio at about 20% is observed for the riskier categories – E, F, G. These figures reveal that the trained model is better at identifying bad loans in these categories as opposed to loans in the higher categories – A, B, C and D.

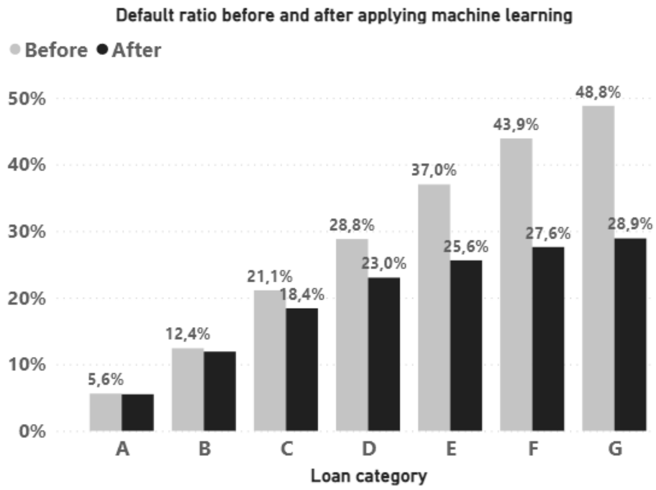

Figure 3. Default ratio by subcategories before and after machine learning implementation

The structure of the loan portfolio by category before and after machine learning implementation is shown on Figure 4. As evident from the figure the structure of the loan portfolio by category is generally maintained. If we assume model implementation for credit approval, the relative shares of loans of categories A, B, C will increase, with the largest increase being seen in category B. The relative shares of loans in the riskier categories D, E, F, G would decrease, with the largest reduction of 2.2 % in category F loans. These changes confirm the hypothesis that the use of machine learning models helps diversify the loan portfolio while improving its structure by increasing the share of credit in the better performing categories.

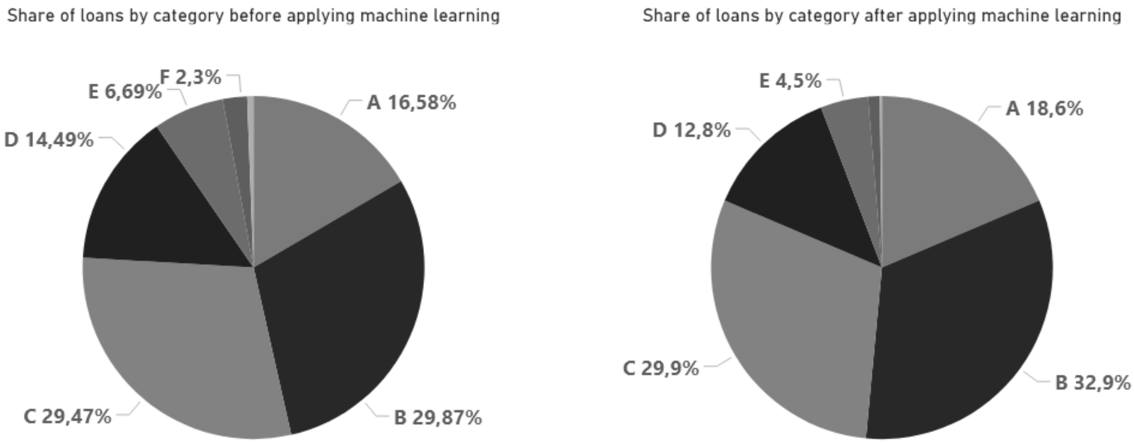

Figure 4. Credit portfolio structure before and after applying machine learning models

Significant changes in the reduction in the share of bad loans are observed after machine learning for the general data set. The default ratio for the initial data set is 20% and as a result of predicting with a trained classification model, it could be lowered to 12%-13%.

Conclusion

Analyzing the impact of applying machine learning models by examining structural changes and changes in the default ratio before and after machine learning shows that P2P lending companies can gain important advantages by implementing machine learning models trained according to the approach we offer. First, the default ratio as a result from more accurate predictions can reduce the share of bad loans overall and by category. Another benefit is that machine learning application can improve the credit portfolio structure by increasing the shares of loans from better categories which are likely to repay credits and reducing the share of the riskier loans. Last, but not least, the calculation of the optimal threshold for prediction generation following our proposed approach can maximize profit for the online Peer-to-Peer lending platform and investors. With regards to this results, we recommend using cost matrix and cumulative profit chart to determine the threshold and at the same time consider the traditional measure for evaluation of binomial classification models.

