Educational Administration: Theory and Practice
2024, 30(5), 12958-12966
ISSN: 2148-2403
https://kuey.net/

Educational Administration Theory and Practice

Research Article

Leveraging Machine Learning Techniques For Developing Robust Credit Scores For Peer-To-Peer Lending Platforms

E.Srinivas Jayaram[1*], Dr.G.Balachandar[2], Dr. Kompalli Sasi Kumar[3]

[1*]Research Scholar, Department Of Business Administration, Annamalai University, Tamilnadu, Es.Jairam@Gmail.Com, 7093025442
[2]Assistant Professor, Department Of Business Administration, Govt. Arts And Science College For Women, Alangulam, Tenkasi, Tamilnadu, Aubalachandar@Gmail.Com, 9952636466
[3]Associate Professor, GITAM School Of Business, Hyderabad, GITAM University, Skompall@Gitam.Edu, 9848192864

Citation: E.Srinivas Jayaram et al. (2024), Leveraging Machine Learning Techniques For Developing Robust Credit Scores For Peer-To-Peer Lending Platforms, Educational Administration: Theory and Practice, 30(5), 12958-12966
Doi: 10.53555/kuey.v30i5.5633

ARTICLE INFO	ABSTRACT

Peer-to-Peer lending platforms have become a widely accepted alternative to conventional banking, providing borrowers and lenders with a direct method of financial transaction. The study commences with an overview of P2P lending platforms, delving into the progression of credit scoring and providing a synopsis of the P2P lending market. A thorough examination of the existing literature establishes the foundation for our inquiry. We describe our data pre-processing procedures, which involve manipulating independent variables and filling in missing values, using a comprehensive dataset obtained from a prominent P2P lending platform. Subsequently, we employ automation to compute discrete and continuous variables to pre-process the data for analysis. The study progresses by splitting data into train and test datasets to calculate the probability of the default (PD) model. The logistic regression machine learning algorithm creates and verifies the credit score model. The model estimation process is subsequently validated using test data, where performance metrics such as Kolmogorov-Smirnov and Gini statistics are employed to estimate predictive accuracy and discriminatory power.
This study substantially improves credit scoring systems in P2P lending platforms by utilizing machine learning techniques. Our study strengthens lenders' risk assessment capacities, providing them with more powerful tools that promote trust and efficiency in lending. The results highlight the significant impact that machine learning may have on credit rating, enabling the sustainable expansion of the P2P lending sector.

Keywords: Credit Scoring, Preparing Credit Score Cards, Machine Learning Algorithms, Peer-to-Peer Lending, Default Risk Modeling, Default Forecasting

1. Introduction

Machine learning, defined by Arthur Samuel in 1959 as the Capacity of a computer to acquire knowledge and improve its performance without explicit programming, has significantly influenced the financial industry. This field's progress has enabled the separating of essential tasks such as processing payments, managing maturity transformation, and distributing risk. Additionally, machine learning has enhanced capital deployment due to the entry of new players like financial transaction facilitators, trading platform integrators, P2P lenders, and advanced exchanges.

The financial services sector has experienced substantial disruption driven by financial technology (FinTech) advancements. These advancements leverage big data innovations and the introduction of algorithm-based automated investment platforms (Bachmann et al., 2011). FinTech refers to the integration of IT to provide technology-enabled services and solutions. Advances in FinTech are reshaping the digital environment, enabling financial institutions to perform inter-platform payment transfers and transaction solutions (Thompson, 2017).

The FinTech peer-to-peer lending market has grown substantially since its inception in 2006 and is gaining popularity among small enterprise owners and consumers. Borrowers can obtain loans at more advantageous interest rates compared to those offered by traditional banks. Lenders can invest their money to achieve more significant returns and utilize advanced investment products. Following the economic meltdown, there was a change in attention from traditional banking to the emerging field of FinTech banking. FinTech banks offer products and services that are more convenient, yield superior returns, and are far more transparent than those provided by traditional banks. Consequently, consumers are increasingly showing interest in FinTech instead of conventional banking. An essential distinction in determining interest rates in this market is that FinTech banks predominantly depend on qualitative and quantitative information, such as the FICO score, loan purpose, education level, employment title, length, loan term, loan grade, annual income, and other features.

With the revolution in the technology sphere and the advent of algorithms used for data analysis, financial institutions can bring efficiency into their operations, thereby becoming competent in identifying the right borrowers, providing better investment avenues to their customers, and bringing rationality to payment services (Bofondi and Gobbi, 2017). Based on the lifestyle changes and rapid transformations that are taking place in the financial technology space, new businesses and business models are evolving (Baldassarre et al., 2017). Small and medium enterprises (SMEs) can source money from non-banking financial companies (NBFCs) to meet their business needs for maintaining adequate working capital, hiring human resources, and purchasing and maintaining the required raw materials (Mills & McCarthy, 2014).

