SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I10P107

Volume 12 Issue 10, 64-72, October 2025

Original Article

A Hybrid AI Model Integrating Decision Trees, LSTM and XGBoost for Financial Risk Forecasting

Nisa Vinodkumar[1], Archana Anandapadmanabhan[2,3]

[1]*Department of Management, College of Business Administration, Princess Nourah bint Abdulrahman University, Saudi Arabia.*
[2]*Turiba University, Riga, Latvia, Europe.*
[3]*Riga Management College, Riga, Europe.*

[1]*Corresponding Author : nisavinodkumar6@gmail.com*

Received: 07 August 2025 Revised: 09 September 2025 Accepted: 08 October 2025 Published: 31 October 2025

Abstract *- The financial sector has been able to automate and improve decision-making processes, especially in risk forecasting, owing to the rapid advancement of Artificial Intelligence (AI). This paper proposes a hybrid AI-based model. It uses XGBoost for strong predictions, Decision Trees (DTs) for sorting data based on rules, and Long Short-Term Memory (LSTM) networks to find patterns over time in financial data. In particular, in highly unpredictable market circumstances, it is required to improve the reliability and precision of financial risk assessment. To train and test the model, a publicly accessible financial dataset that includes records of past loans and transactions is employed. Then its performance is compared to that of conventional single-model methods. When looking at traditional AI models like logistic regression, random forest, and standalone LSTM, the results confirm that the proposed new hybrid model works better, predicts more accurately, and is cheaper. This study emphasizes the possibility of using sequential and ensemble learning frameworks to build data-driven financial forecasting systems that are knowledge-based and impartial.*

Keywords *- Decision Tree, Financial Risk Forecasting, Hybrid AI, LSTM, Time-Series Prediction, XGBoost.*

1. Introduction

For timely risk assessment, strong forecasting frameworks are required owing to the nonlinear and volatile nature of financial markets. This kind of complexity is a common challenge for traditional statistical models, identical to Auto Regressive Integrated Moving Average (ARIMA), which has led to a surge in interest in AI-driven alternatives. When compared to ARIMA and other time-series models, LSTM networks, which aim to handle vanishing gradients in sequential data, offer significant advances and have been successful in financial time-series forecasting [1, 2]. Even if other designs are becoming more popular, a 2024 evaluation reveals that LSTM is still the go-to for modelling financial pricing data's long-term relationships [2-5]. Nevertheless, even though there were great improvements in the performance of each single model, the integration of LSTM with its temporal modeling capability, XGBoost with its ensemble learning, and DT with its interpretability is understudied. The current gap in the development of hybrid AI models encourages the current research.

XGBoost, a highly scalable gradient boosting system, has also demonstrated exceptional performance in credit risk and default prediction tasks. This work used credit card customer default studies, for example. The state-of-the-art findings show an accuracy of approximately 99% and ROC AUC values surpassing 0.99 [6]. According to other research on Small and Medium-sized Enterprises (SME) supply-chain finance, optimized XGBoost outperforms other models in credit-risk categorization [7]. Financial markets are volatile and nonlinear, and older models such as ARIMA are not designed to deal well with those aspects. In this way, there has been an increased necessity for stronger, more interpretable, and reliable risk prediction models, especially financial forecasting data with AI.

Although LSTM, XGBoost, and DTs are all strong models in their own right, there has been little research on hybrid architectures that include all three. But joining LSTM's temporal modelling with XGBoost's ensemble learning and DTs' interpretability and rule-based segmentation bodes well. In investment forecasting tasks, recent arXiv research on 2025 highlights the potential benefits of hybrid sequential plus tree-based models, demonstrating 10 to 15% accuracy boosts over individual models [8-11].

The financial sector urgently needs transparent and high-performing AI. Critical in risk-sensitive fields, explainability and regulatory compliance have been largely ignored in favor of performance in much previous research. Credit scoring frameworks based on explainable XGBoost have been introduced in recent work [12, 13]. These frameworks combine model performance with interpretability by making use of tools like SHAP and clustering techniques.