References

1. Ariza-Garzon, M. et al. (2020) Explainability of a machine learning granting scoring model in Peer-to-Peer lending. *IEEE Access*. vol.8, pp. 64873 – 64890. doi:10.1109/ ACCESS.2020.2984412.
2. Baker, V. et al. (2021) Magic quadrant for cloud ai developer services. [Online] Available from https://www.gartner.com/doc/reprints?id=1-1YGJKJ5P&ct= 200224&st =sb. [Accessed 03/10/2021].
3. Carmichael, D. (2014) Modeling default for Peer-to-Peer loans. *SSRN´s eLibrary*. doi:10.2139/ssrn.2529240.
4. Dimitrov, G., Petrov, P., Dimitrova, I., Panayotova, G., Garvanov, I., Bychko, O. and Petrova, P. (2020) Increasing the classification accuracy of EEG based brain-computer interface signals. *10th International Conference on Advanced Computer Information Technologies*, ACIT 2020 - Proceedings, pp. 386-390. doi:10.1109/ACIT49673.2020.9208944.
5. Federal Reserve Bank of St.Louis (2021) Effective federal funds rate. FRED. Economic Data, [Online] Available from https://fred.stlouisfed.org/series/ FEDFUNDS. [Accessed 17/02/2021].
6. Federal Reserve Bank of St.Louis (2021) Unemployment rate FRED Economic Data, [Online]

Available from https://fred.stlouisfed.org/series/UNRATE. [Accessed 17/02/2021].
7. Gutzwiller, K.J. and Chaudhary, A. (2020) Machine-learning models, cost matrices, and conservation-based reduction of selected landscape classification errors. *Landscape Ecology*. Vol. 35, № 2, pp. 249 – 255. Springer. doi:10.1007/s10980-020-00969-y.
8. H2O.ai (2020) Stacked ensembles. *H2O.ai official documentation*. [Online] Available from https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/stacked-ensembles.html. [Accessed 12.10.2020].
9. Krensky, P., Idoine, C., Brethenoux, E. et al. (2020) Magic quadrant for data science and machine learning platforms, [Online] Available from https://content.dataiku.com/gartner-mq-2021/gartner-mq-21. [Accessed 03/10/2021].
10. Petrov, P., Radev, M., Dimitrov, G., Pasat, A. and Buevich, A. (2021) A systematic design approach in building digitalization services supporting infrastructure. *TEM Journal*, 10(1), pp 31-34, 2021, doi:10.18421/TEM101-04.
11. Polena, M. and Regner, T. (2016) Determinants of borrowers' default in P2P lending under consideration of the loan risk class. *Jena Economic Research Papers*. No.2016-023. [Online] Available from: https://www.econstor.eu/handle/10419/148902 [Accessed 25/08/2020].
12. Polley, E. and Van der Laan, M. (2010). Super learner in prediction. *U.C. Berkeley Division of Biostatistics Working Paper Series*. [Online] Available from http://biostats.bepress.com/ucbbiostat/paper266. [Accessed 12/10/2020].
13. Serrano-Cinca, C., Gutiérrez-Nieto, B., and López-Palacios, L. (2015) Determinants of default in P2P lending. *PLoS ONE*. Vol. 10, pp. 1 – 22. doi:10.1371/journal.pone.0139427.
14. Turygina, V.F., Vasilev, J., Safrygin, A.S., Nizov, A.N. and Panteleeva, N.D. (2019) Comparison of statistical models for assessing the probability of bankruptcy of enterprises. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management - SGEM*, 19(2.1) pp. 175-180.
15. Van der Laan. M., Polley, E. and A. Hubbard (2007) Super learner. *Statistical Applications in Genetics and Molecular Biology*. Vol. 6, № 1. doi:10.2202/1544-6115.1309.
16. Wang, Z., Jiang, C., Ding, Y., Lyu, X. and Liu, Y. (2018) A Novel behavioral scoring model for estimating probability of default over time in Peer-to-Peer lending. *Electronic Commerce Research and Applications*, № 27, pp 74-82.
17. Ziegler, T. and Shneor, R. (2020) The global alternative finance market benchmarking report. Trends, opportunities and challenges for lending, equity, and non-investment alternative finance models. *Cambridge Center for Alternative Finance*, [Online] Available from: https://www.jbs.cam.ac.uk/fileadmin/user_upload/research/centres/alternative-finance/downloads/2020-04-15-ccaf-global-alternative-finance-market-benchmarking-report.pdf, [Accessed on 17/04/2020].
18. Zhou, G., Zhang, Y., and Luo, S. (2018) P2P network lending, loss given default and credit risks. *Sustainability*. 10(4), pp 1-15. doi:10.3390/su10041010.