EAI Endorsed Transactions
on AI and Robotics

Research Article EAI.EU

AI-Powered Predictive Analytics for Financial Risk Management in U.S. Markets

Md Zikar Hossan[1], Muslima Begom Riipa[1], Md Azhad Hossain[1], Sweety Rani Dhar[1], Al Modabbir Zaman[1], Mohammad Hossain[1], Arif Hossen[2], and Hasan Mahmud Sozib[3*]

[1]Department of Business Administration, International American University, Los Angeles, CA 90010, USA
[2]Department of Business Administration, Trine University, Angola, IN 46703, USA
[3]Department of Electrical and Electronics Engineering, Ahsanullah University of Science and Technology, Tejgaon, Dhaka, 1208, Bangladesh

[*]Corresponding Author: sozib2019@gmail.com

Abstract:

In the fast-changing environment of financial complexity, efficient risk management is vital for economic stability as well as for growth. In this study, we present a robust AI-powered predictive analytics framework to improve financial risk classification in U.S. markets. The framework utilizes advanced machine learning techniques, a hybrid CatBoost and SVM model that allows it to solve challenges like class imbalance in a high-dimensional dataset while maintaining interpretable models. To probe errors, we use techniques such as Principal Component Analysis (PCA) and Synthetic Minority Oversampling Technique (SMOTE) for data quality and fairness in classification. Comprehensive experiments on a financial risk dataset are conducted to evaluate the framework at which it achieves high accuracy (95.93%) and F1-score (0.95) when compared to traditional machine learning models such as Logistic Regression and Random Forest. Furthermore, a feature importance analysis identifies important predictors of financial risk such as Total Debt-to-Income Ratio, Loan Duration, and Interest Rate, providing actionable on decision-making. Additionally, the proposed approach is not only highly scalable but it is also interpretable and adaptable to the dynamic demands of financial institutions. This study serves as a benchmark for predicting analytics for dealing with risk-associated challenges, leading to informed decision-making to ensure economic stability by integrating AI and machine learning in financial systems.

Keywords: Financial risk management, AI-powered, predictive analytics, CatBoost, SVM, decision-making, U.S. markets.

Received on 12 June 2025, accepted on 30 July 2025, published on 22 August 2025

doi: 10.4108/airo.9532

1. Introduction

In today's dynamic and complex economic landscape, financial risk has become a significant challenge for businesses and financial institutions. Effectively predicting financial risks is crucial for maintaining economic stability and fostering sustainable development [1]. Over the past few years, the issue of financial risk management has become increasingly important due to several key factors: an escalating macroeconomic environment, heightened compliance demands, intensified market competition, and rising criminal activities [2]. Financial risk, for the most part, is regarded as the likelihood of a company failing to meet its obligations to creditors. To minimize various financial risks and assess the financial condition of companies, predictive models are employed, serving as warning systems to identify specific issues within the analyzed companies [3]. Over the past 75 years in the United States, financial risk tolerance assessment has predominantly focused on five key methodologies: preference between choices, rationalistic value functions, measurable values, prototypical estimates, and perceptual judgment [4].

The effective rating and valuation of financial risk has significant consequences on economic stability and the quality of decision-making processes in the U.S financial market. The task of calculating imbalanced profiles poses challenges to financial institutions because the latter is a multifaceted attribute of financial information containing elements such as the payment history's attributes, credit scores, DTI, and employment status, among others. The rapid advancements in information technology have made data analysis a valuable tool for addressing this issue. As one of the world's largest financial markets, the United States' financial institutions and banks have gained extensive experience and insights in leveraging data analysis for risk management [1]. Among the key advances in the stream of financial auditing and insurance, it is possible to note the use of risk [6]management models: variance, covariance, standard deviation, and value at risk [7]. Adapting these complex, high-dimensional data, handling minor and major classes across risk levels, and providing stable, accurate, and easily explained models for the multiclass problem. Based on the evolving structure of commercial banks, their competitive advantage depends on their ability to identify and mitigate risks promptly and efficiently. Contemporary approaches with the help of risk control based on artificial intelligence (AI), biometrics, and big data are identified as key concerns for financial specialists and scholars [2]. Additionally, the increasing complexity and volume of financial data necessitate more effective solutions. Financial analysis has greatly benefited from AI, and it is now used to find insights, trends, and execute trades [8].

Artificial intelligence and business analytics are being utilized in finance to predict patterns, identify risks, and make informed decisions promptly [9]. Using features such as machine learning and deep learning, these algorithms employ comprehensive analysis of past data and real-time calculations of significant data volume, which previously could not be revealed at such speeds [10].

To overcome the current challenges, this study develops an AI-powered predictive analytics framework for effective financial risk classification. By utilizing advanced machine learning algorithms, including hybrid models, the goal is to enhance predictive accuracy and deliver actionable insights for managing financial risk in U.S. markets [11]. Especially for addressing the financial risk categories to overcome shortcomings of high-dimensional data comprehensively, the an imbalance in class data and a requirement for a strong model. Employing state-of-the-art techniques such as feature engineering, PCA, and oversampling, along with a combined model of CatBoost and SVM, the study improves the model accuracy and offers a modifiable, highly scalable solution for financial risk assessment. The framework enhances decision-making and identifies drawbacks of current risk evaluation techniques, making it useful for bringing new perspectives to U.S. financial institutions [12].

Through rigorous experimentation, this work compares different machine learning models, emphasizing the importance of interpretability in various fields, including finance, and provides valuable insights for risk management. The suggested framework enhances the precision, capacity, and semantic readability of financial risk categorization across organizations, thereby promoting a higher level of objectivity in decision-making [13].

2. Literature Review

As the complications of contemporary financial systems become more indubitable, there has been an unusual research interest in the integration of artificial intelligence (AI) into financial risk management. In this section, we review some of the existing methodologies, which characterize key advances and gaps in statistical, machine learning (ML), and hybrid methods, focusing on multi-class classification.

2.1 Preprocessing Techniques: Ensuring Data Quality

Model performance is heavily dependent on the quality of the financial datasets. Principal Component Analysis (PCA) was, and is still, a standard method for dimensionality reduction. Zhong et al. [14] proved that PCA has the capacity to mitigate overfitting and accelerate the computational efficiency in high-dimensional financial datasets. Multi-class classification problems also require an efficient method for handling class imbalance. SMOTE (Synthetic Minority Oversampling Technique) was introduced by Chawla et al. [15] has been the recommended method of oversampling the minority classes to improve model fairness. Yet, Dong et al. [16] argued that oversampling methods can result in overfitting in more complex models for financial datasets, and so careful evaluation is needed.

2.2 Statistical Models: The Foundation of Financial Risk Prediction

Logistic Regression (LR) and Discriminant Analysis have long been the prevailing statistical techniques of risk prediction in finance [17]. For example, Giovani et al. [18] stated that credit scoring with regard to logistic regression makes prediction of default probability in terms of a sigmoid function that provides interpretable results, assessing the risk calculated from a number of predictor variables. Similarly, Altman [19] made use of discriminant analysis in order to predict corporate bankruptcy, which is a procedure used by later financial research. But these methods are inherently linear and are not powerful enough to model complex, non-linear relationships that exist in big, high-dimensional financial data.