1.1 P2P Lending Market Overview

Zopa's launch in the UK triggered significant growth in the sector, and now, P2P lending prevails across all continents. The 2021 study from Cambridge Centre - Alternative Finance reveals that the worldwide alternate funding market had a consumer segment of approximately USD34.7 billion in 2020 and is likely to achieve USD45 billion in 2030 (GlobalNewswire), while the commercial segment, which focuses on providing loans for business purposes, touched USD115 billion and is projected to touch USD920 billion by 2030 (Allied Market Research). In the United States alone, the aggregate amount of retail financing stood at around $4,142 billion in the third quarter of 2020, according to the Federal Reserve (FED, 2021). The share of the P2P platforms currently stands relatively tiny compared to the overall credit volume. Between 2014 and 2020, the market had a significant increase in growth. Figure 1 illustrates the development in this market.

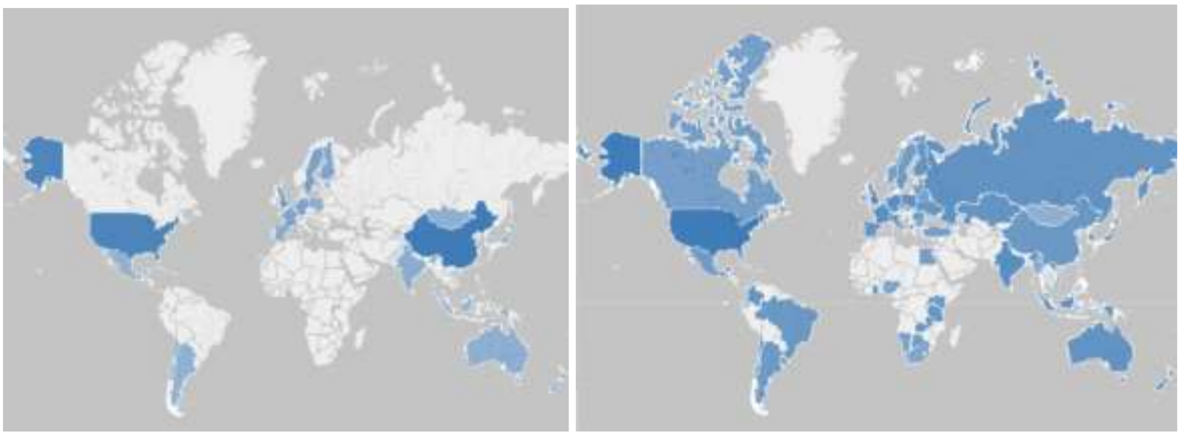

Figure 1: Market Evolution between the Years 2014 to 2020
Source: Olivedi Timea, based on Cambridge Centre - Alternative Finance, 2021

Figure 2 displays the geographical spread of the P2P markets. With 241 P2P platforms, Europe stands on the top, mainly because of its significant presence in the UK and a thriving P2P lending industry in the Baltic; these areas have well-established markets with numerous participants. Western Europe is the region with the second highest number of platforms. In Southern Europe, Italy and Spain combined have twenty-six platforms. North America has fewer P2P platforms than the European market, approximately one-eighth, but the loan volume is considerably high. The primary reason for this is the existence of large investors who enhance total lending volume. Mexico follows, boasting seven P2P platforms. India and Indonesia are the dominant players in the Asian market, with South Korea following behind.

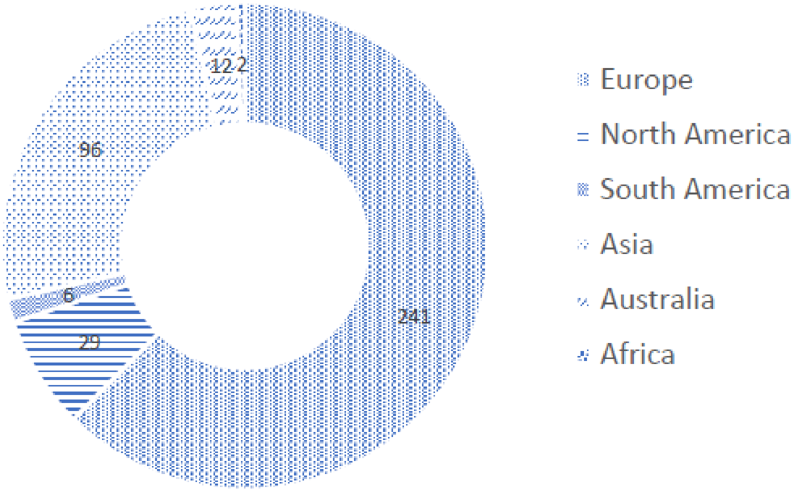

Figure 2: Regional Distribution of P2P Platforms in 2022
Source: Olivedi Timea, p2pmarketdata.com, 2022