To improve financial risk forecasting, this research presents a new hybrid AI architecture that integrates XGBoost, DTs, and LSTM networks. This work is innovative because three proven approaches (DT, LSTM, and XGBoost) are incorporated within the same framework and overcome the shortcomings of prior research, which combines two of the techniques or does not involve interpretability alongside prediction accuracy.

This contribution is as follows:

- A modular pipeline that uses DTs, LSTM, and XGBoost is created.
- The hybrid model is compared to standalone LSTM, Random Forest (RF), and Logistic Regression (LR) on real-world datasets like UCI credit default data, and it is also benchmarked against other hybrid models.
- Ablation research is conducted to measure the relative importance of each part to the whole.
- Explainable AI technologies are integrated that are in line with rising regulatory demands in the financial sector, with an emphasis on interpretability and transparency.

The organization of this paper is as follows: Section 2 reviews related research on AI applications in financial risk forecasting. Section 3 presents the proposed hybrid architecture and algorithms. Section 4 details datasets, preprocessing, metrics, and experimental setup. Section 5 concludes the key findings of this research and provides suggestions for future work.

2. Related Work

2.1. LSTM for Financial Time-Series Forecasting

LSTM's ability to capture long-range relationships while minimizing gradient vanishing difficulties has led to its consistently excellent performance in financial time-series forecasting. In their study, Yu et al. [14] examined LSTM models using S&P 500 returns and twelve financial variables. They found that, despite worries about overfitting, there were constant decreases in validation loss. This suggests that future research should focus on improving ensemble models.

In an early demonstration of fundamental superiority over standard statistical models, Siami Namini et al. [15] revealed that LSTM had error rates that were 84-67% lower than ARIMA on a variety of financial data. Recent research has revealed that LSTM can effectively deal with temporal data, but its combination with other algorithms, such as XGBoost, leads to enhanced accuracy and eliminates the problems of overfitting and errors of generalization.

2.2. XGBoost in Credit Risk Prediction

When it comes to credit scoring, XGBoost is still the ruler. The use of XGBoost in personal loan default prediction was carried out by Chen et al. [16], who achieved an ROC-AUC of around 0.71 and outperformed other classifiers in terms of usability measures.

Using multi-source heterogeneous data for SME finance, Yuwen Zeng et al. [17] combined Recursive Feature Elimination (RFE) with XGBoost, increasing F1 from around 0.882 to 0.915, which is almost 4.8% better than the RF controls. Hyperparameter optimization was shown to be beneficial in supply chain finance risk assessment employing Bayesian-optimized XGBoost, which improved accuracy by more than 91% and had higher AUC and F1 values than untuned models [7]. Nonetheless, the prevailing literature supports sequential learning only in a limited manner since it considers only non-dynamic statistics, e.g., the outstanding fixed amount of the financial risk.

2.3. Hybrid Architectures Combining LSTM and Tree-Based Models

Some new hybrid models are emerging that combine deep sequential networks with boosting or tree-based components. Chang Yu et al. [6] suggested a hybrid architecture that combines LSTM, LightGBM, and CatBoost to enhance stock prediction accuracy by 10-15% compared to individual approaches. To increase the generality and accuracy in consumer lending situations, Zhu et al. [18] constructed ensemble credit-default models employing LightGBM, XGBoost, and local ensemble approaches.

Sun et al. [12] reached their goals of being accurate and transparent by using an explainable credit scoring method that combines XGBoost with K-means explanations and SHAP-style interpretability [12].

2.4. Gaps in the Literature

The literature is still missing fully integrated hybrid architectures for financial risk forecasting that combine rule-based segmentation (DT), temporal modelling (LSTM), and boosted predictive learning (XGBoost) in a unified pipeline, despite significant developments. Few existing studies make use of all three modalities simultaneously in a single framework; most concentrate on pairings of models, such as LSTM + XGBoost or XGBoost + LightGBM.

