ISSN 2581 – 3781

AJCSE
ASIAN JOURNAL OF COMPUTER SCIENCE ENGINEERING

Available Online at www.ajcse.info
Asian Journal of Computer Science Engineering 2025;10(3):1-10

RESEARCH ARTICLE

Machine Learning-driven Risk Management Strategies for Enhancing Stability in the Financial Sector

Neha Upadhyay

Department of Computer Applications, IIS University, Bhopal, Madhya Pradesh, India

Received on: 01-03-2025; Revised on: 15-06-2025; Accepted on: 01-08-2025

ABSTRACT

The increasing sophistication and speed of financial dealings are challenging and require sophisticated risk management mechanisms that cater to large-scale, diversified, and dynamically changing data. This paper introduces a machine learning-powered framework for optimizing financial sector stability, specifically, toward credit risk prediction. Exploiting the German credit data, the strategy uses strict reprocessing, characteristic selection, and data overlapping with the synthetic minority over-sampling technique that operates on data imbalance and enhances the robustness of the models. Random forest (RF) classifier is applied because it handles non-linear patterns, and it is able to avoid overfitting and also provides a readable feature importance. The model performs ultra-high in terms of the predictive performance, as accuracy, precision, recall, F1-score, and receiver operating characteristic curve-area under the curve results show an accuracy of 97.61%. Comparison of accuracy to gradient boosting, support vector machine, and gated recurrent unit shows the outstanding performance of the RF model with respect to categorizing credit risks. The results suggest that the framework that is proposed is scalable and interpretable to provide a solution to proactive risk management. The study advances the research literature in that it offers an effective model that facilitates enlightened choices, fewer theoretical losses, and establishes systemic resiliency within fast-changing financial marketplaces. The work envisaged in the future the incorporation of real-time financial and sentiment data to build an even greater level of predictive ability.

Keywords: Credit risk prediction, financial stability, machine learning, random forest, financial risk management, financial sector

INTRODUCTION

The financial sector is the heart of contemporary economies, as it supports capital allocation, investing, payment systems, and wealth management in the international markets. It includes banks,[1,2] insurance companies, investment firms, and regulatory agencies, and all these are collaborating to enable economic growth and commercial and consumer-based activities.[3] Being a highly integrated system, its efficient functioning is a key to not only the working of each particular institution but also economic well-being at the national and global levels.

***Corresponding Author:**
Neha Upadhyay
Email: neha.upadhyay887@gmail.com

Financial stability is needed in the financial industry since any turmoil may spread rapidly across markets[4] to businesses, governments, and households.[5] Even modest shocks, when unmanaged, can bring about greater systemic crises which destroy confidence and slow down economic growth. It is on this basis that risk management becomes a priority in its effective management, as a guarantee of strength within uncertainty.[6] The risk environment of the financial institutions has been evolving in recent years to be even more complex. There are persistent risks of market volatility, credit defaults, liquidity shortages, cyberattacks, and changing regulatory requirements.[7] The old techniques of risk management such as value at risk, stress testing, and credit scoring have also been useful in the determination and curb of the risks.[8] However, these practices are increasingly not coping well with the fast pace, connected nature, and magnitudinous nature of the contemporary financial frameworks. This has increased the need for solutions that have the capacity to process high quantities of heterogeneous data in real time and provide more in-depth predictive insights.

Financial services digitalization has resulted in a proliferation of data through transaction,[9] market feeds, macro trends, and other novel sources such as social media sentiment and news analytics.[10] Utilizing this data in the most efficient manner possible needs smart technologies that are able to extract patterns and report actionable intelligence with very little latency. It is here that machine learning (ML) has come and changed the picture. A subfield of artificial intelligence[11,12] called ML[13] allows systems to learn based on historical and real-time information and recognize complicated patterns and adapt to changes in conditions without being programmed to do so. Among the benefits it poses in comparison to traditional statistical methods, these are its capacity to process data of high dimensions, non-linear expression, and formulation by integrating structured and unstructured data. Credit default prediction, fraud detection, portfolio optimization, and market risk prediction are all financial applications.[14]

Using ML on various kinds of financial risk, including credit, market,[15] operational, and liquidity, institutions can enhance early warning systems, optimize decision-making processes, and better adhere to regulatory frameworks.[16] Such capabilities go directly toward stability in a sector by decreasing the propensity of systems-wide breakdowns and allowing proactive responses to risk.[17] This paper examines machine-learning-based risk management plans that seek to increase the stability of the financial field. It has offered a conceptual and analytical framework, a comparison of ML techniques to the traditional ones, and practical implications on financial institutions as well as policymakers.

