

A Hybrid Approach for Predicting Probability of Default in Peer-to-Peer (P2P) Lending Platforms Using Mixture-of-Experts Neural Network

Christopher Watitwa Makokha¹, Ananda Kube², Oscar Ngesa³

¹Department of Mathematics, Pan African University Institute of Basic Sciences, Technology and Innovation, Nairobi, Kenya

²Department of Mathematics and Actuarial Sciences, Kenyatta University, Nairobi, Kenya

³Department of Mathematics, Statistics and Physical Sciences, Taita Taveta University, Voi, Kenya

Email: cmakoks2009@gmail.com, kube.ananda@ku.ac.ke, oscangesa@gmail.com

How to cite this paper: Makokha, C.W., Kube, A. and Ngesa, O. (2024) A Hybrid Approach for Predicting Probability of Default in Peer-to-Peer (P2P) Lending Platforms Using Mixture-of-Experts Neural Network. *Journal of Data Analysis and Information Processing*, 12, 151-162.

<https://doi.org/10.4236/jdaip.2024.122009>

Received: January 27, 2024

Accepted: April 23, 2024

Published: April 26, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Peer-to-peer (P2P) lending offers an alternative way to access credit. Unlike established lending institutions with proven credit risk management practices, P2P platforms rely on numerous independent variables to evaluate loan applicants' creditworthiness. This study aims to estimate default probabilities using a mixture-of-experts neural network in P2P lending. The approach involves coupling unsupervised clustering to capture essential data properties with a classification algorithm based on the mixture-of-experts structure. This classic design enhances model capacity without significant computational overhead. The model was tested using P2P data from Lending Club, comparing it to other methods like Logistic Regression, AdaBoost, Gradient Boosting, Decision Tree, Support Vector Machine, and Random Forest. The hybrid model demonstrated superior performance, with a Mean Squared Error reduction of at least 25%.

Keywords

Credit-Scoring, Clustering, Classification, Neural Networks

1. Introduction

Credit refers to money borrowed from financial institutions, which must be repaid with interest, typically in installments [1]. Predicting the risk associated with loans is imperative for financial institutions. This risk represents the likelihood that borrowers will default and fail to repay the entire loan amount. Ac-

curate risk assessment ensures profitability for the institution while avoiding overburdening customers beyond their financial capacity [1]. Credit scoring involves placing individuals on a scale of creditworthiness, indicating their trustworthiness regarding credit. This scale guides decisions on granting credit to applicants [2]. Crowdfunding encompasses web-based fundraising, where users contribute small amounts to finance projects [3]. Crowdfunding can be either commercial, with expected financial returns, or non-commercial, relying on donations [3]. Peer-to-peer (P2P) lending falls under commercial crowdfunding, where online platforms collect funds from the “crowd” to collectively finance higher-value loans for individuals or businesses [4].

Most classification algorithms operate in a batch mode, directly processing the dataset in memory. However, more sophisticated methods often create large data structures, resulting in a significant memory footprint. This can hinder their application to larger datasets. Even when these datasets can be accommodated, the algorithm’s complexity may lead to excessively long classification times. To mitigate these issues, one could allocate additional memory or extend the experiment duration. However, both options come with associated costs in terms of time and money. Furthermore, neither approach guarantees a solution if the dataset exceeds memory capacity at its maximum [5].

This research addresses data mining for modeling credit risk default probabilities while minimizing information loss from large datasets. The proposed approach involves a mixture-of-experts neural network framework to assess default probabilities in P2P lending platforms. By combining unsupervised and supervised machine learning (ML) methods, one study evaluated credit risk for commercial customers. Hybrid models outperformed individual supervised ML models, particularly in complex scenarios with nonlinear predictor-target relationships [6]. Building on these findings, the study anticipates that preceding mixture-of-experts neural networks with an unsupervised clusterer will enhance overall model accuracy.

This study holds significance because it addresses the critical issue of credit risk, which directly impacts financial institutions’ lending capabilities. Given the widespread challenge of borrowers struggling to service their debt, modeling default probabilities becomes crucial for devising effective strategies within peer-to-peer lending platforms. Such efforts have the potential to revolutionize the P2P lending industry by enhancing loan repayment outcomes through informed policy decisions. Furthermore, this research contributes to the existing computer science literature by developing novel ML algorithms or techniques applicable to lending data. These innovations may find broader applications across various industries and fields of research. Future researchers can leverage these findings to extend the body of knowledge in this domain.

