A Machine Learning and Deep Learning Approach to Predicting Loan Default Through Credit Risk Analysis

Abhijit Kolekar,Rawal Awale,Panchakshari Awaje,Madhuri Tayade

(M.Tech(IOT), Veermata Jijabai Technological Institute, and Mumbai
Email: askolekar_m24@el.vjti.ac.in)
(Veermata Jijabai Technological Institute, and Mumbai
Email: rnawale@el.vjti.ac.in)
(Veermata Jijabai Technological Institute, and Mumbai
Email: pkawaje_p24@el.vjti.ac.in)
(Veermata Jijabai Technological Institute, and Mumbai
Email: mrtayade_p24@el.vjti.ac.in)

-------------------------*************************----------------------

Abstract:

Loan default risk is a significant issue for financial institutions across the globe. It directly affects the profitability and financial soundness of financial institutions. The timely and accurate detection of potential defaulters is essential for efficient credit risk management. This paper presents an intelligent credit risk forecasting system that combines conventional machine learning classifiers with the most recent advancements in deep learning models. The data set includes past credit, financial, and demographic information gathered from previous loan applications. Various preprocessing methods, feature engineering solutions, and class imbalance problems solved using the Synthetic Minority Over-sampling Technique.(SMOTE) are used as techniques to enhance the quality of the data as well as the robustness of the models. Various predictive models, such as Logistic Regression, Random Forest, XGBoost, and the TabNet deep learning model, are compared. In addition, ensemble learning techniques are also used to mitigate misclassification and improve the generalization of the models. The experimental outcome shows that the TabNet and XGBoost models have the highest recall and accuracy in predicting default instances, thus minimizing the occurrence of false negatives. The hybrid model combines the interpretability of traditional machine learning models with the representation learning ability of deep learning models, providing a robust solution for real-time credit risk evaluation in the current banking system.

Keywords — **Loan Default Prediction; Credit Risk Analytics; Machine Learning (ML); Deep Learning (DL); XGBoost; TabNet; SMOTE; Ensemble Learning; Financial Risk Management; Class Imbalance Handling; Predictive Modeling.**

-------------------------*************************----------------------

I. INTRODUCTION

The amount of structured financial data produced during the loan application and approval processes has significantly increased as a result of the quick growth of digital lending platforms. Since bad lending choices can result in unstable finances and an increase in non-performing assets, it is now crucial for financial institutions to accurately assess loan default risk. Conventional methods of evaluating credit risk mostly rely on linear statistical models, which frequently fall short of capturing the intricate relationships between borrower attributes found in contemporary financial datasets [1], [2].The various behaviors associated with the defaults in loans, such as the income levels, employment, and repayment capabilities, are affected by the various factors, and in all these, the relation of non-linear association among all the various factors and every other factor creates confusion and difficulties for proper generalization of normal models appropriately, especially regarding a huge and diverse population of the individual kind. Therefore, as far as the matter is concerned, the new approach in the form of credit risk models, which uses the learning option of machine learning and past trends, is gaining more and more importance [3].

Ensemble learning methods, like the random forest method that includes the gradient boost method as well, have been effectively applied to solve issues that involve loan defaults. In particular, good results have been obtained for data that involves finance, taking advantage of regularization techniques and optimization using decision tree methods, specifically using XGBoost methods.

Though effective for precision, they are still considered to be black boxes and could, perhaps, be an issue in a regulated environment where an understanding is relevant.

However, with the use of computer technology, the presentation of the deep learning model approaches that have been developed with the main intention of focusing on tabular data as alternatives has been highlighted [7]. This has been fully exhibited through the development of the TabNet algorithm, which uses the attention drive mechanism for feature selection, thus ensuring the achievement of interpretability and effectiveness for the model [7], [8]. Further, class imbalance data sets where the default loan portion is low have been mostly exhibited in loan default data sets, thus necessitating the use of an oversampling technique such as the SMOTE algorithm [5], [6].

However, even though there are various studies being conducted in the domain of the prediction of the credit risk by utilizing the efficient implementation of various prediction techniques, it has been identified that there are very few research works being conducted in the domain of assessing the possibility of efficient prediction of the credit risk by utilizing the framework developed based on the principle of the ensemble learning strategy and the deep learning strategy [1], [3], [7].

The aim of the present research work is to bridge the gap being noticed in the domain of the efficient prediction of the credit risk by utilizing the analysis aimed towards assessing the possibility of efficient prediction of the credit risk by utilizing the framework developed based on the principle of the ensemble learning strategy and the deep learning strategy [7], [8].