2.3 Machine Learning and Deep Learning: Expanding Predictive Capabilities

Since the advent of ML, financial risk management models have been extended in their predictive capabilities [20].

Both Random Forest (RF) as well as Gradient Boosting (e.g., XGBoost) have been widely used due to their capability to handle non-linear data very robustly. RF presented by Breiman [21] shows smaller overfitting compared to multiple features used and also offers insight into feature importance. Chen et al. [22] further optimized gradient boosting algorithms for classification problems, such as credit risk prediction, with XGBoost. Pai and Hong [23] applied SVM in stock prediction and financial distress, and found that SVM outperformed the traditional methods. Following this track, Hsu et al. [24] used SVM to forecast financial distress and showed that accuracy can be greatly improved compared with the traditional models. SVM, however, being computationally intensive, requires careful parameter tuning. Besides, deep learning architectures have surfaced as neural networks that are used to predict financial risk. Heaton et al. [25] used deep learning to predict creditworthiness and found that it surpassed the predictive power of traditional learning for signaling complex relationships. However, neural networks are 'black boxes' and this lack of interpretability, as well as regulatory compliance issues, are concerns.

2.4 AI-based Models Addressing Classification Imbalance

Liao et al. [26] suggested a model of detecting fraud in public trading companies with the help of the AAERs provided by the SEC and optimizing it through the Nonlinear Activated Beetle Antennae Search (NABAS) algorithm. NABAS effectively finds a fraud with a single search particle and adaptive gradient strategy by decreasing some loss functions. When compared with the traditional approaches, such as logistic regression and SVM-FK, NABAS, with the help of RUSBoost, has proven to be more accurate and computationally efficient when identifying financial fraud. Mena et al. [27] assessed the REMED symbolic classifier in terms of financial risk prediction based on an unbalanced Federal Deposit Insurance Corporation data sample. The research emphasizes the fact that REMED can achieve high accuracy whilst keeping the interpretability with the use of straightforward rules of decision. Relative to J48 and JRip, REMED shows a better performance on class imbalance, which makes it one of the prospective tools of interpretable and effective financial risk assessment. To address the imbalance between the two classes and deliver an efficient feature selection strategy, Amarnadh and Moparthi [28] introduced ROGENet, a new framework based on the Range-Controlled SMOTE algorithm to address the class imbalance issue, and Granular Elastic Net regression with the optimization algorithm of GENGSO to select the features. By efficiently handling sparsity and correlation of the feature as well as enhancing the accuracy of the minority class prediction to 99.4%, 99%, 98.6%, and 97.3%, they are successful at mitigating the problem to a large extent. The findings show that ROGENet significantly outperforms conventional practices in the identification of creditworthiness with greater accuracy and resilience.

Nevertheless, these studies have some limitations. The NABAS algorithm might encounter issues of scalability when used on large real-time data. The rule-based essence of REMED can make it less versatile in times of extreme volatility in the financial realm. Though ROGENet produces accurate results, in the process of minimizing the coefficients tends to add noise or overfit the results unless tuned carefully. These limitations serve as an indicator of a need to conduct additional studies so that the developed models can have a higher level of applicability and generalizability and should be balanced in terms of their complexity and interpretability.

2.5 Hybrid Models: Combining the Best of Multiple Algorithms

As a promising approach, hybrid models have emerged as methods that show how to combine the strengths of different methodologies, while mitigating their respective weaknesses [29]. A large variety of ensemble techniques (e.g., bagging, boosting, stacking) have been used to improve model robustness. LightGBM, which was developed by Ke et al. [30], has been integrated into stacking frameworks for financial applications. Wang et al. [31] demonstrated recent advancements in integrating XGBoost and LightGBM with other techniques that achieve remarkably good results on binary as well as multi-class classification tasks in a credit default prediction use case. Lu et al. **[32]** combined CatBoost with SVM and showed better accuracy and speed in an imbalanced dataset for financial risk prediction. Other advanced hybrid systems consisting of deep learning and conventional ML were also explored. An efficient framework based on convolutional neural networks (CNN) and SVM for financial fraud detection was proposed by Berhane et al. [33] and achieved good performance. Nevertheless, the computational complexity of these systems makes their practical application cumbersome. However, these hybrid architectures have not been widely applied to multiple-class financial risk classifications, including for datasets with high class imbalance [34].

In this work, we focus on addressing these gaps, proposing a new multi-class financial risk classification using a new hybrid CatBoost and SVM model. Dimensionality reduction with PCA and class balancing with SMOTE are utilized to evaluate the model on a broad dataset. Feature importance analysis is detailed, and relevance is emphasized, leading to a proposed framework suitable for real-world financial risk management.

3. Data and Methodology

The methodology involves the use of a detailed set of financial data to categorize risk, incorporating feature engineering, Principal Component Analysis (PCA), and oversampling techniques. The risk assessment incorporated in the AI framework utilizes hybrid models such as CatBoost and SVM to achieve operational and scalable interpretability. The overview of our proposed methodology is shown in Figure 1.

3.1 Dataset Description

The dataset used in this study was created for financial risk scoring and consists of 20,000 records obtained from Kaggle (https://www.kaggle.com/datasets/lorenzozoppelletto/financial-risk-for-loan-approval/data). It includes, but is not limited to, demographic details, character, credit history, employer details, income, assets, liabilities, and other details. The dataset serves two primary purposes: combining a constant predictor to assign a risk score for loan default and a binary classifier for loan approval results. Such basic predictor attributes include applicant age, credit score, loan amount, income, and the debt-to-income ratio, along with the loan approval status, making it easy to create models in the field of financial risk management. The dataset used in our study is outlined in Table 1.

Table 1. Summary of the Key Details of the Dataset

Dataset	Source	Number of Columns	Total Samples	Classification Type
Financial Risk for Loan Approval	Kaggle	34	20000	Binary Classification (Loan Approval)

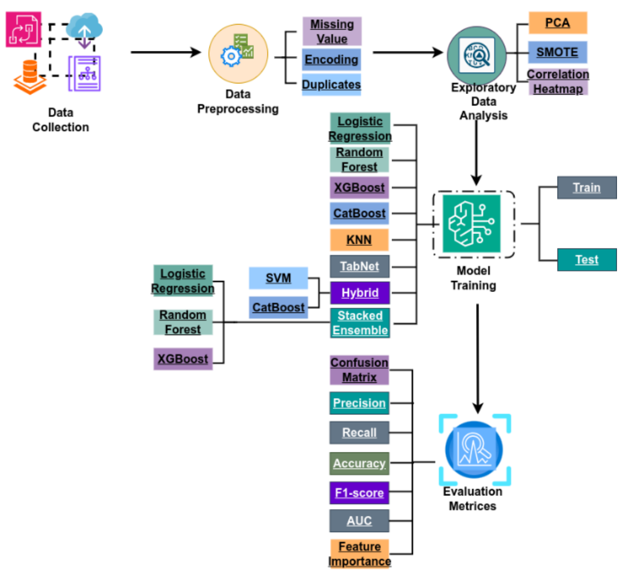


Figure 1. Structure of the overall workflow of this study

3.2 Dataset Preprocessing

Before implementing a profuse number of machine learning designs in this research, several essential procedures were performed on the dataset, which are detailed in the next section of this paper. Firstly, some columns with no use or with repetitive information in the dataset were excluded. Subsequently, missing values were also identified and managed through imputation, where the mean was applied to the respective missing data columns without removing any values. The data was also ensured to be free of duplicate samples in order to eliminate any unhealthy samples. Moreover, other features such as 'EmploymentStatus', 'HomeOwnershipStatus', 'EducationLevel','MaritalStatus', and 'LoanPurpose' were one-hot encoded because these nominal variables require numerical values to support machine learning algorithms using LabelEncoder [35]. The equation for LabelEncoder comes out as

$$X_{encoded} = LabelEncoder(X) = \{y_1, y_2, \ldots, y_n\} \quad (1)$$

Where each y_i is a unique integer corresponding to the category x_i.