Not many studies have combined modular interpretability (made possible by a separate DT component) with temporal sequence modelling in a hybrid architecture. This is particularly true when it comes to studies that have evaluated the architecture on real-world financial datasets from the credit, loan, and investment domains, as well as those that have added explicability to specific XGBoost frameworks. The summary of these existing works is tabulated in Table 1.

Table 1. Summary of key studies

Model / Study	Focus	Key Findings
Chang Yu et al. [6]	LSTM + LightGBM/CatBoost hybrid	Improves forecasting accuracy by 10–15 %
Sun et al. [12]	Explainability in XGBoost	High accuracy with interpretability via SHAP/K-means
Yu et al. [14]	LSTM on S&P 500 returns	Reduced validation error; highlights overfitting risk
Siami-Namini et al. [15]	LSTM vs ARIMA	LSTM lowers error by ~85 %
Chen et al. [16]	XGBoost on credit default	ROC-AUC ~0.71; best among tested models
Yuwen Zeng et al. [17]	RFE + XGBoost for SME risk	F1 improves from 0.882 to 0.915
Zhu et al. [18]	BO-XGBoost in supply chain finance	Accuracy ~91 %; highest AUC via tuning

The proposed study fills important gaps by expanding on these results and proposing a three-part hybrid design that:

- Uses Chosen Option Trees for Rule-Based Data Segmentation that are Interpretable,
- Uses LSTM to create detailed temporal modelling of a financial time series,
- XGBoost enhances the predictive capabilities of the designed features,
- Transparency is ensured by embedding explicable AI approaches, such as SHAP values and decision rules,
- This comparison is made against standards derived from real-world datasets, such as UCI credit default statistics and LendingClub loan records.

3. Proposed Methodology

To enhance the accuracy, interpretability, and resilience of financial risk forecasting, the proposed study presents a hybrid predictive AI model. DT, LSTM networks, and Extreme Gradient Boosting (XGBoost) are three original components that function together in this architecture. DTs offer explainable segmentation, LSTM networks capture the temporal relationships of financial sequences, and XGBoost enhances learning from both sequential and static characteristics. It is required to integrate interpretable machine learning, deep sequential modelling, and ensemble-based prediction into a single pipeline to maximize their respective capabilities.

3.1. Overview of the Architecture

The architecture comprises DTs for feature segmentation, an LSTM for temporal modelling, and an XGBoost for final classification. The first step in processing raw input data is to standardize formats, eliminate missing values, and convert categorical features to numerical representations. This data includes both customer-specific information and temporal repayment records. Two parallel pathways are established for these cleaned inputs.