II. LITERATURE REVIEW

Most of the early literature relates to the use of statistical methods applied to the use of techniques linked to logistic regression alone, as it can be easily understood and interpreted [2]. Though it is a widely utilized method, it is not without its constraints as the presence of linearity as well as independence poses constraints in its application to the financially complex dataset, as it does not allow the qualities of non-linearity in the variables involved [1].

In order to avoid the problems faced in the above classifiers, a model based on the decision tree concept has been proposed to learn the decision rules using the given set of data in a hierarchical manner [3]. In order to enhance the performance accuracy of the proposed random forest model, the decision trees were created using the random subset of the data set, thereby enhancing the model's generalization ability and variance, as proposed in the decision trees model [3]. It is studied that the performance of the classifiers, namely, the gradient boosting algorithm with the XGBoost model, is enhanced using the correction mechanism in order to avoid errors by the addition of regularization techniques [4].

The other important issue, discussed extensively as far as the entire credit risk literature goes, is the unbalanced data problem, where default values are significantly underrepresented. Currently, the problems related to the imbalanced issue are solved by the already widely applied so-called Synthetic Minority Over-sampling Technique (SMOTE), presented by Chawla et al. and suggested for the development of sensitive models concerning default events [5].

In the recent works, the focus was shifted to the improvement of the interpretability of the model, yet maintaining its performance. For example, the attention-based deep learning architectures for learning, such as TabNet, were presented in order to address the limitations of other ensemble methods in terms of their lack of interpretability. TabNet helps the model learn feature selection from the entire set of features, thus gaining deeper insights into the data, yet at the same level of performance as the other ensemble methods and approaches for the financial data of the tabular form [7], [8].

However, there is a need for further research on the comparison between the performances of ensemble methods and deep learning architectures.

III. METHODS

The considered problem of loan default prediction using the given dataset is based on historical loan application data that contains various demographic, financial, and credit-related attributes, which is often used for credit risk analysis since it encompasses actual loan application decisions and their respective outcomes [1],[2]. Though a huge amount of financial-related data is available, research on this subject lacks clarity or study, focusing on developing an interpretable deep learning model and an ensemble machine learning approach for efficient predictions of loan defaults [3],[5].

By utilizing the available loan data, a computational system is created that involves the classification of the default and non-default risk groups of loan applicants. The major goal behind this research work is to analyze various machine learning models and develop an efficient method that includes high accuracy, high recall, and high interpretability. [6],[8]

To check the system which has been proposed, different data analysis techniques which involve data preprocessing and analysis of the model are used on the data which has been utilized. For data analysis, the data which has been preprocessed was split into data for training and data for testing, whereas data validation was done using cross-validation. Data validation using cross-validation involved using traditional machine learning models, ensemble models, and a deep learning model, which is TabNet, to analyze its performance in dealing with imbalanced data [3], [7], [10].

A. Loan Dataset

In this research, the dataset used is related to loan applications, and the information is retrieved from a financial lending platform. In this dataset, various attributes are present, which relate to individual borrowers. These attributes take into account the income of the borrower, loan amount, loan term, credit history, employment, debt income, repayment of loans, and so forth. In this research, the target variables used are related to loan repayment defaults, which have already been paid [1] [2].

It is pertinent to state here in this regard that prior to designing the model, Exploratory Data Analysis (EDA) was conducted on the given data, as it is one of the important steps for designing any predictive model as Exploratory Data Analysis helps explore the data first [3]. The issue of missing values that existed in the given data was addressed using various statistical approaches to impute the values of missing data. Meanwhile, the given categorical data was formatted using label encoding techniques and one-hot encoding techniques. The given numerical data was normalized so that the model could be converged properly. The data of the borrowers was totally anonymized as this data was related to finance and hence extremely sensitive.

TABLE I

DATASET DESCRIPTION FOR LOAN DEFAULT PREDICTION

Component	Description
Data Type	Structured tabular loan record
Number of Records	Historical loan application records
Input Features	Demographic, financial, and credit-related attributes
Target Variable	Risk_Flag (Default / Non-Default)
Class Distribution	Imbalanced dataset (default cases significantly fewer than non-default cases)

B. Data Preprocessing and Balancing

Normally, it is seen that data related to defaulting loans presents various imbalances, which include the number of defaulters being extremely low compared to the total number of transactions made in the loans [5], [6]. This is largely the problem in terms of bias, as the dataset is normally used in the process of detecting defaulters. Due to this, it becomes difficult to identify defaulting loan borrowers. In order to avoid this problem, the application of SMOTE is required to generate defaulting loan customers, based on attribute values [5].