The outcome of the presented preprocessing steps was that the final dataset was clean, highly formatted, and suitable for both further analysis as well as model training.

3.3 Exploratory Data Analysis

When addressed in the Exploratory Data Analysis (EDA) part of this research, several methods were employed to handle issues such as high dimensionality and data imbalance.

1) **SMOTE (Synthetic Minority Over-sampling Technique):** To deal with the problem of imbalance, SMOTE was used to synthesize new samples for the minority class. Before undersampling, the data was distributed unevenly, with 15,220 instances in class 0 and 4,780 instances in class 1. Applying SMOTE resulted in nearly balanced classes, with 12,237 samples in each class, as shown in Table 2. This resampling strategy helped to cancel model bias towards the majority class, increase the probability of a correct forecast for the minority class; and, therefore, enhance the suitability of the database for training models.

Table 2. Class Distribution Before And After Applying Smote

Before Resampling	After Resampling

Class: 0	Class: 1	Class: 0	Class: 1
15220	4780	12237	12237

2) **Principal Component Analysis (PCA):** As a result of applying PCA, it was possible to visualize class distribution from the reduced dimensions of the dataset. The distribution of the points for the two classes of financial risk was clearly distinguishable in the pairwise PCA plot; the 0 class had a more compact distribution than the 1 class. This visualization effectively showcased the separability of the classes, as well as potential features that have been impacting the variance of the data, as illustrated in

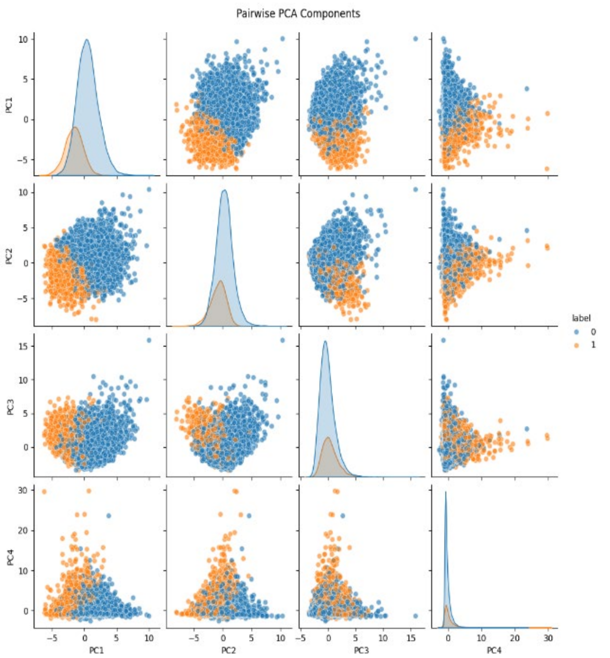


Figure 2. Pairwise PCA Plot

Figure 2. In addition, using PCA, more information about potential outliers was obtained, which helped to improve the model when classifying financial risks.

3) **Heatmap Correlation:** As feature correlations were examined, a heatmap was used to identify multicollinearity. The number of highly correlated features were subjected to removal or transformation as including all of them in the model reduces the model's prediction capability by including features with high correlation. The correlation heatmap is presented in Fig, where the connections between features in the dataset were observed.

3.4 Machine Learning Models

Logistic Regression (LR): A type of regression having to do with predicting a binary outcome based on input features. It produces a probability of an event according to a logistic function of a linear combination of these input variables [36]. It is often the case that the independent variables, used in logistic regression models, may be skewed, and such models are designed for non-linear analysis [37]. For our study, we have fixed the parameters for LR are 'C': [0.01, 0.1, 1, 10, 100], 'penalty': ['l1', 'l2'], and 'solver': ['liblinear', 'saga']. The equation for LR can be expressed as,

$$P(y = 1|X) = \frac{1}{1+e^{-(b_0+b_1X_1+b_2X_2+\cdots+b_nX_n)}} \qquad (2)$$

Random Forest (RF): RF, proposed by Leo Breiman, is an ensemble learning method (a combination of Bagging and random feature selection) where multiple uncorrelated decision trees are concatenated as a classifier. Each tree is created using the Bootstrap technique, which involves a random subset of features and random samples for training each tree. The final prediction is determined by majority voting across all trees, and the trees do not require pruning [38]. For a test set X with C categories and T decision trees, the output is:

$$H(X) = \arg max \{\textstyle\sum_{t=1}^{T} I(h_t(X) = y)\} \qquad (3)$$

where I is an indicator function that checks if the prediction matches the true class.

XGBoost: Extreme Gradient Boosting (XGBoost, Chen and Guestrin [22] is an incredibly powerful ensemble learning technique that utilizes the gradient boosting framework. It sequentially decides about the cost of an error and the classification accuracy. The major strength of XGBoost is its ability to transform weak classifiers into a strong ensemble model, rendering it more efficient and flexible than single models. By applying the algorithm, the classification performance is improved by generating multiple decision trees in multiple iterations and predicting the trees' model sum [39]. The model prediction is represented as:

$$\hat{y} = \textstyle\sum_{m-1}^{M} f_m(x), \;\; f_m \in \mathcal{F} \qquad (4)$$

where f_m is a decision tree in the ensemble and $\mathcal{F}$ is the space of possible trees.
In this study, hyperparameter tuning in XGBoost was performed using RandomizedSearchCV to optimize the variables of `n_estimators` (100–500), `learning_rate` (0.01–0.2), `max_depth` (3–10), and the regularization terms (`reg_alpha`, `reg_lambda`). The best combination is identified through cross-validation and random sampling, ensuring a well-calibrated, high-performance model.

CatBoost: To improve of accuracy and stability, CatBoost has introduced a prior distribution term into the Greedy Target-based Statistics. Such a method helps avoid account noise and low-frequency category data during the model learning, allowing it to deal with unseen feature values appropriately and improve the alignment to the prediction target [40]. The prediction at the t-th iteration is given by:

$$F_t(x) = F_{t-1}(x) + \eta \cdot h_t(x) \qquad (5)$$

Where $F_t(x)$ is the prediction after the t-th iteration, $F_{t-1}(x)$ is the prediction from the previous iteration, η is the learning rate controlling the update size and $h_t(x)$ is the new decision tree added at iteration t.