2. Literature Review

P2P lending platform refers to private individuals originating loans on online platforms, with financial services firms acting solely as intermediaries as mandated by law. The first platform was established in 2005, organized by web-based social networks. The classification of P2P lending platforms is into two categories: commercial and non-commercial. The primary distinction is the lender's overall objective and returns expectations. Lenders aim to maximize their investment returns while managing a specific amount of uncertainty, while loan applicants with varying credit risks seek sources of funds. P2P platforms play the role of mediators, facilitating the connection between different organizations and striving to meet the anticipations of all the stakeholders. Occasionally, investors unite to establish tiny collectives to focus on their shared priorities (Wang & Greiner, 2011; Herrero-Lopez, 2009).

The study by Havrylchyk, Mariotto, Rahim, and Verdier in 2018 discovered a positive correlation between P2P platform growth and population density. Moreover, a direct relationship exists between the applicants' age and education and the extent to which these factors directly influence P2P lending platforms' growth. In addition, other factors influence investors to choose P2P lending platforms over traditional banks for their investments. Investors are motivated to join P2P lending systems because these platforms offer them the chance to fund consumer borrowing, allowing them to gain a better interest rate (Herzenstein, Andrews, Dholakia, & Lyandres, 2008).

Another incentive is the increased yield. P2P platforms have fewer costs associated with facilitating transactions; the borrowers are in an advantageous position as they are provided loans at a reduced interest rate and provide lenders with increased income. The primary factor contributing to the reduced intermediation costs is the absence of deposit collection by P2P platforms, which exempts them from some regulatory requirements (Serrano-Cinca, 2015).

Consequently, borrowers secure loans with more favorable interest rates than a conventional financial institution would offer. At the same time, lenders earn a better rate of interest on the investments. The interest rates on the loans taken from the P2P platforms may sometimes prove higher than those from conventional financial institutions. This disparity primarily stems from the elevated level of risk connected with the loans, as borrowers sometimes pose a higher risk than borrowers from traditional banks. When factoring in risk, P2P platforms' interest rates align closely with what conventional financial institutions offer (De Roure, Pelizzon, Tasca, 2016).

One final incentive that P2P platforms provide individual investors is the autonomy of their investments. Investors can specify their desired interest rate and determine the intended use of the loan they wish to invest in. Previously, smaller investors had access to a limited range of investment options and depended entirely on the options provided by the banks.

Researchers extensively studied the funding efficacy of P2P loans. This section will specifically examine the various aspects that impact the level of funding success for peer-to-peer (P2P) loans. Funding success occurs when investors fully fund a peer-to-peer loan. Herzenstein (2008) primarily examined the factors contributing to obtaining funds in the P2P marketplaces in his article. He conducted research using Prosper.com, another P2P lender. He discovered that demographic characteristics, such as borrowers' ethnicity, cultural background, and gender identity, impact the likelihood of obtaining funds to a smaller extent than the borrower's financial strength. (Herzenstein, Andrews, Dholakia, & Lyandres, 2008).

Furthermore, Herzenstein identified effort as the characteristic with the most significant influence on the loan's successful funding. This metric quantifies the borrower's effort in disclosing their information to the P2P platform to secure credit. This variable indicates that there is still a problem of unequal access to information in the P2P lending industry. Significant borrower effort in providing information leads to reduced information asymmetry. Herzenstein discovered that the likelihood of successfully obtaining finance increases as borrowers exert more effort in providing information. The reference is from Herzenstein et al. (2008).

Furthermore, Herzenstein discovered that the credit rating influenced the funding's success. He determined that peer-to-peer lending groups exhibit less discrimination and greater democratization than regulated institutional lending. The borrower attributes, such as ethnicity, cultural background, and gender identity, hold no significance in P2P platforms than in conventional financial institutions. In P2P markets, the lending community determines which loans get funds and which do not.

In his 2010 study, Herzenstein worked on herding behavior prevalent in investors while providing loans to borrowers. His findings revealed that investors become nervous, come under peer pressure, and give loans without much thinking, and this phenomenon is especially evident during the uncertainty in the P2P economy.

