Decoding risk assessment: Exploring the efficacy of credit scoring vs. profit scoring in P2P lending

Marko Dávid Vateha[1]

Abstract: The conventional practice in the financial industry involves the assessment of creditworthiness through default scoring models, primarily focused on predicting the likelihood of default. However, this research paper challenges the status quo by introducing benefits of the concept of profit scoring – a novel approach that shifts the focus from mere default prediction to forecasting potential returns on a debtor when extending credit to clients. Using publicly available peer-to-peer (P2P) lending data, this study meticulously examines the efficacy of profit scoring compared to traditional default scoring methods. Through the implementation of machine learning (ML) algorithms, our analysis demonstrates that profit scoring outperforms default scoring in average return on portfolio of out-of-sample credit applications. In order to achieve comparable results, the paper uses the same family of algorithms for both tasks: profit scoring (predicting return on a loan) and default scoring (classifying defaults and non-defaults). Furthermore, the paper reveals that the factors influencing loan profitability exhibit disparities when compared to those driving default probabilities. The results underscore the transformative potential of profit scoring, offering lenders a more holistic perspective on risk assessment. By focusing on potential internal rate of return rather than default probability alone, financial institutions can make more informed decisions, leading to enhanced portfolio performance and profitability.

Keywords: P2P lending, loans, credit scoring, profit scoring
JEL Classification: G21, G32, C55

1 Introduction

In the fast-evolving landscape of financial services, the process of risk assessment plays an integral role in the extension of credit to customers. This assessment ensures that lending institutions strike a delicate balance between providing financial opportunities to creditworthy individuals and safeguarding their own interests. Traditionally, default scoring has been the linchpin of this assessment process, leveraging historical credit data and borrower profiles to predict credit risk. Default scoring methods primarily aim to predict the probability of default. Such models are trained on a sample of loans with a binary target variable – Default or Full repayment of all liabilities (Bluhm et al, 2016). This methodology has seen widespread use and has been well-documented in research, highlighting its significance in mitigating default risk (e. g. Bussmann et al, 2021).

However, the modern financial environment has brought in a surge of innovation that has led to the introduction of novel risk assessment methods. Traditionally, credit risk assessment relied on rule-based systems and statistical models. Then, in order to improve this process' accuracy, efficiency, and flexibility, machine learning (ML) techniques have been employed (Khandani et al, 2010). Among other innovations, profit scoring has recently garnered increasing attention (see Verbraken et al, 2014). This approach adopts a holistic perspective, extending beyond the binary notion of default risk to consider the potential profitability of loans (Crook et al, 2007). It seeks to balance risk and reward more effectively in lending decisions and optimize the allocation of capital. The implicit risk management strategy used by the profit scoring models aims to maximize loan portfolio profitability. With the help of the loan application's parameters, this method can forecast the annualized return on a loan in the future (e. g. Serrano-Cinca et al, 2016; Paula et al, 2019). Lenders can either focus on non-negative predictions of the profit scoring model, or work with minimal profit threshold they accept.

Peer-to-Peer (P2P) lending is a burgeoning financial phenomenon that has gained substantial attention in recent years (Wei et al, 2007). In the realm of financial intermediation, P2P lending represents a departure from traditional banking institutions. It facilitates direct lending interactions between individuals or entities seeking to borrow funds (borrowers) and those willing to lend capital (investors or lenders), often through online platforms. This disintermediated approach leverages technology to match borrowers and lenders efficiently, allowing individuals and small businesses to access financing while providing investors with opportunities for diversified lending portfolios (Bachmann et al, 2011).

[1] University of Economics in Bratislava, Faculty of Economics and Finance, Department of Banking and International Finance, Dolnozemská cesta 1, Bratislava 852 31 Slovak republic, MarkoDavid.Vateha@euba.sk

In light of these developments, this research paper embarks on an exploration of risk assessment in the world of P2P lending, aiming to compare the relative efficacy of default scoring and profit scoring methodologies. By scrutinizing real-life lending data and building upon prior research, we endeavour to provide a deeper understanding of which approach, credit scoring or profit scoring, yields superior results. In doing so, we hope to offer valuable insights that will empower decision-makers in navigating the multifaceted landscape of contemporary lending practices.