2. Literature

A Hybrid Expert Neural Network model is a type of artificial neural network that combines different types of neural networks or other machine learning

models to solve complex problems [7]. The structure and working principle of such a model can vary depending on the specific models being combined and the problem at hand. However, a common feature in many hybrid expert neural network models is the use of an expert function and a gating function.

The expert function in a hybrid expert neural network model refers to the individual models or “experts” that are trained to handle specific subtasks within the overall problem. These experts can be any type of model, including but not limited to neural networks, and they are typically designed to be highly specialized in their respective subtasks [7].

The gating function, also known as the router, is responsible for determining which expert(s) should be used for a given input. This is typically done by assigning weights to the outputs of the experts based on the input, effectively controlling the influence of each expert on the final output.

The gating function is a key component of the model because it allows the model to dynamically adapt to different inputs by leveraging the strengths of its various experts. The gating function itself is often implemented as a neural network, and it is trained alongside the experts. The training process involves adjusting the parameters of the gating function and the experts to minimize the difference between the model’s output and the actual output for a set of training data.

A hybrid expert neural network model leverages the strengths of multiple expert models and uses a gating function to dynamically adapt to different inputs. This makes it a powerful tool for tackling complex problems that may be difficult to solve with a single model [7].

Here are some advantages and disadvantages of Hybrid Expert Neural Networks (HENN) models compared to other models:

Advantages:

- Leveraging Strengths: HENNs can leverage the strengths of various neural network types, such as Convolutional Neural Networks (CNNs) for spatial data and Recurrent Neural Networks (RNNs) for sequential data [8].
- Improved Performance: By combining different models, HENNs often achieve better performance on complex tasks that may be beyond the scope of a single model [8].
- Flexibility: They offer flexibility in design, allowing for the creation of custom networks tailored to specific tasks [8].
- Efficiency: HENNs can be more efficient in learning representations from data, as they can process diverse data types and handle complex patterns [8].

Disadvantages:

- Complexity: The increased complexity of HENNs can make them harder to design, train, and optimize [8].
- Interpretability: Like many neural networks, HENNs can act as “black boxes,” making it difficult to interpret how they make decisions [8].
- Computational Resources: They may require more computational resources and data for training due to their complexity [8].

- Risk of Overfitting: With a large number of parameters, there's a risk of overfitting the model to the training data [8].

Hybrid learning combines unsupervised and supervised learning methods. The process occurs in two stages. Initially, an unsupervised method identifies data clusters. In the subsequent stage, data transformation occurs based on the learned clusters, and a supervised learning algorithm constructs a function to differentiate between various class labels [7].

Clustering serves two primary purposes:

- Data Exploration: Clustering methods identify similarities between instances and group them accordingly when such groups are formed, they can be named, and their attributes defined. For instance, one can choose the group mean as the representative prototype for instances within the group, or describe the possible attribute range. This simplifies data understanding [9].
- Preprocessing Stage: Clustering can also be used as a preprocessing step. An advantage of preceding a supervised learner with an unsupervised clustering is that the latter doesn't require labeled data. Labeling data is expensive. By using a large amount of unlabeled data to learn cluster parameters, we can subsequently use a smaller labeled dataset for the second stage of classification or regression [9].

West [10] investigated the credit scoring accuracy of five neural network models. These included multi-layer perceptron, mixture-of-experts, radial basis function, learning vector quantization, and fuzzy adaptive resonance. Real-world datasets were split into training and independent test sets using 10-fold cross-validation. The results were compared to linear discriminant analysis, logistic regression, k-nearest neighbor, kernel density estimation, and decision trees. The study recommends considering Mixture-of-experts (MOE) and Radial Basis Function (RBF) neural network models. The MOE model's superiority may stem from its ability to partition the input space and assign local expert networks to learn specific sub-spaces. This partitioning reduces learning interference, leading to a network that converges to more accurate locations in the error surface [10].