Motivation and Contribution

Credit risk management is essential in the maintenance of financial institutions, minimizing the possibilities of default and maintaining confidence in the market in the long run. Conventional statistical models are sometimes unable to describe non-linear connections, effectively work with unbalanced data, and give explanations to be used when making decisions. Such constraints may cause the classification of borrowers with high risks to be incorrectly, causing a lender financial losses and compliance problems. As the scale of computer-based financial transactions continues to rise and different and high-dimensional datasets become available, the drive to investigate more advanced ML techniques that achieve a combination of these characteristics has become acute. This project aims to use these capabilities to facilitate proactive significant risk identification, enhance regulatory customer compliance, and facilitate data-driven lending initiatives. The primary results of this study are as follows:

- The analysis utilizes the German credit dataset, a publicly available benchmark dataset for credit risk prediction, with extensive pre-processing steps including handling of missing values, removal of outliers, and application of min-max normalization to ensure consistency and uniform scaling across all features.
- A critical feature selection strategy is used to maximize the model's efficiency. The synthetic minority over-sampling technique (SMOTE) is a way to handle class imbalances. It treats high-risk and low-risk individual credit situations equally.
- The application of a random forest (RF) classifier during the model training is because the RFs are robust, interpretable, and able to capture nonlinear relationships in complex financial data.
- Structured evaluation models include numerous important performance indicators to ensure a comprehensive comparison of the model's predictive power. These metrics include accuracy, precision, recall, and F1-score.
- The methodology shows that ML-powered credit risk models would key in better predictability to deliver actionable information that would guide financial institutions in taking better lending decisions, cutting high default rates, and stabilizing the market to good behavior.

Justification and Novelty

The rationale of such a research project is based on the idea that efficient, scalable, and explainable models of predicting credit risk are in high demand in the environment of data complexity, regulatory pressure, and financial volatility. Traditional statistical methods often fail to capture non-linear relationships and handle data imbalance effectively. This study introduces an RF-driven framework enhanced by rigorous preprocessing, feature selection, and SMOTE-based balancing to overcome these limitations. The novelty stems from integrating interpretable feature importance with high predictive accuracy, enabling financial institutions to proactively manage risks, reduce potential losses, and maintain systemic stability through data-driven, real-time decision-making.

Organization of the *Paper*

The structure of the paper is as follows: Section II talks about related work on using ML to predict credit risk and control financial risk. In Section III, they talk about the suggested method, how the model will be made, and the factors for evaluation. Section IV shows the outcomes of the experiment, including a comparison of the results. Section V ends with some important conclusions and ideas for future study.

LITERATURE REVIEW

A recent comprehensive review and analysis of numerous significant studies on risk management for enhancing stability in the financial sector has been conducted. Notable recent works include:

Shi *et al*. (2025) – The hybrid financial risk predictor (HFRP) is a model that was created through an empirical study. Utilizing a hybrid approach that incorporates CNN and LSTM networks enhances financial risk prediction. Results are highly steady and accurate when quantitative and qualitative ratings from financial text analysis are combined, particularly when compared to the HFRP model. Remarkable results include a clear reduction in both the training loss quantity (0.0013) and the testing loss quantity (0.003). They assume in this hypothesis that the chosen HFRP model yields very accurate estimates of sales, net income, and EPS. The model successfully lowers the risk significantly. Compared to the previous levels, operational risk is now at 0.35, credit risk is at 0.25, liquidity risk is at 0.25, and market risk is at 0.30. Findings from the HFRP model indicate that the plan would lead to more secure financial markets.[18]

Raliphada *et al*. (2025) paper estimates borrower default probabilities and effectively manages financial risk. While traditional models, such as LR (92% accuracy) and XGBoost (94% accuracy), primarily rely on structured numerical data, recent advancements in NLP have facilitated the incorporation of sentiment analysis derived from financial news. This study investigates the optimization of credit risk classification using BERT (91% accuracy), RoBERTa (90% accuracy), and LLaMA 3.2 (78% accuracy), in conjunction with VADER sentiment analysis using a dataset of financial news.[19]

Barongo and Mbelwa (2024) – The researchers overcame two obstacles: The limited scope of LR factors and risks in data integrity. The quality of the loan portfolio, market circumstances, asset strategies, and financing arrangements were not included as realistic LR variables in these calculations. This analysis utilized results from 38 different Tanzanian banks that were gathered by the Bank of Tanzania between 2010 and 2021. Eleven features were identified for usage in ML analysis and LR rating after substantial factor experimentation with RF and MLP models. In addition to increasing LR sensitivity and reducing the limitations of RF and MLP models through generalization, it was statistically and practically effective. Reducing Type I and Type II errors in classification was achieved with a discriminant power of 2.61, accuracy ranging from 90% to 96%, precision, accuracy, F1 score, G-mean, Cohen's kappa, Youden's index, and area under the curve (AUC). There was a 0.8%, 9.1%, and 1% chance of the negative happening. Preceding LR examples demonstrated the superior performance of RF-MLP over MLA.[20]

Pritam *et al*. (2024) examine the effectiveness of the RF model in detecting bankruptcy and evaluate the impact of SMOTE in enhancing the accuracy of classification. The RF has significant promise in forecasting bankruptcy, while the use of SMOTE to address imbalanced data yields a favorable effect, bolstering the dependability of financial risk assessment models. An accuracy rate of 97% is reached by taking into account a diverse variety of optimization parameters. Subsequent investigations may prioritize enhancing the precision and promptness of bankruptcy prediction models.[21]