2 Methods

In the subsequent sections of this research paper, we delve deeper into the methodology employed for credit risk assessment, which midpoints on the application of machine learning techniques, particularly regularized linear regression models. We will provide a comprehensive exploration of the methodology's intricacies, including feature selection, model training, and evaluation metrics. Furthermore, we will elaborate on the dataset used, which comprises real-world lending data publicly disclosed by Lending Club, a prominent P2P lending platform in the United States. This rich dataset offers a valuable glimpse into the lending decisions and borrower profiles on the platform, enabling us to conduct a thorough empirical analysis of credit risk assessment practices in the context of P2P lending.

2.1 Statistical models

ML has significantly reshaped the landscape of credit risk prediction (e. g. Baesens et al, 2003; Gambacorta et al, 2019). Recently, it has been transitioning from the research desk to the application stack for credit scoring and a variety of other applications in credit risk after decades of resistance from supervision (Breeden, 2021). This paper explores the application of ML techniques, specifically focusing on two fundamental categories: linear and logistic regression. These methods, extensively studied and refined in the field, play a pivotal role in credit risk modelling. Linear regression is a versatile approach that strives to establish a linear relationship between a set of predictor variables and a continuous target variable, making it feasible to project possible returns on credit application (Khandani et al, 2010). On the other hand, logistic regression is tailored for binary classification tasks intrinsic to credit risk assessment (Dumitrescu et al, 2022).

Distinguishing these two regression variants is essential. While both employ a linear combination of input features, linear regression seeks to predict a continuous outcome, often used for tasks, such as expected internal rate of return on credit provided to a borrower (Khandani et al, 2010). In contrast, logistic regression, designed for binary classification, employs the logistic – sigmoid – function to transform the linear combination into probabilities, making it suitable for tasks like classifying borrowers as high or low credit risks (Dumitrescu et al, 2022).

To enhance the robustness and generalization of regression models, regularization techniques are crucial. One such technique is L1 regularization, often referred to as Lasso regularization. Lasso regularization plays a vital role in mitigating multicollinearity and overfitting. It achieves this by adding a penalty term to the linear regression or logistic regression objective function, forcing some coefficients to be exactly zero. This results in feature selection, where only the most informative variables are retained in the model, enhancing interpretability and potentially improving predictive performance (Friedman et al, 2010). It is believed L1 regularization within the context of credit risk modelling, provides additional value in managing high-dimensional financial data and bolstering predictive accuracy. As well, it is used to identify the most informative parameters in decision making processes, which is of concern to this research paper.

2.2 Data

The dataset employed in the paper provides an anonymized list of accepted credit application (both defaulted and fully repaid) with detailed flow of cash from every loan. It was publicly disclosed by Lending Club for its potential investors. Lending Club is well-known peer-to-peer (P2P) lending platform with headquarters in the US. It functions as an online marketplace that connects investors and borrowers, providing a simple substitute for conventional banking methods. Lending Club data is widely used for these purposes (e. g. Serrano-Cinca et al, 2015; Ye et al, 2018).

The data spans the years 2007 to 2018, during which time the platform issued 2.2 million loans. The dataset contains more than 150 characteristics about applicant at the time of application. The initial collection of records was reduced to 1.2 million loans with more than 530 columns to which we arrived via data transformation procedure.