Furthermore, in auction-based lending platforms, the interest rate remains elevated because bidding drops once the loan gets financed. This situation confers a benefit to the lenders while imposing a drawback upon the borrowers.

One final determinant that impacts the financing of loans is the structure of the P2P platform. In their 2016 study, Wei and Lin discovered that funding arranged by peer-to-peer lending platforms employing an investment-driven approach is quicker than a bidding-driven approach. Nevertheless, using an investment-driven approach also has certain drawbacks. The P2P platforms determine the interest rates on loans in fundraising-based systems. As a result, fundraising-based systems have higher interest rates than auction-based systems. Increasing the interest rate for debtors is associated with an elevated likelihood of loan default (Wei & Lin, 2016). In his study, Carmichael discovered that the variables 'FICO score,' 'inquiries,' 'annual income,' and 'purpose' substantially impact loan default. (Carmichael, 2015).

Research Objective

The objective is to advance strong credit scoring models specifically designed for P2P lending platforms, utilizing state-of-the-art machine learning approaches. The research seeks to transform credit risk assessment in the P2P lending industry by improving the precision and dependability of credit scores. This project aims to enhance credit scoring and credit risk modeling procedures in P2P lending platforms by combining powerful machine learning algorithms with comprehensive data analysis, thereby fostering greater trust and efficiency within P2P lending platforms.

3. Research Methodology

3.1 Significance of the Study

This study is essential for a lot more reasons than just academic ones. It has real-world effects that are felt strongly in the financial sector. The results give credit managers in banks a plan for improving their companies' financial success by effectively managing credit risk. With the information gathered from this study, credit managers can change how they allocate their resources and find where reducing credit risk will have the most impact. By using these findings, banks can strengthen their financial bases, making them more stable and able to handle changing market conditions.

Additionally, investors will greatly benefit from the study's findings, as a deeper knowledge of credit risk will allow them to make intelligent choices to protect their investments against possible downturns. The study also guides everyone involved, explaining the complicated process of managing credit risk and encouraging everyone to work together to reach common financial goals. The study encourages people to act based on their knowledge, leading to an atmosphere of caution, foresight, and prosperity in the financial world.

3.2 Dataset Description

Lenders commonly use credit risk modeling for credit cards and consumer loans, the most prevalent retail goods. Here, we'll concentrate on consumer loans with various modeling-related issues. The data, which includes over 400,000 loans made through 2015, was provided by LendingClub, a well-known US P2P company. To begin with, quickly reviewing the important columns and going over the dataset is a great technique. The information includes the loan ID, amount, term, interest rate, lender details, and borrower details, which include purpose, employment period, yearly income, debt-to-income ratio, house ownership, credit rating, and default status. Loans show a combination of fully paid and defaulted payments.

4. Data Analysis and Interpretation

4.1 Pre-processing Independent Variables

In this section, we'll build the probability of the default (PD) model, followed by the credit scores for the borrowers. We will use logistic regression to construct the PD model as this is a regression model. The pre-processing will convert all independent variables into appropriate categorical dummy variables.

4.2 Examining and Imputing the Missing Values

To determine whether any data point in a data frame is missing, use the specialized pandas' method 'isnull.' If a value is missing, the isnull method returns True. We will examine the variables utilized in our study that have missing values. We can handle the missing values in two ways: either remove the rows containing the missing value or replace the missing value with a different value. We employ the pandas 'fillna' function to infer the missing values.

4.3 Automating Calculations: Discrete and Continuous Variables

To automate these computations and their simultaneous execution for every selected variable, a user-defined function is created to handle the calculations for any given categorical and continuous variable. This function will calculate the WoE value for each category of the independent variable. An efficient method to assess the WoE is to visualize it by loading the pyplot module from the matplotlib package to generate charts. A function named "graph_woe" will be created to represent the data visually, with the categories of the independent variables on the x-axis and the WoE on the y-axis.