Byanjakar *et al.* [11] proposed a credit scoring model using artificial neural networks to classify peer-to-peer loan applications into default and non-default groups. They utilized a publicly available dataset from Bandora, a leading European P2P lending platform. The dataset consisted of 15.77% defaulted loans and 84.23% non-defaulted loans. The neural network architecture included:

- Input layer: 14 neurons, corresponding to the number of independent variables in the data.
- Hidden layer: 5 neurons.
- Output layer: 1 neuron.

During training, the model achieved the following classification accuracy:

- Training data set: Non-default loans: 64.47% and default loans: 74.75%.
- Testing data set: Non-default loans: 62.70% and default loans: 74.38%.

The study's findings suggest that financial attributes play a more influential role than demographic attributes in determining credit risk. Moreover, the

neural network outperformed logistic regression in classifying default loans (74.38% vs. 61.03%). However, it is important to note that this study focused on a single P2P lending case. Byanjakar *et al.* [11] posited that future research should explore similarities and differences across various P2P lending cases operating in diverse financial environments.

In their research, [12] conducted an in-depth analysis of credit scoring methods, including discriminant analysis, logistic regression, neural networks, and Classification and Regression Trees (CART). Their study was based on a dataset from a Turkish bank, which included credit card information. The dataset, comprising 1260 entries, contained nine predictive variables for each customer: gender, age, marital status, level of education, occupation, job position, income, customer type, and ownership of credit cards from other banks. The outcome variable was the creditworthiness of the customer, categorized as good or bad credit. The artificial neural network model they developed included 13 neurons in the hidden layer. The findings of the study revealed that CART outperformed the other methods in terms of the average correct classification rate. However, the neural network model for credit scoring demonstrated fewer Type II errors, which are typically associated with higher misclassification costs, thereby indicating superior overall credit scoring performance.

Fernandes [13] conducted a study comparing Support Vector Machines (SVM) classification with several clustering methods to distinguish human from non-human users on Twitter. Their goal was to identify normal human activity within a random sample of 1000 accounts. The clustering techniques employed included K-Means, DBSCAN, and Gaussian Mixture Models. Here is a breakdown of their approach:

- Initial Classification: SVM with a Radial Basis Function (RBF) kernel was used. The feature set initially consisted of 70 variables. Feature selection reduced complexity and improved generalization performance.
- First Stage Classification: Both classification and clustering approaches achieved similar f_1 accuracy scores of 90%. However, they encountered challenges when classifying human users who exhibited abnormal behavior.
- Second Stage Classification: A second classification step further separates non-human users into categories; brands, celebrities, and promoters/information. The average f_1 accuracy achieved in this stage was 74%.

In summary, both classification and clustering methods effectively separated human and non-human users, but clustering had the advantage of not requiring pre-classified data.

Wang [14] explored multi-class cost-sensitive classifiers for credit rating using data from Lending Club. They selected four well-known classifiers—C4.5, Naive Bayes, Logistic Regression, and SVM—as base models for the cost-sensitive meta-algorithm. The cost matrices included uniform costs, linear costs, and partitioned-linear costs. Key findings include:

- Cost-sensitive SVM performed best in 2017Q2, indicating that predictive

performance varied based on both the cost-sensitive classifier and cost matrix.

- The study emphasizes the importance of considering cost-based performance when selecting a cost-sensitive classifier and assigning costs to minimize losses.

3. Methodology

There are two steps in this model *i.e.* clustering and classification respectively.

In the clustering phase, we apply the Mixture Models clustering technique to reduce large data sets to a size where classification algorithms can be run on them with minimum loss of information vital from those data sets.

In the classification phase, we apply Mixture-of-experts Neural Networks on the output of the clustering phase to classify the defaulting status of an individual.

Let y_i be the outcome variable specifying the defaulting status of individual i . Assuming that X_1, X_2, \dots, X_p are features that are collected together with y_i . The interest is to use the features X_1, X_2, \dots, X_p to make the classification.

3.1. Step 1: Mixture Models Clustering

Let X_1, X_2, \dots, X_n denote a random sample of size n , where X_j is a p -dimensional random vector with probability density function $f(x_j)$ on \mathbb{R}^p .

X_j contains the random variables corresponding to p measurements made on the j th recording of some features of the phenomenon under study.

Let $X = (X_1^T, \dots, X_n^T)^T$, where superscript T denotes vector transpose.