Table 1 Descriptive statistics of annualized rate of return in the full data sample

			Annualized rate of return (in %)			
	Default rate	**Count of loans (in #)**	**Median**	**STD**	**10th percen-tile**	**90th percen-tile**
Full data	19.53%					
Non-Default		1,010,349	12.27	4.86	8.33	19.45
Default		245,252	-40.85	35.83	-97.97	1.27

Source: Own processing

2.3 Data sampling

In the pursuit of advancing credit risk modelling, this research endeavours to harness the power of data-driven methodologies. We initiated our investigation by selecting a substantial dataset comprising 1.2 million loan records. To ensure the robustness of our analysis, we employed a systematic random sampling with replacement strategy, generating six distinct samples of 100,000 loans each. Random sampling with replacement is a method where individual loans are randomly selected from a dataset, and after each selection, the loan is placed back into the dataset, allowing it to potentially be chosen again in subsequent selections (i. e. new random sample of loans). Within this framework, 20,000 loans were reserved for out-of-sample testing, serving as an independent litmus test for our models, while the remaining 80,000 loans were allocated for the rigorous training of our ML models.

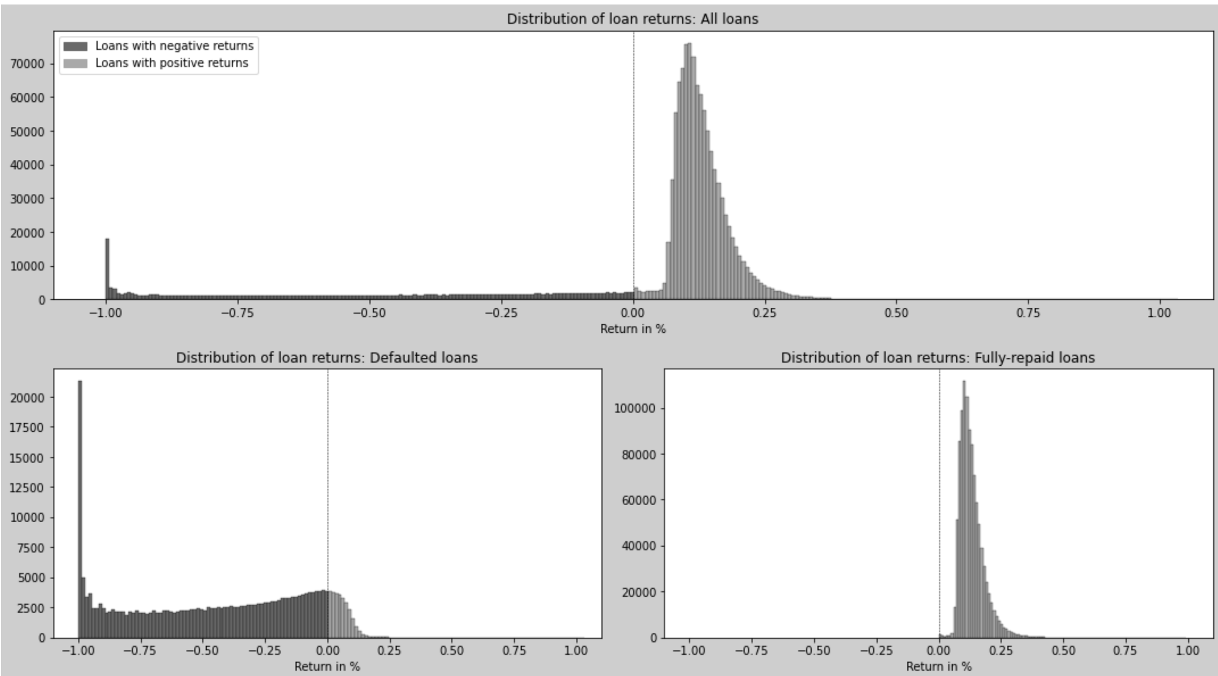
In the domain of default scoring, we employed logistic regression as the predictive model of choice. This statistical method has been widely acknowledged for its effectiveness in discerning the likelihood of loan defaults, aligning with established research in the field. In contrast, our exploration into profit scoring took a different route, focusing on a more holistic approach to lending assessment. Here, we trained our model to predict the Modified Internal Rate of Return (MIRR) which is recommended in the strand of the literature, e. g. Xie (2021). In MIRR calculation, we assume the initial release of funds from lending platform to a loan applicant to be a single payment transferred at the time of approval. Then, we estimate that the debtor (Lending Club) would be able to reinvest cash flow generated from repayments of loan at the median IRR of their loan portfolio realised in the previous 12 months, hence arriving to MIRR. With this method, we can calculate an annualized rate of return (i. e. MIRR) and compare loans with various maturities.