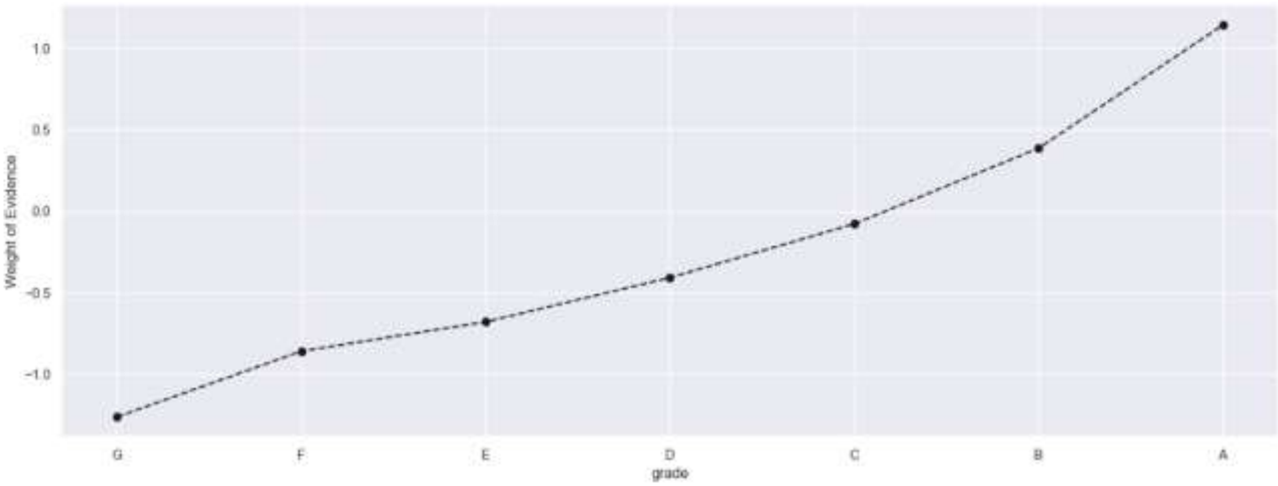

Figure 5: Weight of Evidence Plot for Grade Variable using graph_woe function

When including dummy variables in a regression model, one category will be selected as a reference category to evaluate the influence of the other categories on the outcome, and the category with the highest credit risk will be the reference category.

For example, as shown on the chart for the 'grade' variable, the WoE increases with the external credit rating grades from the worst 'G' to the best 'A.' It implies that loans with greater external ratings are better on average. Based on this WoE pattern, the grade variable's seven categories are retained as independent variables for the PD model. When incorporating the seven dummy variables into a regression model, designate one as a reference category to assess the impact of all others on the outcome. It is advisable to record which dummy variable will serve as a reference category when creating all dummy variables for each original variable. Always keep the category with the worst credit risk as a reference category. In this case, the dummy variable' G,' with the lowest WoE, will be the reference category.

4.4 Data Preparation: Splitting Data

We develop a model for the train data and save a portion of the data for subsequent use. Using the model, we evaluate its correctness and build a confusion matrix based on the test data. One approach to segregate the data into training and testing sets is employing the Sklearn Python module. The train_test_split method from the Sklearn model_selection needs to be set. We define test_size to a particular value, say 0.2. In such a situation, we employ twenty percent of the data for testing and eighty percent for training because too much data kept aside for testing results in the model being trained on less data.

4.5 Estimating the Probability of Default Model

We prefer a model like Logistic regression that is easily understandable and interpreted. The term "logistic regression" is used because the logistic function defines the curve that predicts the probability of an outcome. Interpreting logistic regression involves calculating the odds of an event occurring. These odds are determined by taking the exponential of the linear combination of the independent variables and their coefficients.

4.6 Importing Data and Choosing Features

First, we import the Numpy and Pandas libraries as np and pd, respectively. Next, we will import the data sets: train input and target. We will create the PD model by selecting the required independent variables. We utilize many categorical dummy variables in the PD model to represent each selected original variable. We identified a single reference category for each variable during the pre-processing stage. We store the reference variables in the `ref_categories' and remove them from the data frame using the drop method. The inputs_train data frame contains the remaining variables.

4.7 PD Model Estimation

For the PD model, we use Sklearn's logistic regression from Sklearn.linear_model and create an instance named 'reg.' We fit the model using input_train (independent dummy variables) and input_target (dependent variable), then store the results in 'reg.' The feature names are stored in a variable called feature_name.

To determine the statistical significance of each dummy variable's coefficient, we need multi-variate p-values, which Sklearn does not provide directly. To address this, we create a new LogisticRegression_with_P_values function that inherits from the original logistic regression class and overrides the fit method to compute and store p-values.

When interpreting the coefficients, we retain all dummies for that variable if any dummy variable coefficient p-value is less than 0.05, i.e., is statistically significant; this ensures the model remains conceptually sound. By applying this method, we maintain all significant dummy variables and remove those that are not, providing a robust final model.

4.8 Validating with Test Data

We use the test data to measure the PD model's efficacy by selecting only the relevant columns from the test data frame. We drop the reference categories, utilizing the same eighty-four variables we used in training the model. The default probability for each observation in the test data is estimated with y_hat_test equal to reg2.model.