Zhou (2023) – integrating the BP neural network with the partial least squares (PLS) method, this research constructs the BP-PLS model, which serves as a financial crisis warning system for online retailers. Financial crisis warning signs for e-commerce firms are examined and selected in the experiment before components are extracted using the PLS approach. The BP neural network is then fed vectors containing the data from the component extraction process. Finally, the experiment produced financial crisis warning models for e-commerce firms in T-1, T-2, and T-3 years, respectively, using the BP-PLS model. The experimental data show that the T-1 and T-2 models are more than 90% accurate. An 85% improvement above the mean is achieved by the T-3 model.[22]

Bi *et al*. (2022) – predicting the default behavior of financial markets, reducing the bad debt rate in bank loans and securities investment, and immediately identifying potential dangers are the objectives of this research. The current methods rely heavily on simple linear models and other weighted models. The speed of these algorithms is an advantage. However, these methods also have accuracy issues when working with large samples; thus, ML modelling approaches are needed for training the models. Financial data mining algorithms based on RFs can accurately predict micro behaviors and mitigate financial risks; this study provides a modeling foundation for such algorithms. Recall is 90% and forecast accuracy is 85% for about 90% of the time.[23]

Table 1 provides an overview of current research on financial risk management, showcasing new models, datasets, important results, challenges, and potential future routes.

RESEARCH METHODOLOGY

The study introduces a method that uses ML to improve financial sector risk management tactics. It primarily aims at predicting credit risk using the German credit dataset. Methodology starts with thorough data pretreatment, which includes fixing missing numbers and removing outliers to keep data consistent. Min-Max normalization is subsequently used to have equal feature scaling. After the model's important elements are refined to make it functional, SMOTE is employed to address issues of class imbalance. Partitioning the cleaned data into a training set and a testing set

Table 1: Literature summary of machine learning approaches for financial risk management and stability

Author (s)	Methodology	Dataset	Key findings	Limitations	Future direction
Shi *et al*. (2025)	Hybrid CNN + LSTM (HFRP model) with quantitative and qualitative text data analysis	Financial statements + financial text data	Loss = 0.0013/0.003; reduced credit, liquidity, market, and operational risks by ~ 50–75%; close match to actual revenue, net income, EPS	No details on dataset size/diversity; unclear generalizability to other markets	Expand to cross-country datasets; test against extreme market volatility
Raliphada, *et al*. (2025)	NLP sentiment analysis (VADER) with ML models (XGBoost, Logistic Regression, BERT, RoBERTa, LLaMA 3.2)	Financial news dataset	XGBoost achieved highest accuracy (94%); sentiment data improved credit risk prediction	Limited to text data; ignores structured financial data; some NLP models underperformed	Combine structured and unstructured data; test multilingual news datasets
Barongo and Mbelwa, (2024)	Hybrid RF-MLP for liquidity risk classification	38 Tanzanian banks' data (2010–2021)	Achieved 90–96% across accuracy, precision, recall; improved sensitivity over traditional MLA metrics	LCR and NSFR data missing; limited to Tanzanian banking sector	Include global bank datasets; integrate real-time data feeds for continuous monitoring
Pritam *et al*. (2024)	Random Forest + SMOTE for imbalanced bankruptcy prediction	Bankruptcy datasets (unspecified)	Accuracy 97%; SMOTE improved performance on imbalanced data	Dataset source not disclosed; potential overfitting risk	Test on varied industry sectors; explore deep learning approaches
Zhou (2023)	BP neural network + PLS for feature extraction (BP-PLS)	E-commerce enterprise financial indicators	Accuracy > 90% (T-1, T-2), >85% (T-3); strong training convergence	Specific to e-commerce, it may not generalize to other industries	Apply to multi-sector datasets; incorporate macroeconomic indicators
Bi *et al*. (2022)	Random Forest for financial data mining	Large financial market dataset (unspecified)	Precision 85%, Recall 90%; identified potential risks and reduced bad debt rate	Dataset details missing; only one ML model tested	Compare with other ensemble & deep learning models; use multi-source financial data

allows for fair model testing. An RF classifier's noise management and resilience make it an excellent choice for training with non-linear financial data patterns. Using recall, F1-score, and accuracy and precision as shown in Figure 1, they assess the model's performance. Improved forecast accuracy and more stable, well-informed financial decisions are two significant benefits of this procedure.

Data Collection

The raw data used is a German credit dataset that has been collected through Kaggle and contains 32,581 observations in 12 variables. Information such as person_age, person_income, person_home_ownership, person_emp_length, loan_intent, loan_grade, loan_amnt, loan_int_rate, loan_percent_income, cb_person_default_on_file, and c b person cred hist length are included in its demographic and financial features. The target variable, loan status, is a binary one; it takes on the value 0 when the loan is paid in full and the value 1 when the debt is in default.

Figure 2 displays the class distribution of the dataset. The two groups are shown on the x-axis, while the overall number of observations is shown on the y-axis paid(0) and default(1). The chart clearly shows a significant imbalance in the data. The "Paid" class has a much higher number of observations, nearing 8,000, compared to the "Default" class, which has approximately 2,000 observations. This imbalance indicates that most credit accounts in the dataset were paid on time, while a smaller portion defaulted. Credit risk modelling often faces this problem, which might cause models to favor the majority group.