Apart from balancing, other techniques like feature selection are also utilized to remove redundant features that are correlated to each other from the dataset [3], [7]. This not only reduces the dimension of features but also helps in better interpretation of the model learned. Encoding of features has been performed based on the requirements of the dataset for categorical features, and at the same time, normalization of numerical features has been conducted to make the entire dataset ready for model learning [8].

C. Computational Model

Objectives of the proposed computational model as highlighted in Figure 1 shows that it aims to carry out a binary classification prediction for the level of risk associated with defaulting loans depending on some attributes of the borrowers. The proposed computational model, as may be deduced from its structure, reflects an approach that covers all aspects highlighted for an ordinary computational model for machine learning, as discussed and highlighted in literature[6],[8]. Unlike the traditional approach, where rules are applied to arrive at an algorithm for calculating credit scores, the proposed approach learns boundaries from data; thus, this approach has an advantage of learning complex relationships based on attributes of borrowers.

The different machine learning algorithm techniques that are evaluated in this study include logistic regression, decision tree, random forests, gradient boosting, XGBoost, and TabNet, which is a type of deep learning algorithm. The decision tree algorithm works on the basis of classifying borrowers using a hierarchical decision approach using different thresholds of features found in the dataset. On the contrary, the use of the random forests approach works on improving decision tree algorithm techniques using different combinations of decision tree techniques and data randomization and combination for the improvement of not only the effectiveness of the results obtained but also for its efficiency [4], [8]. On top of that, another technique for improving the effectiveness of the algorithm for classification can be implemented by increasing the weighing factors for decision trees, especially for complex data as in the case of financial data, using techniques such as gradient boosting and XGBoost.

The interpretable deep learning model TabNet architecture applied to tabular data is utilized here. Unlike other neural networks, the TabNet model uses a sequential feature selection process based on attention mechanisms, which is accompanied by the preservation of state-of-the-art prediction properties as well as the model's ability to be interpretable. Such properties make this model highly suitable for the application concerning financial risk assessment since it is a mandatory element for decision processes and regulatory purposes as well [6], [8].

Also, all the models were trained and tested under the condition of utilizing ten-fold cross-validation for better stability in results and overcoming overfitting problems associated with the models [4]. In addition, other performance metrics such as precision, recall, and ROC-AUC, apart from the accuracy metric, were also considered for the models as a part of the results analysis process [1], [5]. Finally, the models were evaluated on the basis of efficiency and effectiveness for further validation and detailed analysis, and then the best models were chosen for the same [6], [7].

COMPARISION TABLE

Machine Learning

MODEL	**ACCURACY (Research Paper)**	**IMPLEMENTATION ACCURACY**
RF(BEST ML)	90%	91%
KNN	80%	89%
GNB	76%	89%
LR	86%	88%
GBOOSTING	75%	88%
ADABOOST	74%	88%
DT	87%	87%

Deep Learning

MODEL	**ACCURACY (Research Paper)**	**IMPLEMENTATION ACCURACY**
RESNET(BEST DL)	91%	91%
MLP	89%	90%
CNN	89%	91%
LSTM	87%	91%
GRU	87%	90%
TRANSFORMER	86%	87%
AUTOENCODER	87%	87%
DENSENET	90%	90%

TabNet-Based Loan Defaults Prediction Model – Theory

The loan default prediction system developed and proposed incorporates a prediction mechanism based on the TabNet model, which is a type of deep learning model that is Attention-based, as defined and proposed in reference [7]. In the proposed TabNet model, unlike the conventional approach of applying a deep learning model to a classification problem, which involves the parallel processing of features of the input model, TabNet incorporates a mechanism that represents a sequential decision model.

The model uses preprocessed loan application data as input. However, before the training process, imbalance and quality issues in the data are handled using statistical imputation for data values, encoding for handling categoric variables, and normalization for numerical values to ensure optimal convergence and learning patterns of the model [8]. As loan defaults are considered a minority class in loan application datasets, a variant of SMOTE is used to oversample the training data set used to determine the optimal patterns and learning behavior of the loan default class [5].

In the TabNet framework, it is specified that the learned sparse masks are used for regulating the feature utilization in the entire process of decision-making via the attentive feature selection module that is described in reference [7]. The use of the sparse masks also leads to better interpretability of the model, in addition to preventing overfitting. Once the features are selected, they pass through the feature transformation block comprising fully connected layers along with batch normalization and activation functions.