The model in reg2 uses the predict method to calculate log-odds, which are exponentiated to get odds and derive the probability of non-default. A cut-off of 0.5 classifies observations: probabilities below 0.5 as bad or zero and the probability of 0.5 or higher as good or one.

4.9 Evaluation of Model Performance: Gini and Kolmogorov-Smirnov (K-S) Statistic

The Gini coefficient and the K-S statistic in credit risk modeling are employed to assess the model's performance. The Gini coefficient measures the disparity between good (non-defaulted) and bad (defaulted) borrowers by plotting their cumulative proportions. The Gini coefficient measures the accuracy of a model in distinguishing between default and non-default borrowers by plotting their cumulative proportions. It looks at the area between the line of perfect classification and the actual performance curve. A larger area indicates that the model is better at predicting defaults. The Gini index value for our model is around forty-one percent, which suggests that our model has an average ability to predict default and non-defaulted borrowers.

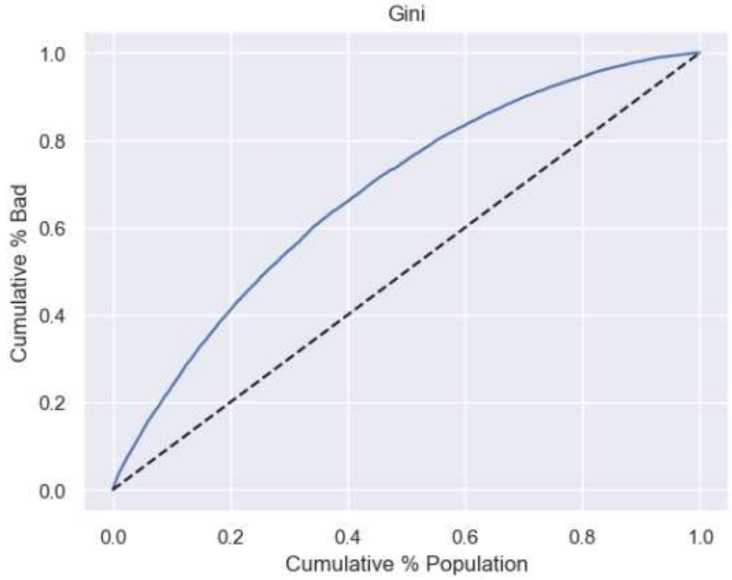

Figure 5: Gini Coefficient

The K-S statistic measures how well our model separates good and bad borrowers. It calculates the distance between two cumulative distribution functions. The further apart they are, the better the underlying feature differentiates them. A perfect model would score 1, while random chance would score 0. The estimated probability of being a good and bad borrower is plotted on the x-axis, with blue and red lines showing cumulative percentages of bad and good borrowers, respectively. A K-S statistic of 30% means our model has satisfactory predictive power.

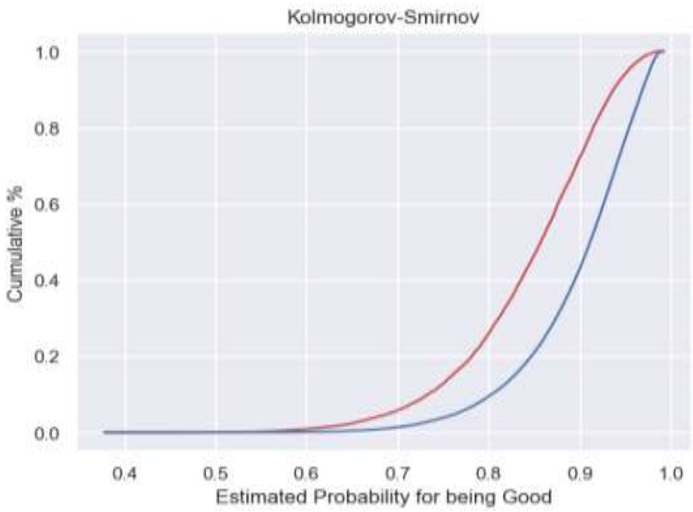

Figure 6: Kolmpgorov-Smirnov Statistic

4.10 Creating a Scorecard for Credit Assessment

We construct a scorecard for credit assessment within a probability of default (PD) modeling framework. The scorecard serves as a bridge between statistical insights and practical creditworthiness assessment. Utilizing regression coefficients from the PD model stored in summary_table. We define the scorecard's operational range, typically between 300 and 850. Rescaling creditworthiness assessments to fit this range involves discerning minimum and maximum assessments from the model and converting coefficients into scores using the formula below. The minimum and maximum sum of coefficients is -1.531 and 5.499.