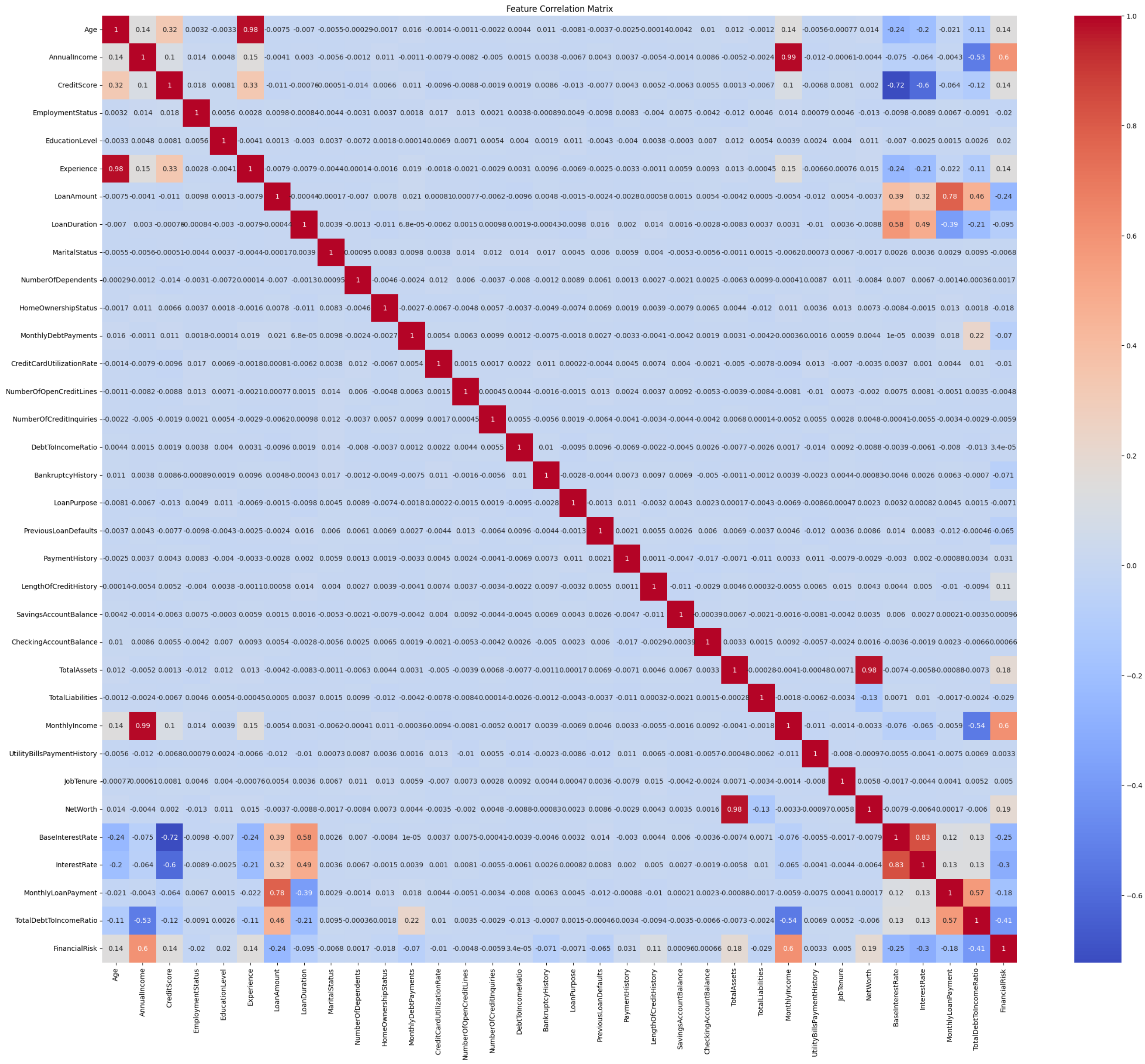


Figure 3. The Feature Correlation Heatmap

KNN: k-Nearest Neighbors (KNN) learns about the data and makes predictions based on a distance; in this case, Euclidean or Manhattan distance is likely to be used. It allocates the test sample to the class that is most frequent amongst its neighbors. Therefore, it does not require a training phase; all it needs for prediction is the stored data [41]. The prediction is given by:

$$y = mode(y_1, y_2, \dots, y_k) \quad (6)$$

Where, $y_1, y_2, \dots, y_k$ are the labels of the k-nearest neighbors.

Stacked Ensemble (LR-RF-XGBoost) Model: The principle of the ensemble method is to utilize multiple models to construct a more comprehensive and accurate model for forecasting financial variables [41]. This has the added advantage of reducing misclassified examples and boosting the results of single classifiers by combining them in groups. In the present work, the meta-learner was used, and it combines any number of base-learners, including RF, LR, and XGBoost, to obtain better classification accuracy [42]. The equation for a stacked ensemble model can be represented as follows: Let the base models be $f_1(X), f_2(X), \dots, f_n(X)$, here we shall denote X as the input features to be included in the model, n is the number of base models. The prediction of these base models is then calculated using a meta-learner g to create the final score of the prediction.

$$y = g(f_1(X), f_2(X), \ldots, f_n(X)) \quad (7)$$

Hybrid (CatBoost+SVM): We introduced a Hybrid model-based approach that enables decision-making to achieve a continuous gradation of the transition toward the use of ML solutions without abruptly shifting to a completely different approach. This hybrid model should be used to enable decision-making to control the role of the ML model at the time of decision-making as the system transitions [43]. The experiments demonstrated that the increased application of the ML-based solution enhances overall decision-making effectiveness, thereby improving financial risk management. Namely, the proposed hybrid model, which integrates the features of CatBoost and SVM in a parallel configuration, outperforms the individual predictors in financial risk evaluation. The hyperparameter distribution is considered according to TABLE 3.

Table 3. Hyperparameters for Hybrid Model

iterations	Learning rate	depth	Random seed	verbose
500	0.05	6	42	0

TabNet: TabNet [44] is a new deep learning architecture that can solve tabular data, being both highly performant and interpretable, which differentiates it from unexplainable neural nets. Thus, TabNet not only brings a length-wise improvement to models, but also provides an interpretable process for how the predictions are made, especially for false negative minimization such as the default prediction. It also provides a flexible performance based on the size and quality of the dataset and is intended for various data types and characteristics. Additionally, TabNet imposes significantly less data preprocessing, addresses null values, handles different data types and is faster in training compared to tree-based methods. Founded on attention mechanisms and decision trees, the TabNet model is specifically designed for tabular data. While the exact formulation is more complex due to the layers and attention, a simplified representation of the core decision process can be written as:

$$\hat{y} = f(X) = Attn(X) \cdot W + b \quad (8)$$

Where, X is the input feature matrix, $Attn(X)$ represents the attention mechanism applied to the input, W is the learned weight matrix, b is the bias term, and $\hat{y}$ is the predicted output.

4. Results and Discussions

4.1 Performance Comparison of Different Models

In Table 4, we showed the performance of various AI-based models used to predict large-scale financial risk in U.S. markets. Each model was then evaluated on metrics like accuracy, precision, recall, and F1 score, which give away great insights as to the model's ability to deal with the multi-class classification task. We report an accuracy of 94.88% and an F1-score of 0.93, indicating Logistic Regression is a reliable baseline model. However, its sensitivity and specificity are inadequate, with precision and recall metrics (0.93). Despite Random Forest having an accuracy of 92.88%, precision of 0.90, and F1-score of 0.91, it was unable to efficiently tackle the dataset's complexity.