We suppose that the density $f(x_j)$ of X can be written in the form $f(x_j)$ is as follows.

$$f(x_j) = \sum_{i=1}^g \pi_i f_i(x_j) \quad (1)$$

where $f_i(x_j)$ are densities and the π_i is non-negative quantities that sum to one; that is

$$0 \leq \pi_i \leq 1; (i = 1, \dots, g) \quad (2)$$

and

$$\sum_{i=1}^g \pi_i = 1 \quad (3)$$

The quantities π_1, \dots, π_g are called the mixing proportions or weights.

The $f_i(x_j)$ is called the component densities of the mixture.

In this formulation of the mixture model, the number of components g is considered fixed.

The mixture density $f(x_j)$ can then be written as

$$f(x_j; \Psi) = \sum_{i=1}^g \pi_i f_i(x_j; \theta_i) \quad (4)$$

where the vector Ψ containing all the unknown parameters in the mixture

model can be written as

$$\Psi = (\pi_1, \dots, \pi_{g-1}, \xi^T)^T \quad (5)$$

where ξ is the vector containing all the parameters in $\theta_1, \dots, \theta_g$ known in advance as distinct.

With a mixture model-based approach to clustering, it is assumed that the data to be clustered are from a mixture of an initially specified number g of groups in various proportions. That is, each data point is taken to be a realization of the mixture density Equation (4), where the g components correspond to the g groups. On specifying a parametric form for each component density $f_i(x_j; \theta_i)$, the vector Ψ can be estimated by maximum likelihood (or some other method).

Once the mixture model has been fitted, a probabilistic clustering of the data into g clusters can be obtained in terms of the fitted posterior probabilities of component membership for the data. An outright assignment of the data into g clusters is achieved by assigning each data point to the component to which it has the highest estimated posterior probability of belonging.

3.2. Step 2: Mixture-of-Experts Neural Network Classification

The task at hand is now simplified to using r_1, r_2, \dots, r_k for the classification of y_i . In this context, we employ a neural network classification algorithm known as the mixture-of-experts. This algorithm comprises K experts and a coordinating unit referred to as a gating network, which serves as a mediator among the expert networks. The assumption here is that different networks excel in different areas of the input space, hence the necessity for the gating network. Considering a regression problem where a regressor r yields a response represented by the random variable D , with a specific realization of this variable denoted by d , we opt for a scalar form of regression for ease of representation. Given the scalar nature of the regression problem, each expert network is composed of a linear filter [7].

As shown in **Figure 1**, the output produced by expert k is the inner product of the input vector r and synaptic weight vector w_k of this neuron.

$$y_k = w_k^T r, k = 1, 2, \dots, K \quad (6)$$

The gating network is structured with a single layer composed of K neurons, where each neuron is designated to a particular network. The neurons within the gating network exhibit non-linearity, and their activation functions are determined by

$$g_k = \frac{\exp(u_k)}{\sum_{j=1}^k \exp(u_j)} \quad (7)$$

where u_k is the inner product of the input vector r and synaptic weight vector a_k .