To evaluate the performance of our models, we conducted a multifaceted analysis. We gauged our models' efficacy by scrutinizing the mean and median values of profit derived from the out-of-sample portfolio. Furthermore, we scrutinized the overall financial performance of our models, considering the total earnings achieved relative to the credit extended within our sampled universe. These assessments culminate in a comprehensive view of model performance, shedding light on their effectiveness in minimizing default risk and maximizing the financial returns for lending institutions in real-world scenarios.

3 Research results

The paradigm of profit scoring stands out as a strong rival in the field of credit risk assessment because it provides a nuanced viewpoint that goes beyond traditional default scoring's binary framework. The motivation behind this paradigm change is definitely strong, as shown by empirical data that contradicts accepted thinking. The discovery of a subset of "defaulted" loans that, intriguingly, show a positive annualized return on investment is a key finding of our investigation. It emphasizes that credit institutions' main goal is to maximize the performance of their portfolio by utilizing the potential profitability present in some loans, rather than merely categorically identifying default.

Figure 1 Profit distribution of P2P lending loan portfolio of Lending Club



Source: Own processing

Figure 1 delves deeply into the loan portfolio's annualized rate of return distribution. Dark grey is used to indicate loans that have negative internal rates of return. The lower panel offers a split perspective on the segment of defaulted (left side) and fully repaid loan (right side) while the upper panel shows distribution of the entire portfolio of accepted credit applications. The drill-down on non-performing loans shown in the left bottom panel may tempt some investors to target even defaulted loans because some of them are profitable (light grey).

As a result, profit scoring plays a crucial part in empowering institutions to make well-informed lending decisions that not only reduce risk but also tap into hidden opportunities for profit. In order to maximize profits in the dynamic and competitive world of modern lending, this holistic perspective puts standard credit assessment procedures to the test and forces a re-evaluation of how creditworthiness is determined.

3.1 Profit scoring vs. Default scoring

Figure 1 reveals a crucial insight that disrupts the accepted norms of credit risk assessment: there is a subset of defaulted loans that not only generates profitability but also attractive returns for lending institutions. Our research builds on this intriguing finding by using a machine learning architecture that includes logistic regression (LogReg) for default scoring and linear regression (LinearReg) for profit scoring, both supported by L1 lasso regularization to effectively zero out unnecessary parameters. This methodological uniformity allows for a thorough and directly comparative investigation of default and profit scoring. To ensure robustness, we apply these models consistently to six randomly drawn samples, each with replacement, from our extensive dataset of 1.2 million loans.

Table 2 Descriptive statistics of annualized rate of return in the full data sample

	(in %)	MIRR (in %) for entire portfolio			(in mil. $)
	Invested loans	**Median**	**Mean**	**STD**	**Total profit**
Credit scoring - LogReg					
Sample 1	61.63	7.80	2.88	17.27	8.84
Sample 2	61.56	7.70	2.79	17.12	8.64
Sample 3	61.48	7.75	2.87	17.30	8.77
Sample 4	61.57	7.79	2.89	17.08	8.54
Sample 5	61.60	7.70	2.90	17.00	8.93
Sample 6	62.10	7.84	2.84	17.24	8.87
Profit scoring - LinearReg					
Sample 1	71.43	8.97	3.26	19.40	10.84
Sample 2	69.71	8.61	3.06	19.04	10.36
Sample 3	71.75	8.94	3.17	19.62	10.81
Sample 4	71.10	8.87	3.06	19.65	10.13
Sample 5	69.49	8.67	3.25	18.84	10.67
Sample 6	69.92	8.72	3.11	19.04	10.16

Source: Own processing

A convincing conclusion emerges from the analysis: profit scoring regularly produces greater median and mean returns on loan portfolios, frequently outperforming default scoring by about 10%. Those findings are presented in Table 2 numerically. Higher standard deviations, which are a sign of higher risk, are present along with this improved profitability, however. Notably, this increase in profits appears to be largely attributable to the approval of more loan applications, highlighting the critical role that profit scoring plays in streamlining lending decisions and boosting financial returns in the complex and dynamic environment of modern credit risk assessment.