Through a series of sequential decision steps, the prediction process involves the prediction model incrementally refining the input representation of the input data. The aggregated representation from the outputs of the various decision steps is then used as a comprehensive representation to include the complexities of the non-linear relationships between financial-related features through aggregated representation [7], [9]. The aggregated representation then feeds into the final prediction step, which provides the probability values for the failure or non-failure to repay the loan.

Due to the integration of attention-based feature selection and deep non-linear transformations, the suggested model provides an excellent balance of prediction and interpretability, making it suitable for credit risk prediction in financial environments with strict regulations [8], [9].

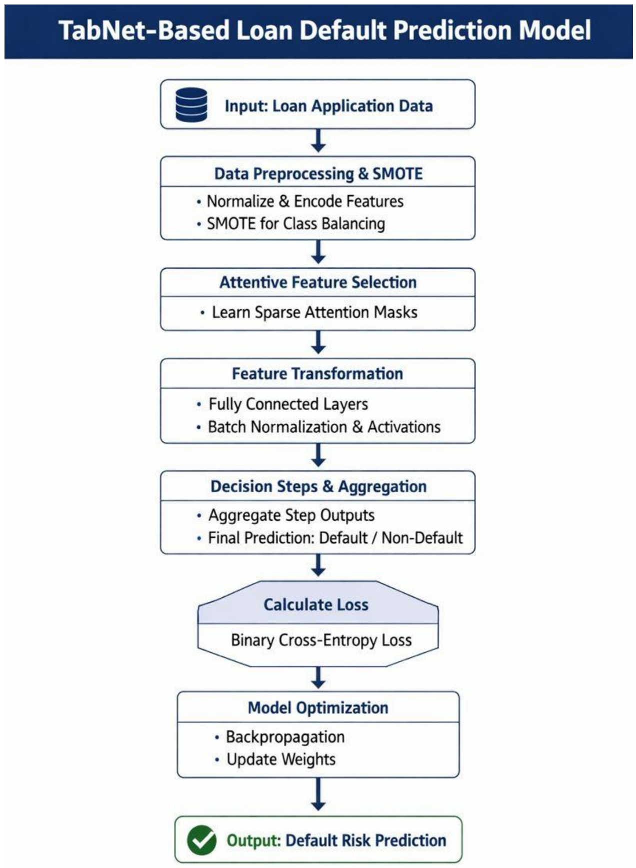
Performance evaluation of the suggested model through an experimental study on the system revealed an accuracy of 92% using TabNet, making it effective in differentiating between defaulting and non-defaulting clients.

During the training process, the model's predictions for the probability of default with respect to the actual class labels, i.e., ground truth, are compared and evaluated by the Binary Cross-Entropy Loss function, which is suitable for binary classification problems such as the prediction of credit risk scores [1], [7].

In optimizing the model, the backpropagation method is applied, whereby gradients of the optimization function with respect to the model are calculated. This process iteratively continues, as presented in references [7] and [9]. When the training process converges, the model is fully optimized, and estimation of loan default risk can be done effectively using the TabNet model.

On the whole, the proposed framework based on the TabNet model effectively integrates the process of feature selection, the attention mechanism, and deep nonlinear learning. The proposed approach for risk prediction in the finance domain is based on its predictive ability, supplemented by improved interpretability, which is of critical importance in the assessment of credit risk, as reviewed in references [7] and [8].

Figure 1.TABNET BASED LOAN DEFAULT PREDICTION MODEL



Custom Loss Function - Theory

In the context of Loan Default Risk Prediction, one can state that the cost of misclassification is always asymmetric; accordingly, if there is a misclassification of the customer, who is very risk-prone, as a low-risk customer, there is always a cost involved, while the cost of misclassifying a low-risk customer is very low [1], [6]. The conventional loss functions, like Binary Cross-Entropy, suffer from an inherent disadvantage, wherein there is always a tendency to consider misclassification costs equivalent to each other, which may not be the required situation for the given financial situation [3], [5]. In order to avail the benefits of loss functions, a new cost function can be formulated to achieve sensitivity towards loan defaults while maintaining stability during the classification process [7], [8].