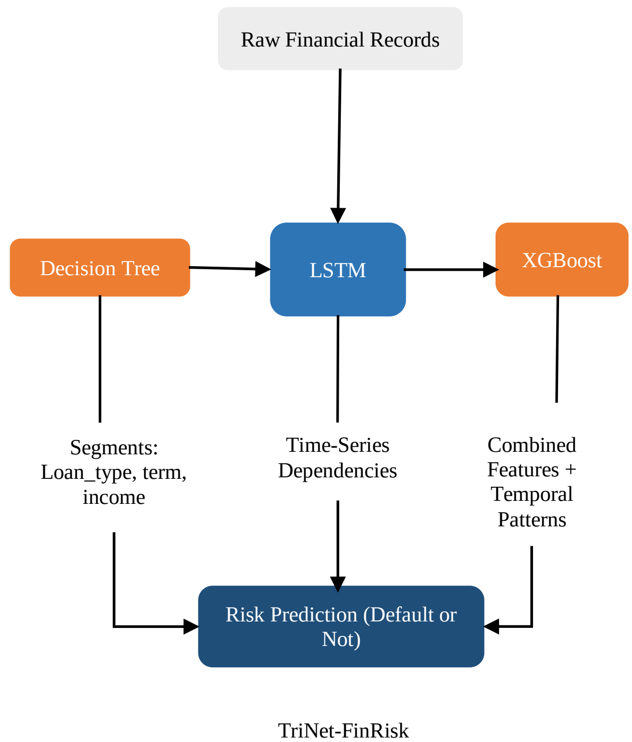

Fig. 1 A hybrid tri-stream architecture for financial risk forecasting

First, a shallow DT is used to analyze financial data like yearly salary, credit grade, loan duration, and job tenure to arrange consumers into understandable categories. To uncover latent temporal characteristics, the second route inputs time-dependent information into an LSTM network. These records may include installment payments, loan installment dates, and balance progression. Finally, the XGBoost classifier is fed the combined outputs of the two routes together with the initial static characteristics to make a prediction about financial risk, such as the chances of loan default. The flow of the proposed research is pictorially illustrated in Figure 1.

3.2. Data Preprocessing

The hybrid architecture's performance is highly dependent on efficient preprocessing. This study utilizes the LendingClub dataset, which includes more than 2 million records, including loan purpose, income, length, interest rate, debt-to-income ratio, credit history, and job history, among other characteristics. Among the many data cleaning activities carried out is the elimination of records when vital fields, such as income or loan status, are either blank or include incorrect information. Based on the cardinality, categorical data like loan grade and job title are encoded using either label encoding or one-hot encoding. Loan amounts and yearly income are two continuous variables that are normalized utilizing min-max scaling. To facilitate LSTM training, serial formats of set length are used for temporal information, such as loan issuance dates and installment payment timeframes. To build a binary target variable, the loan status must first be mapped to a number between 1 and 0. For defaulted loans, the number is 1 (for example, "Charged Off"), and for fully paid loans, the number is 0 [19].

3.3. Decision Tree-Based Feature Segmentation

As an initial step in the hybrid pipeline, a lightweight DT is trained to classify borrowers into separate groups according to risk-relevant characteristics. DTs' inherent interpretability and rule-based structure make them ideal for this particular purpose. The tree depth is limited to 3 in this implementation to keep it readable and prevent overfitting. Loan terms, income levels, and credit score bands are some of the characteristics used for segmentation. Every instance is given a segment name based on the borrower category represented by each leaf node of the tree. When the input set is prepared for downstream learning, this label is added as an extra category feature. By providing explicable boundaries between risk classes and making the rationale of borrower classification understandable to human analysts, this module adds to the model's transparency.

3.4. LSTM for Temporal Feature Extraction

An LSTM network is used to model the temporal component of financial risk. One kind of Recurrent Neural Network (RNN) that outclasses at modelling sequential data is the long-term dependent type LSTM. The LSTM network model is trained utilizing sequences of recurring monthly installment payments and total amounts owed. A multivariate time series with one timestep every month frames each borrower's history. The time-series sequences are fed into the network via an input layer, which is then followed by a 128-hidden-unit LSTM layer and a dropout layer that prevents overfitting. To capture the temporal dynamics of the repayment behaviour, the LSTM produces a fixed-dimensional vector as an output. The hidden representation, referred to as h_t, records patterns that are predictive of financial risk, such as late payments, early closures, or lost payments. The mathematical foundation of the LSTM unit is as follows. At each timestep t, the network computes:

$$i_t = \sigma\,(w_i \,.\, [h_{t-1}, x_t] + b_i \text{ (Input Gate)} \quad (1)$$

$$f_t = \sigma\,(W_f \,.\, [h_{t-1}, x_t] + b_f) \text{ (Forget Gate)} \quad (2)$$

$$\widetilde{c_t} = tanh(W_c \,.\, [h_{t-1}, x_t] + b_c) \text{ (Candidate Memory} \quad (3)$$

$$c_t = f_t * c_{t-1} + i_t * \widetilde{c_t} \text{ (Cell State)} \quad (4)$$

$$h_t = o_t * tanh(c_t) \text{ (Hidden State)} \quad (5)$$

$$o_t = \sigma\,(W_o \,.\, [h_{t-1}, x_t] + b_o) \text{ (Output Gate)} \quad (6)$$