$$\text{Variable score} = \text{Variable\_coef} * \frac{(\text{max\_score} - \text{min\_score})}{(\text{max\_sum\_coef} - \text{min\_sum\_coef})}$$

After calculating scores for all regression coefficients, we store the results in the 'Score_Calculation' data frame. Since the intercept is a constant part of the creditworthiness assessment, its corresponding score is naturally low. This score mirrors the minimum credit score a borrower would receive in the worst-case scenario. Hence, we aim to adjust the intercept's value to align closely with the minimum credit score, typically 300. To achieve this alignment, we utilize a specific equation to determine the score corresponding to the intercept.

$$\text{intercept\_score} = \frac{(\text{intercept\_coef} - \text{min\_sum\_coef}) * (\text{max\_score} - \text{min\_score})}{(\text{max\_sum\_coef} - \text{min\_sum\_coef})} + \text{min\_score}$$

We now have an intercept value of 312, which looks much more like we expected. We'll also check if the minimum and maximum possible scores equal the desired minimum and maximum scores.

4.11 Calculation of Credit Scores for an Individual Borrower

To determine a borrower's credit score, we need to add the credit scores associated with each category of the dummy variables. For this, we use the borrower with an index of 362514; we start with 313 because that's the score corresponding to the intercept. We see that the external grade category of the observation is 'c,' and we add the corresponding score of 54 from the scorecard to the intercept score and get a total of 367 points. Now, we will look at the house-ownership status, which is a mortgage, and we add a score of 8 from the scorecard, taking the total to 375 points. Knowing that the borrower will reside in California, we add a score of 5 points from the scorecard, bringing the total score to 380. We sum up the relevant categories to get a final credit score of 611 points. We follow the same procedure to arrive at the credit scores for the remaining observations. Similarly, we can calculate the credit scores for the remaining borrowers.

S.No	Feature Name	Coefficients	Score-Calculation	Category	Credit Score
1	Intercept	-1.30015	312	1	313
2	loangrade:C	0.691568	54	1	54
3	house_ownership:MORTGAGE	0.105745	8	1	8
4	stateaddress:CA	0.059683	5	1	5
5	loanpurpose:major_purch__car__ house_impr	0.266617	21	1	21
6	months_since_issue_date:40-41	0.711379	56	1	56
7	interest_rate:12.025-15.74	0.299441	23	1	23
8	months_since_earliest_credit_line: 165-247	0.033718	3	1	3
9	inquiry_last_6mths:0	0.657848	51	1	51
10	annual_income:60K-70K	0.223884	18	1	18
11	debt-to-inc:7.7-10.5	0.328779	26	1	26
12	months_since_last_delinq:Missing	0.110687	9	1	9
13	months_since_last_record:Missing	0.324867	25	1	25
				Total	**611**

Conclusion

This study emphasizes the notable progress that machine learning approaches contribute to credit risk modeling, specifically in Peer-to-Peer (P2P) lending platforms. Through a thorough examination of P2P lending, we have highlighted the shortcomings of conventional credit rating techniques and showcased the effectiveness of machine learning-driven models. We employed a comprehensive methodology that included meticulous data pre-processing, such as filling in missing values and transforming variables; this ensured a reliable basis for precise model estimation.

We utilized logistic regression, decision trees, random forests, and gradient boosting machines to create customized probability of default (PD) models designed explicitly for the P2P lending industry. Validation of these models using test data demonstrated their improved prediction skills, as seen by performance metrics such as Gini coefficients and Kolmogorov-Smirnov (K-S) statistics, which showed considerable enhancements compared to traditional techniques.

The results of this study have significant ramifications for the peer-to-peer lending sector. Through machine learning, lenders can attain more accurate evaluations of borrower risk, leading to a decrease in default rates and enhancing the general dependability of the platform. Consequently, this promotes increased confidence and involvement from both borrowers and investors, thus contributing to the long-term expansion of P2P lending platforms.

Furthermore, this study establishes a foundation for future research to investigate and enhance these models in greater detail. The evolution of machine learning and data analytics will inevitably lead to an expanded application of these techniques in credit risk modeling, resulting in improved accuracy and efficiency. Future research endeavors may prioritize incorporating supplementary data sources, investigate alternative machine learning techniques, and tackle growing obstacles in peer-to-peer lending.

Ultimately, incorporating sophisticated machine learning methods into credit scoring is a crucial advancement for the P2P lending sector. By adopting these advancements, P2P platforms can strengthen their strategies for managing risk, foster financial inclusion, and propel the next phase of expansion and progress in the financial technology industry.