3.2 Importance of parameters in credit applications

Furthermore, in our comprehensive analysis, we delved into the inner workings of our credit risk models for both profit and default scoring. Our exploration sought to identify the key parameters driving lending decisions, shedding light on the factors that contribute to loan approval or denial. Notably, our findings revealed intriguing disparities between these two models in terms of the composition of the most influential parameters derived from loan applications. It is crucial to emphasize that all parameters underwent a rigorous normalization process to ensure comparability, with their coefficient values reflecting relative importance. However, it is important to acknowledge that direct comparisons of coefficient amplitudes across the models are not feasible due to their distinct target variables – default vs. annualized return – each operating on different scales.

Table 3 Importance of 7 most important parameters in Default scoring and Profit scoring models

	Default scoring (Logistic regression)		**Profit scoring (Linear regression)**	
Descending order	**Parameter**	**Value of coefficient**	**Parameter**	**Value of coefficient**
1	Interest rate on loan	0.48	Interest rate on loan	-0.03
2	Term length	0.17	Home (mortgage)	0.02
3	FICO score	-0.15	Listed for whole amount founding	-0.02
4	Home (mortgage)	-0.09	Term length	-0.01
5	Debt-to-Income	0.09	Home (rent)	-0.01
6	Loan amount to annual income	0.03	Debt-to-Income	-0.01
7	Home (rent)	0.001	Employment length over 10 years	0.01

Source: Own processing

The analysis presented in Table 3 highlights notable differences in parameter importance between default scoring and profit scoring models. While the disparities between these two approaches are not substantial, there are some small yet intriguing nuances that shed light on their distinct objectives and priorities.

In the context of default scoring, the FICO score emerges as a one of the most significant determinants, with a negative coefficient, indicating its pivotal role in assessing credit risk. The lower the FICO score, the higher the risk of the default of the applicant for a credit (target variable default is classified as 1, hence, an inversely proportional relationship between them). FICO score is popular in US and is based on an individual's credit history and provides lenders with a numerical representation of their credit risk, aiding in the decision-making process for extending credit. Meanwhile, interest rate on loans and term length hold considerable importance, aligning with conventional credit assessment factors. Interest rate assigned to the credit application according to internal risk categories of Lending Club is the most important distinctive factor.

In contrast, the profit scoring model, focused on maximizing financial returns, exhibits slightly different parameter priorities. Interest rate on loans, still the most impactful, carries a negative coefficient, suggesting that a lower interest rate is preferred for achieving higher returns, at least on average. Home-related variables, such as mortgage and rent, play a pivotal role in profit scoring (although they appear also in default scoring model, here in profit scoring, they are more dominant), underscoring the significance of property ownership in influencing lending decisions. Profit scoring comparing to credit scoring weights on length of employment of credit applicant. Length of employment provides insights into a borrower's financial stability, reliability, and commitment to meeting financial obligations, making it an essential factor in credit decision-making. A surprising variable to stand out in profit scoring model was a parameter about funding. When a loan is "Listed for whole amount founding" it means that the loan was initially listed in the so-called whole market. In this context, the term "whole" signifies that the loan was proposed to investors with the intention of being fully funded all at once. This has a negative influence on overall profitability. To comprehensively understand how this variable influences profitability, further investigation into its inner mechanics and its impact on loan performance will require additional time and analysis.