The heatmap displays the key variables in the credit dataset in Figure 3, highlighting relationships between demographic, financial, and loan-related features. Most correlations are relatively low, indicating limited linear dependency among variables. This is because cb_person_cred_hist_length and person_age have a high positive correlation of 0.86, indicating that applicants with longer credit histories are more likely to be older. Default risk can be predicted by loan status, loan interest rate, and loan percentage of income, all of which have moderate positive relationships. The matrix indicates that while some features are moderately related, there is no

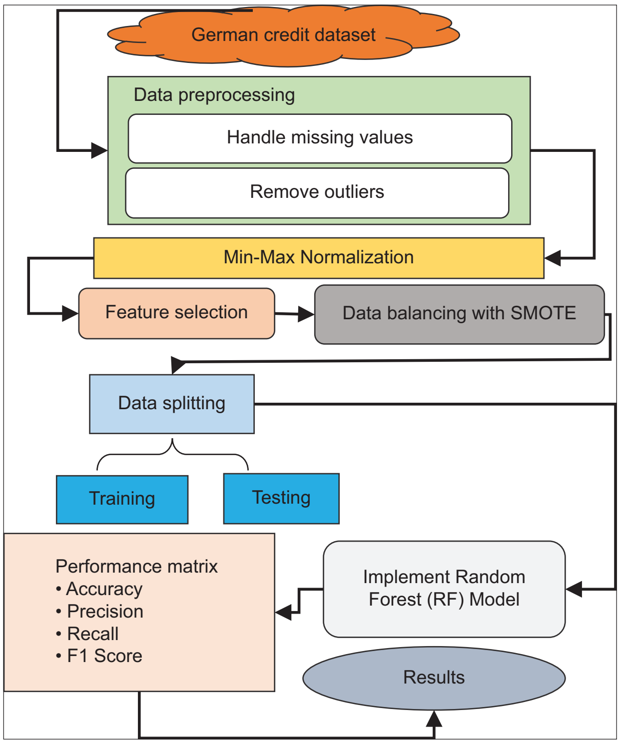

Figure 1: Proposed flowchart for risk management in the financial sector

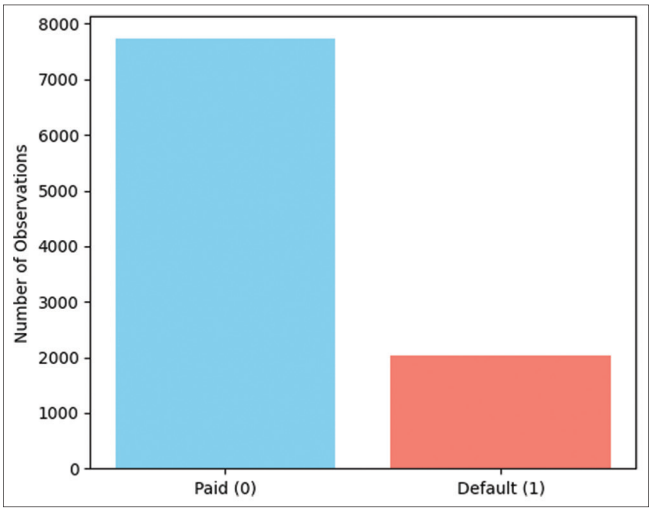

Figure 2: Data distribution of the dataset

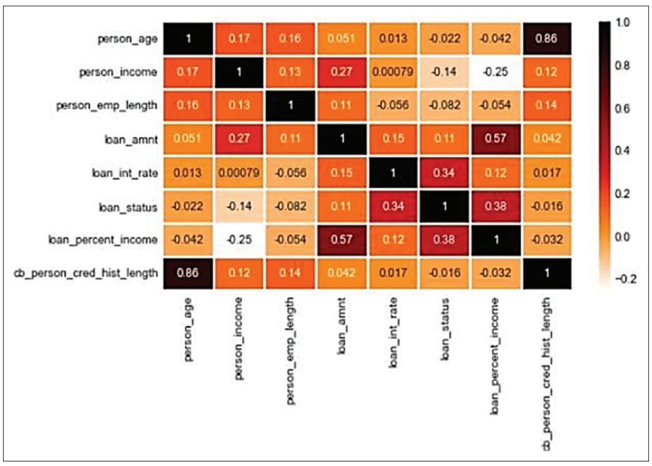

Figure 3: Correlation heatmap of the dataset

excessive multicollinearity, making them suitable for predictive modeling.

The variables that have the greatest impact on predicting loan status are highlighted in the bar chart in Figure 4, which shows the feature importance scores from the RF model. The fact that Loan_percent_income is the most important predictor suggests that the amount of money going toward paying off loans is the most important factor in categorization. person_income, loan_int_rate, and loan_amnt follow, demonstrating the vital importance of both income level and loan terms. Even if person_emp_length and person_age are not very important, they do help the model anticipate outcomes. The two most important factors in determining a borrower's creditworthiness, according to the bar chart, are their income level and the conditions of their loan.