$$u_k = a_k^T r, k = 1, 2, \dots, K \quad (8)$$

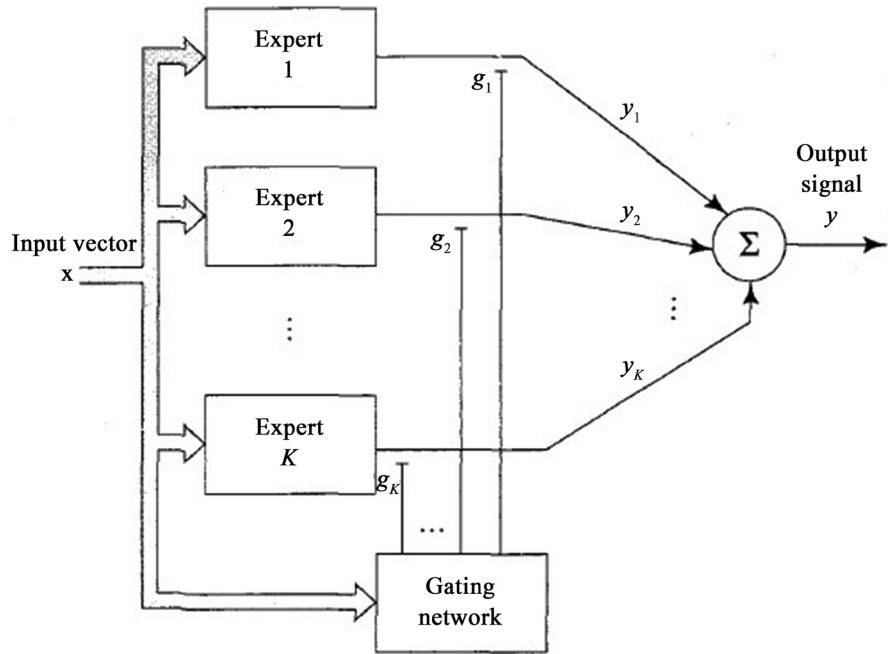


Figure 1. Mixture-of-experts neural network 1 [7].

g_k satisfies the following conditions

$$0 \leq g_k \leq 1 \tag{9}$$

$$\sum_{k=1}^K g_k = 1 \tag{10}$$

Let y_k denote the output of the k th expert in response to the input vector r .

The overall output of the MOE model is

$$y = \sum_{i=k}^K g_k y_k \tag{11}$$

where g_k is a nonlinear function of r .

Consider an MOE model with two experts, and a gating network with two outputs denoted by g_1 and g_2 .

The output of g_1 is defined by

$$g_1 = \frac{\exp(u_1)}{\exp(u_1) + \exp(u_2)} = \frac{1}{1 + \exp(-(u_1 - u_2))} \tag{12}$$

Let a_1 and a_2 denote the two weight vectors of the gating network.

$$g_1 = \frac{1}{1 + \exp(-r^T (a_1 - a_2))} \tag{13}$$

$$g_2 = 1 - g_1 = \frac{1}{1 + \exp(-r^T (a_2 - a_1))} \tag{14}$$

Thus both g_1 and g_2 are in the form of a logistic function.

The orientation of g_1 is determined by the direction of the difference vector $(a_1 - a_2)$, whereas the orientation of g_2 is determined by the direction of the difference vector $(a_2 - a_1)$.

Along the ridge defined by $a_1 = a_2$ we have $g_1 = g_2 = \frac{1}{2}$ and the two experts contribute equally to the output of the MOE model.

Away from the ridge, one or the other of the two experts assumes the dominant role.

4. Results

4.1. Description of the Data

The dataset, sourced from Kaggle, encompasses details about previous loan applicants and their default status. The objective is to discern patterns that suggest a likelihood of default. Such patterns could inform decisions like refusing the loan, decreasing the loan amount, or imposing a higher interest rate on loans to risk-prone applicants. The dataset comprises 1,382,352 records, each with 149 characteristics, gathered from 2007 through the fourth quarter of 2018. The large number of attributes provides a detailed view of the borrowers' profiles, which can lead to more accurate predictions of loan defaults. Having a wide range of attributes helps in identifying the most significant predictors of default and can improve the performance of the classification models [15].

Moreover, the use of such detailed data can aid in the design of financial products like credit derivatives, which can be used to hedge against the risk of default [15]. This approach to credit risk analysis and the subsequent financial product design is crucial for the sustainability and profitability of peer-to-peer lending platforms.

The selected time frame is significant because it includes a variety of economic conditions, from a major recession to a recovery phase, and periods of both expansion and contraction [16]. This diversity can enhance the robustness of a model by training it on data from different phases of the economic cycle, which is essential for creating a model that can generalize well to future conditions [16].

Table 1 summarizes the performance of classification algorithms preceded by unsupervised clustering. The results indicate that preceding a classifier with an unsupervised clusterer improves the performance. Furthermore, there was at least 25% improvement in the performance of the mixture-of-experts neural network model by preceding it with an unsupervised clusterer.

4.2. Discussion

The provided table displays metrics for both individual and hybrid models during validation. These models were assessed using various metrics, including accuracy, precision, recall, F1-score, and mean-squared error. Notably, the mixture-of-expert neural network model achieved the highest accuracy of 0.92, followed by the logistic regression model with an accuracy of 0.89. Conversely, the decision tree model had the lowest accuracy at 0.78. Additionally, the precision, recall, F1-score, and mean-squared error metrics exhibited variations across the different models.

Table 1. Metrics for validation: individual and hybrid models.