Table 2. Notations used in LSTM equations

Symbol / Term	Description
x_t	Input vector at time step t (e.g., installment amount, balance)
h_{t-1}	Hidden state vector from previous time step $t-1$
h_t	Current hidden state at time step t and LSTM output at t
c_t	Cell state at time step t, maintaining long-term memory
c_{t-1}	Cell state from the previous time step
f_t	Forget gate activation, which decides what to discard from the cell memory.
i_t	Input gate activation, which chooses the new data to store
o_t	Output gate activation that controls which part of the memory goes to h_t
$\tilde{c}_t$	Candidate cell state — potential keeps information about the cell state.
$\sigma(\cdot)$	Sigmoid activation function
$tanh(\cdot)$	Hyperbolic tangent function, outputs between -1 and 1
W_i, W_f, W_c, W_o	Input, forget, candidate, and output gates weight matrices
b_i, b_f, b_c, b_o	Input, forget, candidate, and output gates bias vectors
$*$	Element-wise (Hadamard) product
$[h_{t-1}, x_t]$	Concatenation of the hidden state along with the input vector

The descriptions of all notations employed in LSTM are given in Table 2. The f_t decides the details about the c_{t-1} to keep. i_t and $\tilde{c}_t$ work together to update the cell memory. The o_t identifies how much of the cell state is exposed to the next layer or as output. This mathematical formalism allows the LSTM to selectively retain or forget patterns over lengthy periods. Efficiently modelling borrower behaviour over time relies on these calculations, which enable the network to selectively store or forget data between time steps.

3.5. XGBoost-Based Final Prediction

Predicting if a borrower will default is the last classification job that is carried out using an XGBoost classifier. XGBoost's abilities to handle heterogeneous features, resist overfitting, and take advantage of complicated feature relationships led to its selection. The XGBoost input includes initial static characteristics (such as income, term, or grade); the segment label obtained from the DT module; and the temporal embedding vector created from LSTM. The XGBoost model is given these characteristics as a single input vector.

Grid search is used to tweak the hyperparameters, and a binary logistic objective function is used to train the classifier. It is concluded that 200 estimators, a high tree depth of 5, and a 0.05 learning rate are the best parameters to use. The dataset's underlying class imbalance is reduced by adjusting scale_pos_weight in accordance with the default rate ratio.

3.6. Explainability and Interpretation

By employing shapley Additive Explanations (SHAP) for feature attribution, the model enhances interpretability and builds confidence in the AI system [20]. To measure how much of an effect each input characteristic has on the final prediction, its SHAP value is calculated.

This enables analysts to determine the most influential factors for each prediction, whether they are static, temporal, or segment-based. For each group of borrowers, the DT module provides explanations based on rules; to show how time affects things, the LSTM can optionally use attention processes or saliency maps.

4. Experimental Setup

The proposed TriNet-FinRisk model was evaluated using a detailed testing process that involved comparing it to basic models, setting up the model, preparing the dataset, and performing preprocessing tasks. This section gives the experimental pipeline step-by-step that created the proposed hybrid architecture for training, validation, and benchmarking.

4.1. Dataset Description

The experiments use the LendingClub Loan Dataset, a publicly accessible financial dataset that has gained recognition in credit risk modelling research. This dataset includes comprehensive information on borrowers and loans, including records at the loan level issued between 2007 and 2020. To ensure computational feasibility and class balance, a stratified sample of 50,000 records was used for this investigation.