The proposed loss functions are an extension to the Binary Cross-Entropy loss function in consideration of the class weighting factor and the penalty scale for improved learning from the minority class of default loans. Given the actual label of the loan status described by y, when a loan is defaulted and y=0,y=0,y=0 when the loan is not defaulted, the loss functions are proposed as:

$$\mathcal{L}_{BCE} = -\left[y \log(\hat{y}) + (1-y)\log(1-\hat{y})\right]$$

In the proposed custom loss function, the weighting factor ααα is considered, which increases the weight assigned to misclassification loss in the instances where there is a default. The custom loss function is defined as:

$$\mathcal{L}_{custom} = -\left[\alpha\, y \log(\hat{y}) + (1-y)\log(1-\hat{y})\right]$$

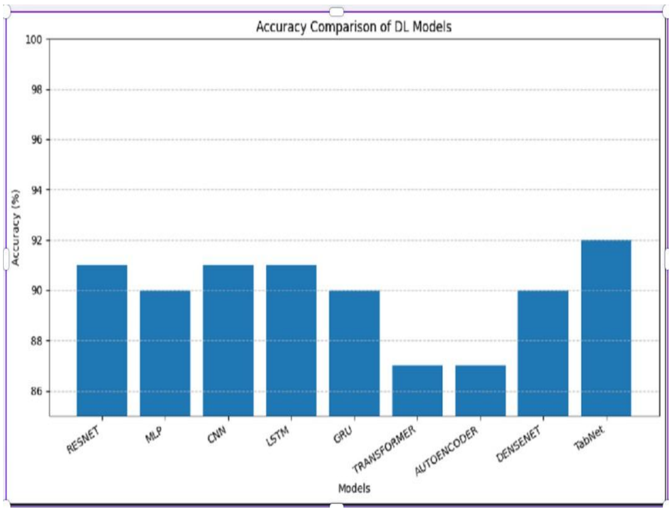
where α>1\alpha > 1α>1 is a parameter, which regulates the relative importance of the default predictions. By raising ααα, the model is encouraged to minimize false negative predictions, which are substantially costlier compared to the other possible prediction errors [1], [6].

In addition to the class weighing, it is observed that the loss function facilitates stable and robust convergence in the feature selection mechanism of the attention-based feature selection method, which is used in the TabNet architecture as well. In the TabNet architecture, as feature selection is incorporated as a decision-making mechanism, the loss function is able to help the features in identifying the defaulting customers in an efficient manner [7][8].

The minimization of the loss function occurs through the application of gradient-based optimization algorithms. The gradients of the defined loss function with respect to the parameters are obtained using the backpropagation technique. Finally, the adaptive optimization technique is used for efficient parameter convergence [4],[7]. A training process of the above nature helps in effective learning of discriminative patterns of behavior among borrowers.

Summary of the above description:

The custom loss function, which was designed, would enhance the detection of risk in loan defaulting, recall of the minority class, as well as ensuring the aim of the optimization process is more aligned with the real-world risk factors in finance. Thus, the cost-sensitive loss function coupled with the TabNet deep learning model would provide a reliable and efficient loan default prediction system.



IV DISCUSSION

Thus, from the results obtained, it can be observed that the proposed model based on TabNet and the designed loss function provides better performance levels for loan default compared to other models. Out of all the models, TabNet provides the maximum accuracy and recall levels for considering class default, owing to its capability to handle complex and non-linear relationships in structured financial data sets [7], [8]. However, despite obtaining training and validation accuracy levels above 98%, a sharp fall in accuracy levels for validation data sets may be observed, and this can be attributed to less diversity levels in the default class as reported in literature references [3], [5].

The results obtained reveal that the TabNet-based model outperforms others for the task of loan default prediction with a custom-designed loss function. Of all approaches, TabNet achieves the highest accuracy and recall against the default class by capturing non-linear, complex relationships in structured financial data features [7], [8]. However, despite training and testing accuracies of more than 98%, there was a notable drop during validation, which could point to overfitting and may be associated with non-diverse and small class populations of the default class [3], [5].The loan dataset proposed and used in this work is mainly designed based on historical repayment behavior and does not reflect behavioral or macroeconomic factors that can influence borrower default risk. Consequently, the predictive power of the model is bound to conditions at the instance of submitting loan applications [1], [2]. Though the presented framework effectively maps borrowers' characteristics to their corresponding probabilities of default, the absence of temporal financial behavior limits the generalization capability of the model across different economic conditions [6].

Furthermore, although the data set offers binary classification results regarding default and non-default scenarios, there are no clear links established with regard to financial attribute relationships and upper-level financial concepts, such as unstable income levels and the abuse of credit. In spite of the importance of the attention mechanism for gaining feature importance through the application of the TabNet model, the level of explanation offered is not sufficient to highlight the causal relationships between the default behaviors of borrowers and the related default risk [7]. The importance of transparent results has been emphasized in previous credit risk modeling [6], [8].