Data *Preparation*

Data preparation involved the process of collecting data, which was financial data, to concatenating the dataset and cleaning the data to maintain its consistency. Applicable features were extracted, and then, the dataset was pre-processed by deleting missing values and outliers. Thereafter, data transformation and normalization processes were carried out. The major steps taken to pre-process are as follows:

- Remove missing value: One of the methods of data pre-processing, which has been applied in preparation for data analysis or ML models, is the deletion of missing values. This is carried out by identifying and working on situations of data absenteeism and incompleteness.
- Remove outliers: Data pre-processing Remove outliers Removal of outliers involves detecting and removing or at least moderating the effects of data points that are far out of line with most of the data. The purpose of this process is the enhancement of accuracy and the reliability of the following analyses or ML models.

Data Normalization

Normalization of records was performed under the method of the min-max technique to bound the values between 0 and 1. This was done with a view to streamlining the performance of the classifiers under consideration and to reducing the influence of the outliers.[24] Normalization was carried out by use of the following mathematical formula (1):

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

where X indicates that the feature was originally, X' is the normalized value, X_{min} is the minimum value and the same feature, and X_{max} is the maximum value of the same.

Feature Selection

Feature selection refers to the selection of a subset of features of interest in the initial set of features of a data set. It is an important feature of the ML process since it seeks to enhance the performance of the models and make them simpler and easier to understand by eliminating irrelevant or redundant features. ML the selection of a subset of the relevant features (input variables) within a dataset to be employed in the construction and training of an ML model is called feature selection. This is to eliminate the uninformative features and keep the most informative features relevant, thus eliminating the irrelevant or redundancy of the feature.

Data Balancing with SMOTE

Balancing data methods for calculating the data set's class distribution are utilized primarily in ML applications dealing with unbalanced datasets. The data balancing by SMOTE is used to solve the problem of an unbalanced dataset in the context of ML, in which only samples of a specific class (say, one of them is the minority class) are not adequate to make the analysis of ML.

Data Splitting

The training and testing sets were taken. The general split was the training of the model and estimation of its parameters using the first 80% of the data and 20% to test and evaluate the efficiency of the model.

Proposed RF Model

RF is a type of ensemble ML method that builds and connects different DTs to make classification or regression work better. To generate variety, the training data for both trees are randomly selected (bagging), and at each split, the trees switch up which features to use, further increasing the randomization.[25] For classification, the final prediction is made by casting a majority vote; for regression, it is an average over all trees. Data are partitioned in a recursive fashion. A query on an attribute is used to do the split at a specific node.[26] The splitting criterion is chosen according to certain impurity metrics, like the Gini impurity or Shannon entropy. The Gini impurity is used as the function to find the quality of the split in each node. Equation (2) gives the Gini impurity at node N:

$g(N) = \sum_{i \neq j} P(w_i) P(w_j)$ (2)

Where the percentage of the population that falls into class i is denoted as $P(_{wi})$. As an additional metric, Shannon entropy can be utilized to evaluate the split's quality. It quantifies the disarray in the data. Quantify the unpredictability of the information contained in a given node of a decision tree using its Shannon entropy. With the help of Equation (3), determine the entropy of a node N:

$$H(N) = \sum_{i=1}^{i=d} P(w_i) log_2(Pw_i)) \quad (3)$$

Where P(*wi*) is the percentage of the population that is labeled as i and d are the number of classes that are being looked at. The maximum entropy is observed when all classes are present in an equal proportion within the node. When a node is pure, meaning it has exactly one class, it has the lowest connectivity. At each node, the most apparent heuristic for selecting the optimal splitting option is the one that minimizes impurity. To rephrase, the optimal split is defined by the biggest improvement in purity or the greatest improvement in information gain. An equation for the information gain from a split is Equation (4):

$\Delta I(N) = I(N) - P_L * I(N_L) - P_R * I(N_L)$ (4)

With I(N) representing node N's Gini or Shannon entropy, P_L and P_R denoting the proportions of node N's population that go to N's left and right children, respectively, following the split. When they examine N, they find N_L on the left and N_R on the right.

Evaluation Metrics

Several performance criteria were used to assess the efficacy of the suggested design. The trained models predicted outcomes and values were contrasted with the actual ones. The outcomes were classified into four groups: False positives (FP), true negatives, true positives (TP), and FN. Below are defined and determined the performance metrics that follow:

Accuracy

The ratio of instances in the dataset (input samples) that the trained model accurately predicted relative to the total assortment of examples in the dataset.[27] The expression in Equation (5) shows it to be:

$$Accuracy = \frac{\mathrm{TP} + TN}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \quad (5)$$

Precision

Accuracy is how many out of every 100 positive occurrences that the model was able to correctly predict. Accuracy shows. How good the classifier is in predicting the positive classes is expressed as Equation (6):

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \quad (6)$$

Recall

The ratio of events that were accurately predicted as positive to all instances that should have proved positive. In mathematical form, it is shown in Equation (7):

$$Recall = \frac{\mathrm{TP}}{TP + FN} \quad (7)$$

F1 score

The F1-score aids in maintaining a healthy equilibrium between recall and precision since it is a composite of the harmonic mean of the two. It can take values between zero and one. The equation for it is provided by mathematics (8):

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

Receiver operating characteristic curve (ROC)

The ROC plots, for a set of decision cut-off points, show the ratio of successfully categorized cases to those that were wrongly classified. FPR is equal to 1-specificity, but TPR is often called sensitivity or recall.