The original class distribution is preserved in the chosen subset, with the remaining loans designated as Fully Paid and about 18–20% of those classified as Defaulted (Charged Off). A binary encoding of Fully Paid as 0 and Charged Off as 1 is used for the dependent variable loan status to facilitate categorization. The dataset comprises a variety of categorical, numerical, and temporal variables relevant to financial risk prediction. Table 3 provides an overview of important characteristics.

Table 3. Lending club dataset summary

Feature Type	Attributes	Examples
Numerical	Loan amount, interest rate, annual income	loan_amnt, int_rate, annual_inc
Categorical	Loan term, purpose, grade, home ownership	term, grade, purpose
Temporal	Payment history, issue date, installment trend	issue_d, installment, payment_time
Target	Loan status (binary classification)	loan_status → 0 (Paid), 1 (Default)

This dataset was selected for testing both static and sequential learning models because of the variety of borrower demographics, financial backgrounds, and temporal repayment habits.

4.2. Data Preprocessing

In order to create the data for time-series investigation, a multi-step procedure was used to fill in missing values, normalize numerical features, encode categorical variables, and build sequences. To begin, the median value was used to impute missing values in crucial numerical variables, including annual_inc, emp_length, and mort_acc. The dataset was cleansed of records that included inconsistent time-series entries or null target labels. Label encoding was used to encode low-cardinality categorical variables like term and grade, but fields with high cardinality, such as emp_title, were removed owing to sparsity and noise.

To guarentee that the LSTM model, which is sensitive to feature magnitudes, could work with continuous features like loan amount, DTI, interest rate, annual income, and min-max normalization was used to scale them. Information pertaining to installments was translated into sequences of up to 36 monthly time steps for the purpose of the temporal modelling component. To ensure that the time-series component of the model was accurate, only the track of borrowers who had at least six months of payment history was considered. By using stratified sampling, the final dataset was split into three sets 70% for training, 15% for validation 15%, 15% for test 15%.

4.3. Training Configuration

The DT, LSTM, and XGBoost components of the TriNet-FinRisk architecture were trained independently prior to their incorporation into the whole pipeline. Monthly payment habits were modelled with the LSTM module. A thick layer with 64 neurons and ReLU activation follows a single LSTM layer with 128 hidden units. To reduce the likelihood of overfitting, a 0.3-rate dropout layer was used. The Adam optimizer was used to create the model with a learning rate of

0.001. It was trained for 50 epochs using binary cross-entropy loss, and early stopping was based on validation AUC. It was configured to use 64 batches.

The borrowers were categorized according to loan length, grade, and income using a shallow DT classifier that was built for segmentation. To keep things simple and prevent overfitting, we kept the tree depth to 3. The XGBoost classifier, which makes the final risk prediction, received three types of information: fixed details about the borrowers, time-related features from the LSTM, and the category label created by the DT. The ideal settings are shown in Table 4, which were obtained by tweaking the hyperparameter using grid search.

Table 4. XGBoost training parameters

Parameter	Value
Learning rate	0.05
Number of estimators	200
Max tree depth	5
Objective	Binary: logistic
Class weight	scale_pos_weight = 4.3

The models were trained on a system with an NVIDIA RTX 3060 GPU and 16 GB RAM, and they were developed using Python 3.10, TensorFlow 2.14, XGBoost 1.7.6, and Scikit-learn 1.3.2.

4.4. Baseline Models for Comparison

Several baseline models were developed to benchmark the proposed hybrid model performance. This includes LR, which is a linear classifier with L2 regularization. RF, an ensemble of 100 DTs with a maximum depth of 10. Standalone LSTM, where the same architecture as used in the hybrid model, is trained independently. Flat XGBoost – XGBoost is trained only on static features without temporal or segment data. These models show how a boosted ensemble using segmentation and temporal learning may improve performance incrementally.