Through XGBoost, which has already become a robust model, an accuracy of 95.8% was achieved combined with an F1-score of 0.94, scoring high on precision (0.94) and recall (0.94). What's more, XGBoost could be beaten by CatBoost, which got an accuracy of 95.85% and an F1-score of 0.95, showing a strong predictive ability and balance across all metrics. In contrast, K-Nearest Neighbors (KNN) did not deliver on its promise, achieving only 85.50% accuracy, with lower precision (0.80) and recall (0.79), as it is unable to handle high-dimensional financial data.

Using Logistic Regression, Random Forest, and XGBoost in the stacked ensemble model gave us 94.92% accuracy and 0.93 F1 score, but it can't outperform the standalone CatBoost or the hybrid model. The best-performing approach was the hybrid model in which CatBoost and SVM are combined together and this gives the best accuracy (95.93%) and F1 score (0.95). The integration of gradient boosting and hyperplane separation techniques with precision and recall metrics (0.95) shows how such integration has the ability to leverage the complementary strengths of gradient boosting and hyperplane separation techniques. Although innovative, TabNet did not come close to the accuracy level of 94.67%

Table 4. Performance Metrics Of Different AI-Powered Models For Predicting Financial Risk In The U.S. Markets

Model	Accuracy (%)	Precision	Recall	F1-score
LR	94.88	0.93	0.94	0.93
RF	92.88	0.90	0.91	0.91
XGBoost	95.38	0.94	0.94	0.94
CatBooost	95.85	0.94	0.95	0.95
KNN	85.50	0.82	0.79	0.80
Stacked Ensemble	94.92	0.93	0.93	0.93

(LR-RF-XGBoost)				
Hybrid (CatBoost +SVM)	**95.93**	**0.95**	**0.95**	**0.95**
TabNet	94.67	0.92	0.94	0.93

A comparative visualization of the performance metrics (accuracy, precision, recall, and F1-score) of the evaluated models to predict financial risk is presented in Figure 4. This graph represents the strengths and weaknesses of the model outputs in one figure and is nice because it is easy to understand their predictive capabilities.

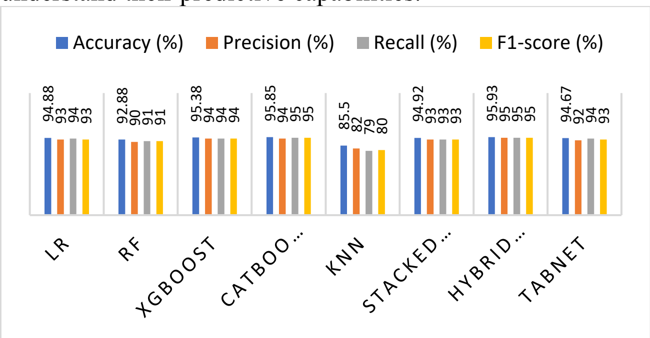


Figure 4. Comparison of performance metrics across different ML models

4.2 Confusion Matrix

Figures 5 and 6 reveal confusion matrices for the AI-driven models we have evaluated: Logistic Regression (LR), Random Forest (RF), XGBoost, CatBoost, KNN, Stacked Ensemble, Hybrid (CatBoost+SVM), and TabNet to predict the financial risk of U.S. markets. Through a detailed comparative analysis of these models, we offer insights into their strengths and weaknesses in appropriately classifying risk and non-risk cases, which are essential for the management of the U.S. financial landscape's financial stability.

Logistic Regression (Figure 5) has a fair ability to identify financial risks with 949 true positives (TP) and 68 of the 950 financial risks being correctly predicted (false negatives, FN), out of the 1,950 total predicted instances. Nevertheless, the 137 false positives (FP) indicate a conspicuous skewed tendency to label stable market conditions as at risk, thereby generating excessive caution and hence overreaction to financial problems. Likewise, the Random Forest Classifier (Figure 5) works well and is only outperformed by Logistic Regression (FN = 118 and FP = 167), which makes it less dependable for sensitive financial risk assessment in volatile U.S. markets. However, for example, XGBoost (Figure 5) only requires 78 FN and only 107 FP, which provides more favorable results in the trade-off between sensitivity and specificity. Because precision is essential in U.S. financial sector decision-making, this model is particularly effective in reducing false alarms. However, the most effective standalone model is CatBoost (Figure 5), with the lowest in the counts of FN (62) and FP (104), implying high trustworthiness in a risk case or a non-risk case detection. The superior performance of CatBoost makes it a robust solution for financial risk prediction in those highly dynamic U.S. markets.

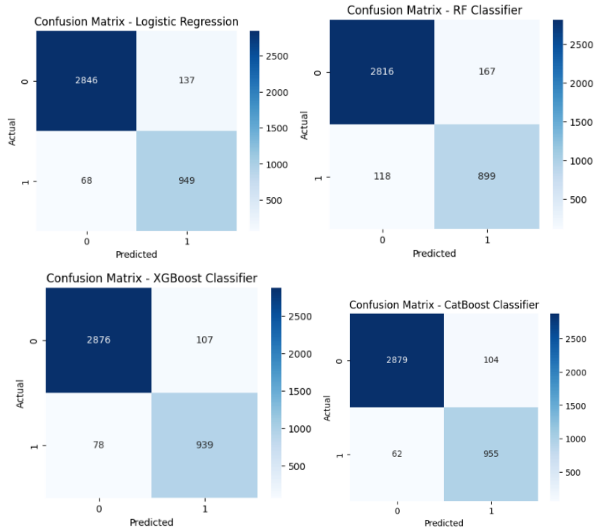
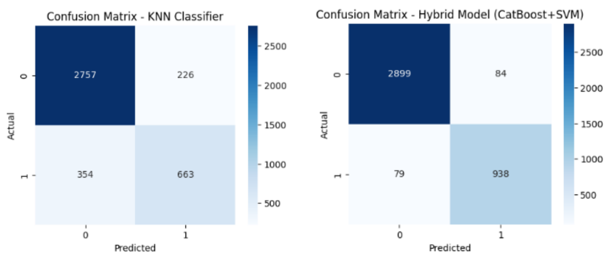


Figure 5. Confusion matrices for Financial Risk prediction models: LR, RF, XGBoost, and CatBoost Classifier

Figure 6 shows that among models, KNN is a poor performer with 226 FP and 354 FN, which is a severe unit weakness in this model's predictive capability. However, such high error rates make it unsuitable for financial risk management in U.S. markets (because accurate prediction is essential to maintain market confidence and stability). On the other hand, the Stacked Ensemble model (Figure 6) combines the strengths of LR, RF, and XGBoost and achieves one of the best performances. It has an amazingly low FP count (97) and a competitive FN count (106), which proves it is able to avoid making false alarms while still providing a robust risk detection rate. From the Hybrid Model, Figure 8 shows the top performing one, achieving the best true negative count (2899), best FP count (84), and also the best FN count (79).