RESULTS AND DISCUSSION

This study delves into the use of ML models in the financial sector, specifically looking at how

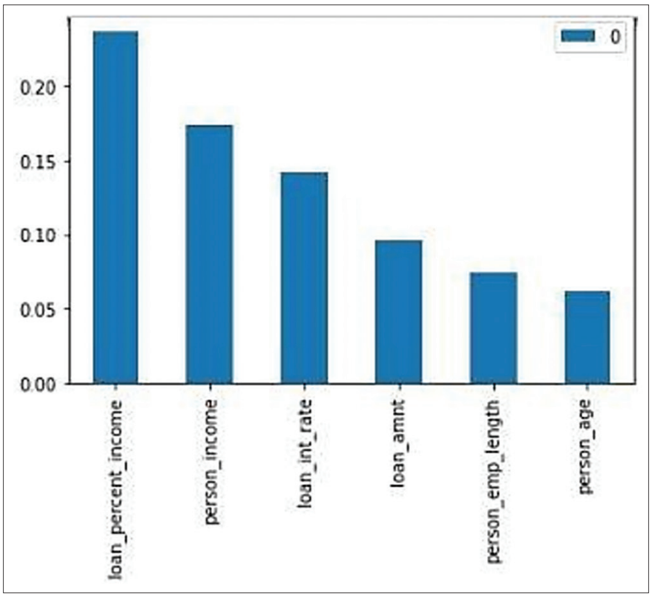

Figure 4: Feature importance using a random forest model

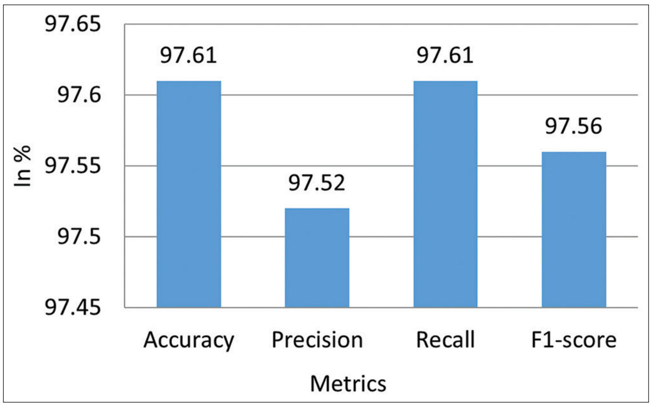

Figure 5: Bar chart of random forest model performance for risk management

Table 2: Experiment result of the proposed model for risk management for enhancing stability in the financial sector

Performance metrics	Random forest model
Accuracy	97.61
Precision	97.52
Recall	97.61
F1-score	97.56

they might improve risk assessment and decision-making. Experiments were conducted using Python with the Scikit-learn framework on an enterprise-grade computing environment running a 64-bit Windows Server OS with 64 gradient boosting (GB) of RAM, optimized for large-scale financial data analytics. Table 2 summarizes the results, which show that the RF model performed exceptionally well, with 97.61% accuracy, 97.52% precision, 97.61% recall, and 97.56% F1-score. These results demonstrate that the model can reliably detect dangers and handle complicated patterns. The results underline the possibilities of the strategies using RF which may be used to reinforce financial risk management systems, helping organizations proactively counteract threats and stabilize, making informed decisions in a rapidly changing market environment.

Figure 5 is a bar chart displaying the output of an ML model, most likely used for financial risk management purposes. With values ranging from 97.46 to 97.62 on the y-axis, the provided chart lists four key metrics: Accuracy, Precision, Recall, and F1-score on the X-axis. With a precision of 97.52%, the accuracy and recall are both 97.61%. The F1-score amounts to 97.56%. Such percentages at the high-level mean that the RF model performs well in the identification and management of risks, which means it has the potential to contribute to stability in the financial sector.

A comparison of the model's forecasts with the observed results is shown in the matrix. From what I see, the model properly identified 1801 genuine-positive cases and 1783 true-negative cases (i.e., transactions that were not considered dangerous). There were a total of 45 FN and 44 FP, indicating that the model made very few mistakes. The RF model's dependability and accuracy in recognizing and controlling risks are supported by its low false-negative rate and low number of misclassifications. As a result, the financial sector is more stable.

Figure 7 shows the ROC curve for an RF model, an important criterion for evaluating binary classification models used in financial risk management. A graph showing the ratio of TP to FP is displayed by the curve. You can see the RF model's performance by looking at the blue line in the top left corner of the graph; this line shows high prediction ability. The dotted diagonal line represents a random classifier, and as the ROC curve is shifted away from the model, its performance improves. With an estimated value of 0.977, the AUC is nearly identical to 1.0. This model's capacity to effectively identify and address risks while simultaneously mitigating stability is a result of its high AUC, which shows how well it distinguishes between positive and negative classifications.