4 Conclusions

In conclusion, the paradigm of profit scoring emerges as a formidable contender in the realm of credit risk assessment, offering a nuanced viewpoint that transcends the standard binary framework of default scoring. The motivation behind this paradigm shift is undeniably strong, backed by empirical data that challenges accepted norms. Notably, the discovery of a subset of "defaulted" loans that yield positive annualized returns is a key finding of our investigation (see Figure 1), emphasizing that credit institutions' primary goal is to maximize portfolio performance, not merely categorically identify default.

The results are compelling, as the profit scoring consistently outperforms default scoring, often yielding higher median and mean returns on out-od-sample loan portfolios by approximately 10%. However, it is important to note that this enhanced profitability comes with higher standard deviations, indicative of increased risk. Significantly, this profit boost appears to be largely driven by the approval of more loan applications, underscoring the pivotal role that profit scoring plays in streamlining lending decisions and enhancing financial returns in the dynamic world of credit risk assessment.

In our detailed analysis, we also delved into the inner mechanics of both credit risk models, uncovering disparities in parameter importance between default and profit scoring. While the differences are not substantial, they illuminate the distinct objectives and priorities of these models. Factors such as the FICO score, interest rates, and term length play essential roles in default scoring, while profit scoring places greater emphasis on interest rates, home-related variables, and the length of employment. Additionally, the variable "Listed for whole amount founding" emerged as a noteworthy factor in profit scoring, indicating its influence on overall profitability.

In summary, our research challenges conventional credit assessment procedures and underscores the critical importance of profit scoring in enabling institutions to make well-informed lending decisions that maximize returns. The nuanced insights gained from this study have the potential to reshape credit risk assessment practices and drive more profitable lending strategies in the modern financial landscape.

Acknowledgement

This paper is an output of the science project VEGA 1/0221/21 Interest rates in the environment with central bank digital currency.

References

Bachmann, A., Becker, A., Buerckner, D., Hilker, M., Kock, F., Lehmann, & M., Funk, B. (2011). Online peer-to-peer lending-a literature review. *Journal of Internet Banking and Commerce*, 16(2), 1.

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. (2003). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the operational research society*, 54, 627-635.

Bluhm, Ch., Overbeck L., Wagner Ch. (2016). *Introduction to credit risk modeling*. CRC Press, Boca Raton (2016).

Breeden, J. (2021). A survey of machine learning in credit risk. *Journal of Credit Risk*, 17(3).

Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics*, 57, 203-216.

Crook, J. N., Edelman, D. B., & Thomas, L. C. (2007). Recent developments in consumer credit risk assessment. *European Journal of Operational Research*, 183(3), 1447-1465.

Dumitrescu, E., Hué, S., Hurlin, C., & Tokpavi, S. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, 297(3), 1178-1192.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1), 1.

Gambacorta, L., Huang, Y., Qiu, H., & Wang, J. (2019). How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *BIS Working Papers* No 834. December, 2019.

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34(11), 2767-2787.

Paula, D., Artes, R., Ayres, F., & Minardi, A. (2019). Estimating credit and profit scoring of a Brazilian credit union with logistic regression and machine-learning techniques. *RAUSP Management Journal*, 54, 321-336.

Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. *PloS one*, 10(10), e0139427.

Serrano-Cinca, C., & Gutiérrez-Nieto, B. (2016). The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending. *Decision Support Systems*, 89, 113-122.

Verbraken, T., Bravo, C., Weber, R., & Baesens, B. (2014). Development and application of consumer credit scoring models using profit-based classification measures. *European Journal of Operational Research*, 238(2), 505-513.

Wei, Z., & Lin, M. (2017). Market mechanisms in online peer-to-peer lending. *Management Science*, 63(12), 4236-4257.

Xie, M. (2021). Research on the modified internal rate of return. *Turkish Journal of Computer and Mathematics Education* (TURCOMAT), 12(11), 4087-4090.

Ye, X., Dong, L. A., & Ma, D. (2018). Loan evaluation in P2P lending based on random forest optimized by genetic algorithm with profit score. *Electronic Commerce Research and Applications*, 32, 23-36.