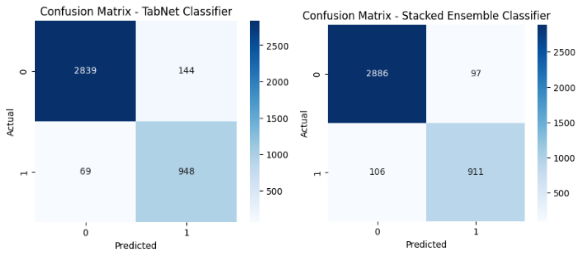


Figure 6. Confusion matrices for Financial Risk prediction models: KNN, Stacked Ensemble (LR-RF-XGBoost), Hybrid (CatBoost+SVM), and TabNet Classifier

With its ability to provide a good balance between specificity and sensitivity, this model is especially useful for predicting financial risks in the U.S., where sharply underpredicted or overpredicted risks will have wide-ranging economic consequences. Combining a Hybrid Model and a TabNet Classifier also shows strong performance (Figure 6), with a lower FN count (69) but slightly higher FP count (144) than CatBoost and the Hybrid Model, indicating a similar, albeit slightly suboptimal, performance for handling non-risk categorizations.

4.3 Feature Importance Analysis

To understand the major factors contributing to the financial risk prediction, we performed an analysis of feature importance with the help of the CatBoost component, a component of the hybrid CatBoost+SVM model. The findings, as depicted in Figure 7, show the comparative importance of every feature in the decision-making process of the model.

Interestingly, 'TotalDebtToIncomeRatio', 'InterestRate', and 'LoanDuration' were found to be the best predictors, which means that the burden of debt on the borrower, the lending rates, and the terms of the loan largely determine the credit risk. Probably these factors indicate the financial soundness of the borrower and his/her ability to repay. Other minor contributors are 'MonthlyIncome', 'NetWorth', and 'AnnualIncome', which reveal how the model is sensitive to income and asset-based characteristics. The variables connected to credit, which include 'CreditScore', 'LengthOfCreditHistory', and 'TotalAssets', are also found to be quite significant, which matches the official approaches to financial risk management.

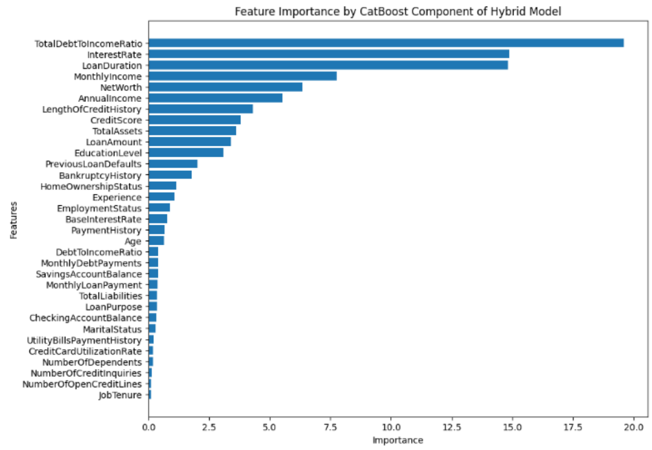


Figure 7. Feature importance analysis for financial risk prediction using the CatBoost component of the Hybrid Model (CatBoost+SVM), highlighting the most influential factors in U.S. markets

Conversely, the features like 'JobTenure', 'NumberOfOpenCreditLines', and 'CreditCardUtilizationRate' did not significantly contribute to the prediction model, implying that they were either not strongly related to default risk or they were redundant because multicollinear with more important variables. Such an analysis not only proves that the model aligns with financial intuition, but it also gives visibility into the factors most proximate to the U.S. credit market, which further proves the explainability of the hybrid manner.

The CatBoost component's ability to identify critical variables in this feature importance analysis not only validates the effectiveness of the CatBoost component in predicting financial risks in the U.S. markets but also is consistent with financial knowledge, strengthening the model's legitimacy. These insights could enable financial institutions to place the greatest emphasis on the most important risk factors, refine their risk assessments, and develop concentrated areas of mitigation to the threat of financial vulnerabilities.

4.4 Model Training and Computational Efficiency

In Table 5, we compared training times for different types of AI-powered models for financial risk prediction in U.S. markets. The result that has the best performance with evaluation metrics like confusion matrices, ROC-AUC scores, and feature importance analysis is the Hybrid Model (CatBoost+SVM), which takes 11 seconds to train. The additional training time is around this moderately higher value, with KNN (0.5 seconds) and Random Forest (9 seconds) as simpler models, as a consequence of the cost incurred on hyperparameter tuning and the use of CatBoost for feature extraction and SVM for classification.

Table 5. Training Time For Different AI-Powered Models

Model	Training Time (s)
LR	95
RF	9
XGBoost	155
CatBoost	4
KNN	0.5
Stacked Ensemble	220
Hybrid	11
TabNet	56

Due to their complexity, ensemble models like XGBoost (155 seconds) and Stacked Ensemble (220 seconds) take a lot of time, though they are still performing similarly. For instance, both the models Logistic Regression (which takes 95 seconds to train) and TabNet (56 seconds to train), generate a balance between training time and accuracy.

Interestingly enough, CatBoost, alone, runs so efficiently (4 seconds of training time) that it seems to be the best choice in case of computational efficiency. Yet its slightly higher training time is outweighed by its superior predictive capability in U.S. financial risk prediction, which makes it the most preferred model in this application.

4.5 ROC Curve and AUC Score

A comparative evaluation of the performance of different models used to predict financial risk in US markets is conducted based on the Receiver Operating Characteristic (ROC) curve and the corresponding Area Under Curve (AUC) scores (Figure 8).

Logistic Regression, XGBoost, and CatBoost also show excellent performance with an AUC score of 0.99, a very near-perfect distinction between positive and negative classes. These models work wonderfully to minimize both false positives and false negatives, making them very reliable models for financial risk analysis.

The second model, Random Forest, also does well with its AUC score of 0.98, while not as high as the above-mentioned model, but still quite predictive based on the result. On the other hand, a score of AUC as high as 0.95 is achieved by the Hybrid Model (CatBoost+SVM), which represents a compromise between sensitivity and specificity.

AUC score of 0.89 is based on the K-Nearest Neighbors (KNN) model, which is much less than any other models, actually indicating that it has much less discrimination power as opposed to others. We should note that neither the KNN model can handle complex financial datasets promisingly.

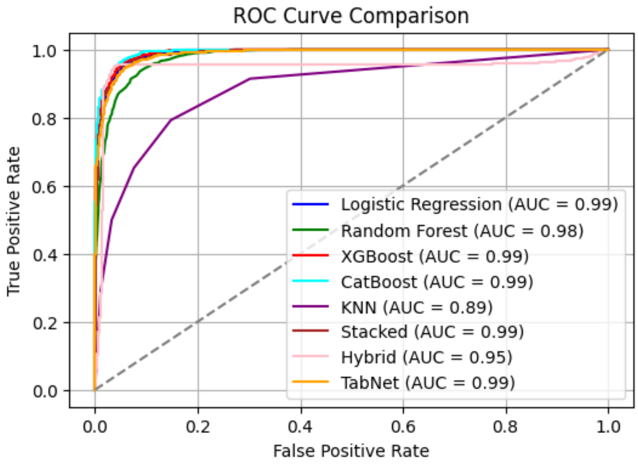


Figure 8. ROC curve comparison with AUC scores for various models in financial risk prediction in the U.S. markets

Overall, ROC curve comparison shows that advanced ensemble models (CatBoost, XGBoost, and Tabnet) perform better than the global ensemble model (AdaBoost) in high-precision finance risk prediction in US markets. These results validate the value of utilizing sophisticated algorithms in order to improve predictive accuracy in financial analytics.