Comparative Analysis

This section analyses ML models on risk management strategies to be used in ensuring

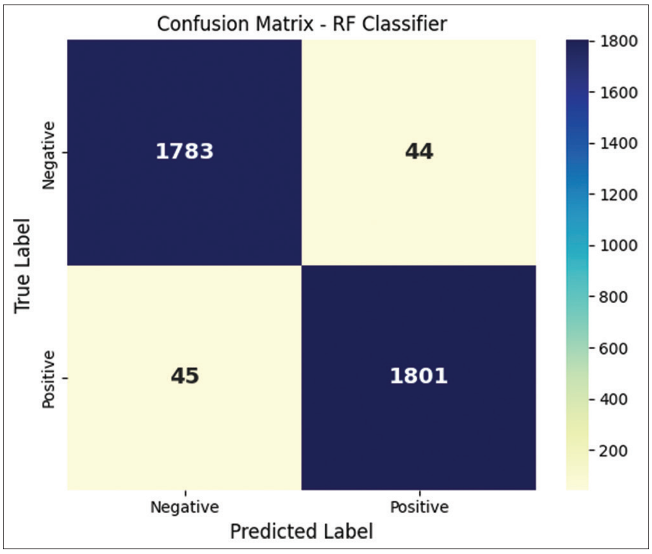

Figure 6: Confusion matrix for random forest model for risk management

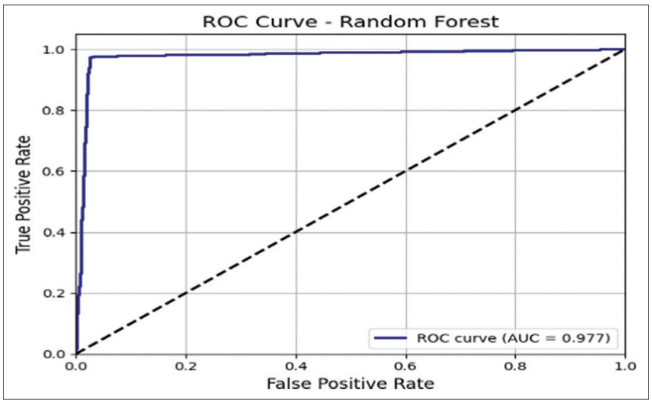

Figure 7: Receiver operating characteristic curves of the random forest model for risk management

that financial stability is achieved, with the case being the measure of accuracy. According to the results summarized in Table 3, RF model with an accuracy of 97.61% proved to be the most accurate in detecting instances of existing financial risks and assisting in providing effective risk mitigation efforts. This was later replicated with the GB model with an accuracy of 94% indicating a high degree of predictive but not as high as RF in terms of reliability. Support vector machine (SVM) model demonstrated relative potential in modeling risk patterns, achieving 88.53% accuracy, and gated recurrent unit (GRU) model had the lowest failure of 76.54% which implies inefficiency in modeling financial associations, hence, the complexity of the practice. These findings emphasize the promise of RF as one of the prominent ML strategies toward risk management in the financial industry to provide tactical information to strengthen financial stability and strategic decision-making.

Table 3: Accuracy comparison of different predictive models for risk management in the financial sector

Models	Accuracy
GRU[28]	76.54
SVM[29]	88.53
GB[30]	94
RF	97.61

GRU: Gated recurrent unit, SVM: Support vector machine, GB: Gradient boosting, RF: Random forest

ML models used to manage financial risks have created new forms of facilitating gains in predictive precision and sector stability. The proposed RF model was especially successful, and it was much more effective than other methods in spotting pattern formation risks in the credit behavior. Its pattern of ensemble learning can process complicated financial data and mitigate the overfitting as well as discover the importance of its features, vital risk factors. Such a union of strength, accuracy, and explainability allows RF to produce significant predictions, which is valuable in the process of decision-making and rational risk management plans within the financial sector.

CONCLUSION AND FUTURE STUDY

RF algorithm through ML-based methods offers a scalable, interpretable, and effective solution toward credit risk prediction in the financial market. The proposed framework was able to solve the problems of class imbalance and overfitting and capture the complicated non-linear relationships present in financial data by the use of rigorous pre-processing, feature selection, and SMOTE-based balanced. The RF model attained a significant accuracy of 97.61% compared to GB, SVM, and GRU and produced great precision, recall, and F1-scores. An importance analysis of the features displayed some factors related to the income and loan terms as feature weights, which could prove valuable in providing insight when arriving at the determination of risk. This pairing of predictive performance and interpretability can help financial institutions to enhance systemic stability, compliance, and minimize possible losses, to the extent of using data-driven decisions.

Future research will be based on adding real-time transactional, macroeconomic, and sentiment data to the framework to enhance responsiveness and flexibility. The application of hybrid models comprising both the ensemble methods and the deep learning architectures might further improve the processing of temporal and unstructured data. Furthermore, it will be important to test a framework across various territories and industries. Different financial products will make it generally applicable, resilient to changes in the market risks, and it is ready to be implemented in various operational contexts.