Clusterer	Classifier	Accuracy	Precision	Recall	f1-score	MSE
	LR	0.86	0.79	0.74	0.88	0.7548
	AdaBoost	0.89	0.80	0.78	0.89	0.7264
	Gradient Boosting	0.88	0.84	0.80	0.86	0.8288
	Decision Tree	0.78	0.85	0.82	0.85	0.8587
	Random Forest	0.77	0.76	0.84	0.82	0.8326
	SVM	0.80	0.82	0.85	0.81	0.7213
	MOE	0.81	0.83	0.86	0.80	0.6258
Mixture Models	LR	0.89	0.75	0.80	0.86	0.6385
	AdaBoost	0.86	0.83	0.79	0.87	0.7385
	Gradient Boosting	0.81	0.81	0.80	0.88	0.8325
	Decision Tree	0.80	0.84	0.91	0.81	0.8690
	Random Forest	0.79	0.86	0.90	0.82	0.8630
	SVM	0.82	0.90	0.88	0.85	0.7018
	MOE	0.92	0.90	0.86	0.86	0.4321

The hybrid expert neural network model was chosen for its ability to handle complex and varied data through a combination of specialized networks.

The empirical contributions of this study results to the risk management and decision-making of P2P lending platforms are significant. They provide valuable insights that can help platforms:

- Identify Key Risk Factors: Research helps in identifying the variables that significantly influence the risk associated with lending, such as loan interest rates, loan duration, and borrower characteristics [17].
- Improve Interest Rate Setting: Studies have shown how interest rates can be adjusted based on loan size, duration, and other factors, which is crucial for balancing risk and return [17].
- Enhance Regulatory Frameworks: Empirical findings can guide regulators and policymakers in designing responses to market challenges, particularly during economic crises like the COVID-19 pandemic [17].
- Facilitate Platform Management: Insights from research encourage government participation in platform management, promoting a healthier development of the P2P lending industry [17].

As for the impacts on related fields:

- Financial Inclusion: P2P lending research contributes to financial inclusion by providing credit access to underserved populations, supporting small businesses, and fostering economic growth [15].
- Economic Resilience: It extends the coverage of credit circulations and improves economic resilience, even during financial crises [15].
- Competitiveness: Findings affect the competitiveness of P2P platforms by

highlighting factors like volume, operation time, and investor numbers that give platforms a competitive edge [18].

In essence, empirical research in P2P lending not only enhances the operational strategies of lending platforms but also has broader implications for financial inclusion, economic stability, and the competitive landscape of financial services.

Based on the results, we recommend further research in the following areas:

- Investigating the impact of different feature selection techniques on the performance of the models.
- Exploring the use of deep learning models for this classification task.
- Evaluating the performance of the models on larger datasets to determine their scalability.
- Comparing the performance of the models with other state-of-the-art machine learning algorithms.

Overall, the results suggest that the mixture-of-expert neural network model is the best-performing model, followed by the logistic regression model. However, the choice of model will depend on the specific needs of the user and the context in which the model will be used.

Based on the findings, we recommend further research in the following areas:

- Feature Selection Techniques: Investigate the impact of different feature selection methods on model performance. Understanding which features contribute most significantly can enhance model accuracy.
- Deep Learning Models: Explore the use of deep learning models for this classification task. Techniques like neural networks with multiple layers may offer improved predictive capabilities.
- Scalability Evaluation: Assess the models' performance on larger datasets to determine their scalability. Ensuring robustness across varying data sizes is essential.
- Comparison with State-of-the-Art Algorithms: Compare the performance of the models with other advanced machine learning algorithms. Understanding how they fare against established methods provides valuable insights.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Alpaydin, E. (2004) Introduction to Machine Learning (Adaptive Computation and Machine Learning Series). Vol. 14, The MIT Press, Cambridge.
- [2] Blanco, A., Pino-Mejías, R., Lara, J. and Rayo, S. (2013) Credit Scoring Models for the Microfinance Industry Using Neural Networks: Evidence from Peru. *Expert Systems with Applications*, **40**, 356-364. <https://doi.org/10.1016/j.eswa.2012.07.051>
- [3] Lenz, R. (2017) Peer-to-Peer Lending: Opportunities and Risks. *European Journal of Risk Regulation*, **7**, 688-700. <https://doi.org/10.1017/S1867299X00010126>