4.6 Discussion

We evaluated advanced machine learning approaches for predicting financial risk models for U.S. markets, showing high predictive accuracy. Among the tested models, we observe that the Hybrid Model (CatBoost+SVM) consistently achieved a robust balance across all the metrics, outperforming others. With superior precision and recall, its AUC score of 0.95 makes it fit for real-world use in applications to complex financial datasets. The hybrid model integrates SVM for classification, dominated by its performance, and CatBoost for feature extraction, dominated by interpretability, enabling it to benefit from both interpretability and classification strength in its risk management applications.

While KNN (0.5 seconds) training times are faster than our proposed models, their performance was significantly behind, with its lower AUC score (0.89) and confusion matrix metrics. Fun with models such as Stacked Ensemble and XGBoost gave great AUC scores (0.99) but took more time to train (220 seconds), spending the most time in training. While it is trained slower than the other models thanks to hyperparameter tuning (11 seconds), the Hybrid Model achieves this trade-off by providing a good predictive performance with respect to time and computational demand, and by striking a good balance between the two.

Finally, blending feature extraction with a powerful classification method is shown to be beneficial for financial risk prediction through the Hybrid Model. It solves the problem of a scalable, interpretable, and effective solution to predictive analytics in dynamic financial environments, creating a benchmark for advancing predictive analytics in future dynamic financial environments.

5. Conclusion

The findings from this study highlight how transformative predictive analytics enabled through AI can be for helping financial institutions cope with the challenges of their U.S. financial risk management business. The research provides a robust solution by building a hybrid CatBoost and SVM model, which is highly accurate, computationally efficient, and interpretable. Since this also ensures balanced and reduced-dimensional data and scalability for real-world applications, we combined these with advanced techniques such as SMOTE and PCA. Feature importance analysis further improves the practicality of this approach by highlighting critical factors for driving financial risk, which are 'TotalDebtToIncomeRatio' and 'InterestRate', and therefore can drive targeted and effective risk mitigation strategies.

Nevertheless, the proposed hybrid model may be less effective in situations involving a highly volatile or non-stationary financial environment, where significant changes in the data pattern can hinder the model's generalization capacity. Moreover, the computational overhead of the CatBoost and SVM blend can limit its use in high-performance financial institutions, such as those with large-scale or low-latency systems.

Aside from its technological contributions, this study provides practical policy implications useful to policymakers, analysts, and financial institutions with regard to the role of combining AI capabilities with domain-specific knowledge. The proposed framework is adaptive to changing financial environments and should serve as a keystone to future improvements in predictive analytics and risk management. It paves the way for further development of AI in intermediation financial systems and for continuous innovation and development in coping with the multifaceted sources of financial risks.

References

[1] J. Xiao, J. Wang, W. Bao, S. Bi, and T. Deng, "Research on the application of data analysis in predicting financial risk," *Financial Engineering and Risk Management,* vol. 7, no. 4, pp. 183–188, 2024.

[2] M. S. Murugan and S. K. T, "Large-scale data-driven financial risk management & analysis using machine learning strategies," *Measurement: Sensors,* vol. 27, p. 100756, 2023/06/01/ 2023, doi: https://doi.org/10.1016/j.measen.2023.100756.

[3] K. Valaskova, T. Kliestik, L. Svabova, and P. Adamko, "Financial Risk Measurement and Prediction Modelling for Sustainable Development of Business Entities Using Regression Analysis," *Sustainability*, vol. 10, no. 7, doi: 10.3390/su10072144.

[4] J. Grable and R. H. Lytton, "Financial risk tolerance revisited: the development of a risk assessment instrument☆," *Financial Services Review,* vol. 8, no. 3, pp. 163–181, 1999/01/01/ 1999, doi: https://doi.org/10.1016/S1057-0810(99)00041-4.

[5] S. Ahmed, M. M. Alshater, A. E. Ammari, and H. Hammami, "Artificial intelligence and machine learning in finance: A bibliometric review," *Research in International Business and Finance,* vol. 61, p. 101646, 2022/10/01/ 2022, doi: https://doi.org/10.1016/j.ribaf.2022.101646.

[6] M. C. Dyball and R. Seethamraju, "Client use of blockchain technology: exploring its (potential) impact on financial statement audits of Australian accounting firms," *Accounting, Auditing & Accountability Journal,* vol. 35, no. 7, pp. 1656–1684, 2022, doi: 10.1108/AAAJ-07-2020-4681.

[7] T. P. Nugrahanti, "Analyzing the evolution of auditing and financial insurance: tracking developments, identifying research frontiers, and charting the future of accountability and risk management," *West Science Accounting and Finance,* vol. 1, no. 02, pp. 59–68, 2023.

[8] S. Pattyam, "'Ai-driven financial market analysis: Advanced techniques for stock price prediction, risk management, and automated trading," *African Journal of Artificial Intelligence and Sustainable Development,* vol. 1, no. 1, pp. 100–135, 2021.

[9] H. A. Javaid, "Ai-driven predictive analytics in finance: Transforming risk assessment and decision-making," *Advances in Computer Sciences,* vol. 7, no. 1, 2024.

[10] M. Schmitt, "Automated machine learning: AI-driven decision making in business analytics," *Intelligent Systems with Applications,* vol. 18, p. 200188, 2023/05/01/ 2023, doi: https://doi.org/10.1016/j.iswa.2023.200188.

[11] B. Biswas, "IT in Improving Integrity and Productivity in Supply Chain Management.," *Transactions on Banking, Finance, and Leadership Informatics,* vol. 1, no. 1, pp. 1–7, 2024, doi: https://doi.org/10.63471/tbfli24001.

[12] M. J. H. Mani Prabha, Jarin Tias Meraj, "AI-Driven Financial Security: Innovations in Protecting Assets and Mitigating Risks.," *Advances in Machine Learning, IoT and Data Security,* vol. 1, no. 1, pp. 14–23, 2024, doi: https://doi.org/10.63471/amlids24004.

[13] M. A. A. M. Md Saddam Hosain, Dr. Joseph P. Siegmund, "Systemic Risk and Financial Stability: Measurement and Policy Implications.," *Transactions on Banking, Finance, and Leadership Informatics,* vol. 1, no. 1, pp. 8–11, 2024, doi: https://doi.org/10.63471/tbfli24002.

[14] X. Zhong and D. Enke, "Forecasting daily stock market return using dimensionality reduction," *Expert Systems with Applications,* vol. 67, pp. 126–139, 2017/01/01/ 2017, doi: https://doi.org/10.1016/j.eswa.2016.09.027.

[15] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *J. Artif. Int. Res.,* vol. 16, no. 1, pp. 321–357, 2002.

[16] H. Dong, R. Liu, and A. W. Tham, "Accuracy Comparison between Five Machine Learning

Algorithms for Financial Risk Evaluation," *Journal of Risk and Financial Management*, vol. 17, no. 2, doi: 10.3390/jrfm17020050.

[17] A. H. Md. Mehedi Hasan, "Cybersecurity Strategies for Businesses: Protecting Data in a Digital World.," *Advances in Machine Learning, IoT and Data Security,* vol. 1, no. 1, pp. 24–29, 2024, doi: https://doi.org/10.63471/amlids24005.