References

1. Riza Emekter, Yanbin Tu, Benjamas Jirasakuldech & Min Lu (2015) Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending, Applied Economics, 47:1, 54-70, DOI: 10.1080/00036846.2014.962222
2. Governance and "slum" upgrading in cities in the South: best practice knowledge transfer, community organization, and knowledge generation, and Google. (n.d.). https://www.ippapublicpolicy.org/.
3. Douglas Cumming, Lars Hornuf, Marketplace lending of small- and medium-sized enterprises (September 2021), https://doi.org/10.1002/sej.1411
4. Zhang, B. Z., Baeck, P., Ziegler, T., Bone, J., & Garvey, K. (2016, January 1). Pushing Boundaries: The 2015 UK Alternative Finance Industry Report. Social Science Research Network. https://doi.org/10.2139/ssrn.3621312
5. Oren Rigbi, 2013. "The Effects of Usury Laws: Evidence from the Online Loan Market," The Review of Economics and Statistics, MIT Press, vol. 95(4), pages 1238-1248, October.
6. The Global Alternative Finance Market Benchmarking Report. (n.d.).
7. https://www.jbs.cam.ac.uk/.
8. The 2nd Global Alternative Finance Market Benchmarking Report. (n.d.).
9. https://www.jbs.cam.ac.uk/.
10. Zhang, H., Zhao, H., Liu, Q., Xu, T., Chen, E., & Huang, X. (2018, March 1). Finding potential lenders in P2P lending: A Hybrid Random Walk Approach. Information Sciences. https://doi.org/10.1016/j.ins.2017.12.017
11. Gao, Y., Yu, S. B., &Shiue, Y. (2018, July 1). The performance of the P2P finance industry in China. Electronic Commerce Research and Applications.
12. https://doi.org/10.1016/j.elerap.2018.06.002
13. Chen, X., Zhou, L., & Di-Fang, W. (2016, January 1). Group social capital and lending outcomes in the financial credit market: An empirical study of online peer-to-peer lending. Electronic Commerce Research and Applications.
14. https://doi.org/10.1016/j.elerap.2015.11.003
15. Milne, Alistair K. L. and Parboteeah, Paul, The Business Models and Economics of Peer-to-Peer Lending (May 5, 2016). ECRI Research Report, 2016, No 17, Available at SSRN: https://ssrn.com/abstract=2763682 or http://dx.doi.org/10.2139/ssrn.2763682
16. Sara Hsu, Jianjun Li, Hong Bao, P2P lending in China: Role and prospects for the future (2020) The University of Manchester and John Wiley & Sons Ltd
17. https://doi.org/10.1111/manc.12332
18. Governance and "slum" upgrading in cities in the South: best practice knowledge transfer, community organization and knowledge generation, and Google. (n.d.).
19. https://www.ippapublicpolicy.org/.
20. Lessmann, S., Baesens, B., Seow, H., & Thomas, L. C. (2015, November 1). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. European Journal of Operational Research. https://doi.org/10.1016/j.ejor.2015.05.030
21. Serrano-Cinca C, Gutiérrez-Nieto B, López-Palacios L (2015) Determinants of Default in P2P Lending. PLoS ONE 10(10): e0139427.
22. https://doi.org/10.1371/journal.pone.0139427
23. Jiang, C., Wang, Z., Wang, R. et al. Loan default prediction by combining soft information extracted from descriptive text in online peer-to-peer lending. Ann Oper Res 266, 511–529 (2018). https://doi.org/10.1007/s10479-017-2668-z
24. Aliano M., Alnabulsi K., Cestari G., Ragni S. (2023). Peer-to-Peer (P2P) Lending in Europe: Evaluating the Default Risk of Borrowers in the Context of Gender and Education. European Scientific Journal, ESJ, 19 (7), 60.
25. https://doi.org/10.19044/esj.2023.v19n7p60
26. Feng, Y., Fan, X., & Yoon, Y. (2015). LENDERS AND BORROWERS'STRATEGIES IN ONLINE PEER-TO-PEER LENDING MARKET: AN EMPIRICAL ANALYSIS OF PPDAI. COM. Journal of Electronic Commerce Research, 16(3), 242.
27. Ala'raj, M., &Abbod, M.F. (2015). A systematic credit scoring model based on heterogeneous classifier ensembles. 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA), 1-7.
28. Boon, K., Lim, Lo, C.Z., Yeo, S.F., Ling, C., & Tan Understanding of Peer-to-Peer Lending Platform Intention: Evidence among Millennials.