- [4] Ariza-Garzón, M.J., Camacho-Miñano, M.D.M., Segovia-Vargas, M.J. and Arroyo, J. (2021) Risk-Return Modelling in the p2p Lending Market: Trends, Gaps, Recommendations and Future Directions. *Electronic Commerce Research and Applications*, **49**, 1-16. <https://doi.org/10.1016/j.elerap.2021.101079>
- [5] Evans, R., Pfahringer, B. and Holmes, G. (2011) Clustering for Classification. 2011 *7th International Conference on Information Technology in Asia: Emerging Convergences and Singularity of Forms - Proceedings of CITA'11*, Sarawak, 12-13 July 2011, 1-8. <https://doi.org/10.1109/CITA.2011.5998839>
- [6] Machado, M.R. and Karray, S. (2022) Assessing Credit Risk of Commercial Customers Using Hybrid Machine Learning Algorithms. *Expert Systems with Applications*, **200**, 116889. <https://doi.org/10.1016/j.eswa.2022.116889>
- [7] Guti, P.A. (2011) Hybrid Artificial Neural Networks: Models. 177-184. https://doi.org/10.1007/978-3-642-21498-1_23
- [8] Albahri, A.S., Alnoor, A., Zaidan, A.A., Albahri, O.S., Hameed, H., Zaidan, B.B., Peh, S.S., Zain, A.B., Siraj, S.B., Masnan, A.H.B. and Yass, A.A. (2022) Hybrid Artificial Neural Network and Structural Equation Modelling Techniques: A Survey. *Complex and Intelligent Systems*, **8**, 1781-1801. <https://doi.org/10.1007/s40747-021-00503-w>
- [9] Baştanlarm Y. and Ozuysal, M. (2014) Introduction to Machine Learning. In: Yousef, M. and Allmer, J., Eds., *miRNomics: MicroRNA Biology and Computational Analysis. Methods in Molecular Biology*, Vol. 1107, Humana Press, Totowa, NJ, 105-128. https://doi.org/10.1007/978-1-62703-748-8_7
- [10] West, D. (2000) Neural Network Credit Scoring Models. *Computers and Operations Research*, **27**, 1131-1152. [https://doi.org/10.1016/S0305-0548\(99\)00149-5](https://doi.org/10.1016/S0305-0548(99)00149-5)
- [11] Byanjankar, A., Heikkila, M. and Mezei, J. (2015) Predicting Credit Risk in Peer-to-Peer Lending: A Neural Network Approach. *Proceedings - 2015 IEEE Symposium Series on Computational Intelligence (SSCI 2015)*, Cape Town, 7-10 December 2015, 719-725. <https://doi.org/10.1109/SSCI.2015.109>
- [12] Ince, H. and Aktan, B. (2009) A Comparison of Data Mining Techniques for Credit Scoring in Banking: A Managerial Perspective. *Journal of Business Economics and Management*, **10**, 233-240. <https://doi.org/10.3846/1611-1699.2009.10.233-240>
- [13] 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**, 74-82. <https://doi.org/10.1016/j.elerap.2017.12.006>
- [14] Fernandes, M.A., Patel, P. and Marwala, T. (2015) Automated Detection of Human Users in Twitter. *Procedia Computer Science*, **53**, 224-231. <https://doi.org/10.1016/j.procs.2015.07.298>
- [15] Weinstock, R.L. (2023) Introduction to U.S. Economy: The Business Cycle and Growth-Congressional Research Service.
- [16] Moran, O. and Ido, T. (2021) An Empirical View of Peer-to-Peer (P2P) Lending Platforms. *Berkeley Business Law Journal*. (Forthcoming)
- [17] Frigui, H. (2008) Clustering: Algorithms and Applications. 2008 *First Workshops on Image Processing Theory, Tools and Applications*, Sousse, 23-26 November 2008, 1-11. <https://doi.org/10.1109/IPTA.2008.4743793>
- [18] Yan, X.X. and An, Q.S. (2020) Research on the Factors Affecting the Competitiveness of P2P Online Loan Platform. *Advances in Economics, Business and Management Research*, **126**, 270-276. <https://doi.org/10.2991/aebmr.k.200306.047>