[18] G. Valdrighi *et al.*, "Best Practices for Responsible Machine Learning in Credit Scoring," *arXiv preprint arXiv:2409.20536,* 2024.

[19] E. I. Altman, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *The Journal of Finance,* vol. 23, no. 4, pp. 589–609, 1968, doi: 10.2307/2978933.

[20] M. K. K. S. Syed Nazmul Hasan, Oli Ahammed Sarker, Md Redwan Hussain, Jarin Tias Meraj, "Blockchain Based Security Solutions for Banking Information Technology," *Transactions on Banking, Finance, and Leadership Informatics,* vol. 1, no. 1, pp. 25–29, 2024, doi: https://doi.org/10.63471/tbfli24005.

[21] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, no. 1, pp. 5–32, 2001/10/01 2001, doi: 10.1023/A:1010933404324.

[22] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, 2016. [Online]. Available: https://doi.org/10.1145/2939672.2939785.

[23] P.-F. Pai and W.-C. Hong, "Support vector machines with simulated annealing algorithms in electricity load forecasting," *Energy Conversion and Management,* vol. 46, no. 17, pp. 2669–2688, 2005/10/01/ 2005, doi: https://doi.org/10.1016/j.enconman.2005.02.004.

[24] M.-F. Hsu and P.-F. Pai, "Incorporating support vector machines with multiple criteria decision making for financial crisis analysis," *Quality & Quantity,* vol. 47, no. 6, pp. 3481–3492, 2013/10/01 2013, doi: 10.1007/s11135-012-9735-y.

[25] J. B. Heaton, N. G. Polson, and J. H. Witte, "Deep learning for finance: deep portfolios," *Applied Stochastic Models in Business and Industry,* vol. 33, no. 1, pp. 3–12, 2017/01/01 2017, doi: https://doi.org/10.1002/asmb.2209.

[26] B. Liao, Z. Huang, X. Cao, and J. Li, "Adopting Nonlinear Activated Beetle Antennae Search Algorithm for Fraud Detection of Public Trading Companies: A Computational Finance Approach," *Mathematics*, vol. 10, no. 13, doi: 10.3390/math10132160.

[27] L. J. Mena *et al.*, "Enhancing financial risk prediction with symbolic classifiers: addressing class imbalance and the accuracy–interpretability trade–off," *Humanities and Social Sciences Communications,* vol. 11, no. 1, p. 1540, 2024/11/14 2024, doi: 10.1057/s41599-024-04047-5.

[28] V. Amarnadh and N. R. Moparthi, "Range control-based class imbalance and optimized granular elastic net regression feature selection for credit risk assessment," *Knowledge and Information Systems,* vol. 66, no. 9, pp. 5281–5310, 2024/09/01 2024, doi: 10.1007/s10115-024-02103-9.

[29] M. R. M. Md. Mehedi Hasan, " Cloud Computing in Banking Flexibility and Scalability for Financial Institute.," *Transactions on Banking, Finance, and Leadership Informatics,* vol. 1, no. 1, pp. 17–24, 2024, doi: https://doi.org/10.63471/tbfli24004.

[30] G. Ke *et al.*, "LightGBM: a highly efficient gradient boosting decision tree," presented at the Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA, 2017.

[31] Y. Wang, Z. Wu, J. Gao, C. Liu, and F. Guo, "A multi-level classification based ensemble and feature extractor for credit risk assessment," (in eng), *PeerJ Comput Sci,* vol. 10, p. e1915, 2024, doi: 10.7717/peerj-cs.1915.

[32] H. Lu and X. Hu, "Enhancing Financial Risk Prediction for Listed Companies: A Catboost-Based Ensemble Learning Approach," *Journal of the Knowledge Economy,* vol. 15, no. 2, pp. 9824–9840, 2024/06/01 2024, doi: 10.1007/s13132-023-01601-5.

[33] T. Berhane, T. Melese, A. Walelign, and A. Mohammed, "A Hybrid Convolutional Neural Network and Support Vector Machine-Based Credit Card Fraud Detection Model," *Mathematical Problems in Engineering,* vol. 2023, no. 1, p. 8134627, 2023/01/01 2023, doi: https://doi.org/10.1155/2023/8134627.

[34] M. S. I. C. Rabeya Khatoon, "Predictive Analytics in Customer Relationship Management in the USA.," *Journal of Business Venturing, AI and Data Analytics,* vol. 1, no. 1, pp. 37–43, 2024, doi: https://doi.org/10.63471/jbvada24005.

[35] Y. N. Prajapati and M. Kumar, "A review paper on cause of heart disease using machine learning algorithms," *Journal of Pharmaceutical Negative Results,* pp. 9250–9259, 2022.

[36] U. Altunöz, "Prediction of Banking Credit Risk Using Logistic Regression and The Artificial Neural Network Models: A Case Study of English Banks," *Sosyal Araştırmalar ve Davranış Bilimleri,* vol. 10, no. 21, pp. 862–887, 2024.

[37] K.-H. Shih, C.-C. Cheng, and Y.-H. Wang, "Financial information fraud risk warning for manufacturing industry-using logistic regression and neural network," *Romanian Journal of Economic Forecasting,* vol. 14, no. 1, pp. 54–71, 2011.

[38] Z. Zheng, "Financial Risk Early Warning Model Combining SMOTE and Random Forest for Internet Finance Companies," *J. Cases Inf. Technol.,* vol. 26, no. 1, pp. 1–21, 2024, doi: 10.4018/jcit.356504.

[39] B. Tan, Z. Gan, and Y. Wu, "The measurement and early warning of daily financial stability index based on XGBoost and SHAP: Evidence from China," *Expert Systems with Applications,* vol. 227, p. 120375, 2023/10/01/ 2023, doi: https://doi.org/10.1016/j.eswa.2023.120375.

[40] Z. Zhuo, "Research on Credit Card Overdue Risk Prediction based on CatBoost Model," " presented at the in Proceedings of the 5th Management Science Informatization and Economic Innovation Development Conference, 2024.

[41] M. Abdar *et al.*, "A new nested ensemble technique for automated diagnosis of breast cancer," *Pattern Recognition Letters,* vol. 132, pp. 123–131, 2020/04/01/ 2020, doi: https://doi.org/10.1016/j.patrec.2018.11.004.

[42] M. A. Muslim *et al.*, "New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning," *Intelligent Systems*

with Applications, vol. 18, p. 200204, 2023/05/01/ 2023, doi: https://doi.org/10.1016/j.iswa.2023.200204.
[43] Á. González-Prieto, A. Brú, J. C. Nuño, and J. L. González-Álvarez, "Hybrid machine learning methods for risk assessment in gender-based crime," *Knowledge-Based Systems,* vol. 260, p. 110130, 2023/01/25/ 2023, doi: https://doi.org/10.1016/j.knosys.2022.110130.
[44] K. McDonnell, F. Murphy, B. Sheehan, L. Masello, and G. Castignani, "Deep learning in insurance: Accuracy and model interpretability using TabNet," *Expert Systems with Applications,* vol. 217, p. 119543, 2023/05/01/ 2023, doi: https://doi.org/10.1016/j.eswa.2023.119543.