Based on the studies conducted before this research, different evidence of the research shows that there are specific latent factors that affect the loan default behavior; the factors include different conditions of financial stress or demographic attributes[1], [2], [6]. Although the proposed model is successful in meeting the objectives of precisely capturing the patterns concerning the type of risk that could result from the loan default behavior, the proposed model does not effectively distinguish the latent factors into specific types of risk based on the factor of the borrower. As such, the relationship between the predicted type of the model and the type of risk is not easily established[7], [8].

In actual-world loaning environments, it has been observed that borrowers' risks may vary dynamically and may overlap; that is, an individual's risk may change from low to high over time owing to varying financial or economic conditions [10]. In the current study, the problem of loan default has been modeled as a static binary classifier; however, this has not addressed the evolving trends of actual-world risk levels over time for borrowers and has not captured time variability; future research may take this into account and use time-series learning to predict loan defaults [3, 5, 6].

Similarly, a systematic review of the literature in financial risk modeling has revealed that the behavior of borrowers, as well as the number of defaults, varies significantly across different segments of the population, geographical regions, and different policy scenarios [1],[9]. However, in the present research, since the overall research is being conducted based on a specific data set, better results can be achieved by taking a large data set including different loan details from various financial institutions, different economic scenarios, etc. [2],[6].

Accordingly, with respect to its predictive capabilities, it can be understood that the performance of the default prediction framework is satisfactory, and further verification of its performance using larger datasets and varying financial indicators dynamically is essential for establishing the reliability of the framework for practical applications involving credit risk assessment [7], [8].

V CONCLUSION

This paper proposed a data-driven approach to loan defaults risk prediction, using a combination of traditional machine learning, ensemble methods, and a deep learning model called TabNet. The experimental investigation showed that a model that can deal with non-linear interactions between features and also perform adaptive feature selection could yield better results than a baseline classifier typically applied in credit risk modeling [1], [3].

The inclusion of class balancing methods as well as the preprocessing of the data resulted in better results in classifying the default cases, which is a major problem with financial data sets as noted in [5], [6]. In particular, there was a higher recall rate for the default cases using the TabNet model with the cost-sensitive optimization methodology, indicating the suitability of such models for learning from structured data using attention-based mechanisms, as noted in [7].

Despite the powerful empirical performance of our proposed framework, there are certain limitations. The model is based on a single historical dataset and does not incorporate temporal borrower behavior or macroeconomic indicators, which have been seen to drive credit risk dynamics [2], [10]. Additionally, the static formulation of the default prediction task may not fully capture shifting borrower risk profiles over time and is, therefore, susceptible to generalization in real-world lending environments.

Such limitations can be addressed by conducting further research through the incorporation of time-series financial information, borrower behavior patterns, and other external economic indicators to ensure robustness and adaptability in such a framework as suggested by reference [6]. In conclusion, therefore, the current study clearly depicts the potential use of interpretable deep-learning-based models in the conduct of reliable and transparent credit risk assessment activities within the modern financial world as suggested by references [7], [8].

REFERENCES

[1] Machine Learning and Deep Learning for Loan Prediction in Banking: Exploring Ensemble Methods and Data Balancing by Muhammad Asif Zahoor et al. (IEEE Access 2024)
[2] Design of Personal Credit Risk Prediction Model and Financial Risk Legal Prevention(IEEE Access 2024)
[3] Improved ADASYN Sampling and Optimized LightGBM for Credit Risk Prediction, Journal of Social Computing 2024
[4] Towards a Machine Learning-based Model for Corporate Loan Default Prediction IJACSA 2024
[5] Predicting Default Risk on Peer-to-Peer Lending Imbalanced Datasets (IEEE Access, 2021)
[6] A Novel Hybrid Model for Loan Default

Prediction in Maritime Finance Based on Topological Data Analysis and Machine Learning (IEEE Access 2025)

[7] A Deep Learning Approach for Credit Scoring of Peer-to-Peer Lending Using Attention Mechanism LSTM (IEEE Access, 2023)

[8] AI-Based Hybrid Models for Predicting Loan Risk in the Banking Sector (IEEE Access, 2023)

Indicator of collective Insights(IEEE Access, 2021)

[9] Formal Specification and Verification of Smart Contract-Based Loan Management System Using TLA+(IEEE ACCESS,2025)

[10] Crowd Dynamics in Online Lending:Unveiling