REFERENCES

1. Mishra J, Biswal BB, Padhy N. Machine Learning for Fraud Detection in Banking Cyber security Performance Evaluation of Classifiers and Their Real-Time Scalability. In: 2025 International Conference on Emerging Systems and Intelligent Computing (ESIC). IEEE; 2025. p. 431-6.
2. Patel D. Enhancing banking security: A blockchain and machine learning-based fraud prevention model. Int J Curr Eng Technol 2023;13:576-83.
3. Garg D, Shelke NA, Kitukale G, Mehlawat N. Leveraging Financial Data and Risk Management in Banking sector using Machine Learning. In: 2024 IEEE 9th International Conference for Convergence in Technology (I2CT); 2024. p. 1-6.
4. Balasubramanian A. Proactive machine learning approach to combat money laundering in financial sectors. Int J Innov Res Eng Multidiscip Phys Sci 2019;7:1-15.
5. Majumder RQ. A review of anomaly identification in finance frauds using machine learning systems. Int J Adv Res Sci Commun Technol 2025;5:101-10.
6. Malali N. Exploring Artificial Intelligence Models for Early Warning Systems with Systemic Risk Analysis in Finance. In: 2025 International Conference on Advanced Computing Technologies (ICoACT). IEEE; 2025. p. 1-6.
7. Liu Z. Application of machine learning in financial risk classification and account verification optimization strategy. Econ Manag Innov 2025;2:64-70.
8. Chaudhari B, Verma SC. Synergizing generative AI and machine learning for financial credit risk forecasting and code auditing. Int J Sci Res Comput Sci Eng Inf Technol 2025;11:2882-93.
9. Verma V. Deep learning-based fraud detection in financial transactions: A case study using real-time data streams. ESP J Eng Technol Adv 2023;3:149-57.
10. Alarfaj FK, Shahzadi S. Enhancing fraud detection in banking with deep learning: Graph neural networks and autoencoders for real-time credit card fraud prevention. IEEE Access 2024;13:20633-46.
11. Malali N. AI ethics in financial services: A global perspective. Int J Innov Sci Res Technol 2025;10:71-8.
12. Dattangire R, Vaidya R, Biradar D, Joon A. Exploring the Tangible Impact of Artificial Intelligence and Machine Learning: Bridging the Gap between Hype and Reality. In: 2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET). IEEE; 2024. p. 1-6.
13. Majumder RQ. Machine learning for predictive analytics: Trends and future directions. Int J Innov Sci Res Technol 2025;10:3557-64.
14. Kali H. Optimizing credit card fraud transactions identification and classification in banking industry using machine learning algorithms. Int J Recent Technol Sci Manag 2024;9:85-96.
15. Prajapati N. The role of machine learning in big data analytics: Tools, techniques, and applications. ESP J Eng Technol Adv 2025;5:???. AQ6
16. Shi S, Tse R, Luo W, D'Addona S, Pau G. Machine learning-driven credit risk: A systemic review. Neural Comput Appl 2022;34:14327-39.
17. Verma V. Security compliance and risk management in AI-driven financial transactions. Int J Eng Sci Math 2023;12:107-21.
18. Shi X, Zhang Y, Yu M, Zhang L. Deep learning for enhanced risk management: A novel approach to analyzing financial reports. PeerJ Comput Sci 2025;11:e2661.
19. Raliphada P, Olusanya MO, Olukanmi S. Optimizing Credit Risk Model Classification Using BERT, RoBERTa, and LLaMA 3.2. In: IEEE EUROCON 2025 - 21st International Conference on Smart Technologies. IEEE; 2025. p. 1-6.
20. Barongo RI, Mbelwa JT. Using machine learning for detecting liquidity risk in banks. Mach Learn Appl 2024;15:100511.
21. Pritam N, Malhotra S, Aluvala S, Gill KS, Devliyal S. Bankruptcy Detection Using Random Forest Classifier and its Data Balancing Using Smote Analysis. In: 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0; 2024. p. 1-4.
22. Zhou Z. Machine learning in financial risk forecasting and management for trading firms. Int J Comput Syst Eng 2023;7:159-67.
23. Bi H, Fang S, Bi X. Prediction method of financial market risk behavior based on big data mining algorithm. Acad J Comput Inf Sci 2022;5:78-84.
24. Malempati M. Machine learning and generative neural networks in adaptive risk management: Pioneering secure financial frameworks. Kurd Stud 2022;10:691-701.
25. Aysan AF, Ciftler BS, Unal IM. Predictive power of random forests in analyzing risk management in Islamic banking. J Risk Financ Manag 2024;17:104.
26. Nimmala R. Enhancing financial liquidity risk management through machine learning-navigating climatechange impacts. J Technol Innov 2022;3:1-5.
27. Khunger A, Anand K, Shukla C, Dnyandev Jagdale A, Chinnakannan A, Dbritto C. Deep learning for financial stress testing: A data-driven approach to risk management. Int J Innov Stud 2022;6:80-92.
28. Li T, Dai X. Financial risk prediction and management using machine learning and natural language processing. Int J Adv Comput Sci Appl 2024;15:211-9.
29. Yang D, Ma H, Chen X, Liu L, Lang Y. Design of financial risk control model based on deep learning neural network. Comput Intell Neurosci 2022;2022:5842039.
30. Daiya H. Financial risk assessment in fintech using ensemble learning methods. J Emerg Technol Innov Res 2023;10:1-13.

Author Queries???

AQ6: Kindly provide page number