4.5. Evaluation Metrics

A set of classification measures was employed to evaluate the model performance thoroughly:

- Accuracy: The proportion of instances when predictions were accurate relative to the total.
- Precision: The ratio of precise default forecasts to all anticipated defaults is measured.
- Recall (Sensitivity): The proportion of real defaults that were accurately predictable.
- F1-Score: It is helpful for data with imbalances since it is the harmonic mean of recall and precision.
- ROC-AUC: Discriminatory capacity is measured by the area under the receiver operating characteristic curve (ROC-AUC).
- Cost Effectiveness: A domain-specific metric that measures the reduction of false positives per 100 loan approvals, which is used as a proxy for financial savings.

The proposed model performance was compared to standard benchmarks using 10-fold stratified cross-validation, and the importance of these changes was checked with McNemar's test.

4.6. Results and Discussion

Here is a detailed look at how well the TriNet-FinRisk model predicts outcomes, saves costs, and what each part contributes, compared to standard industry measures. The findings demonstrate that financial risk forecasting is significantly enhanced by integrating sequential modelling, tree-based segmentation, and ensemble boosting.

4.6.1. Overall Performance Comparison

Table 5 shows the average performance of the existing works and the proposed model after 10 rounds of stratified cross-validation. Compared to all baseline models, the TriNet-FinRisk model consistently performs better on all key measures. When compared to all baselines, the TriNet-FinRisk model always performs better on all important criteria.

Table 5 shows that TriNet-FinRisk has the greatest F1-score, ROC-AUC, and maximum accuracy of 91.43%. The model demonstrates better cost reductions from fewer false approvals and better defaulter classification.

Table 5. Proposed model performance comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)	Cost Effectiveness*
LR	78.62	65.30	58.42	61.66	76.41	1.00
RF	82.15	71.84	68.11	69.93	81.73	1.35
Standalone LSTM	84.68	74.29	71.42	72.83	85.21	1.48
Flat XGBoost	86.12	76.33	73.88	75.08	86.74	1.55
TriNet-FinRisk	**91.43**	**83.64**	**81.75**	**82.68**	**91.02**	**1.83**

*Cost Effectiveness: Reduction in false positives per 100 approvals (normalized baseline = 1.00)

4.6.2. ROC Curve Analysis

The ROC curves for all the models are illustrated in Figure 2. The TriNet-FinRisk curve is quite convex, suggesting that sensitivity and specificity are well-balanced. The Area Under the Curve (AUC), which reaches 91%, confirms its strength in differentiating between non-defaulting and defaulting debtors.

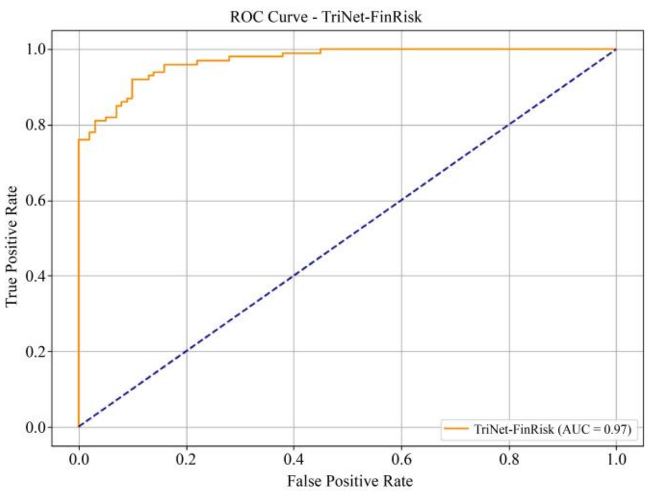

Fig. 2 ROC curves of trinet-finrisk and baseline models

4.6.3. Ablation Study

An ablation study was carried out by methodically deleting one component from the TriNet FinRisk system at a time to better assess the contribution of each model component. The outcomes are given in Table 6.

Table 6. Ablation study results

Configuration	Accuracy (%)	F1-Score (%)	ROC-AUC (%)
Full TriNet-FinRisk	**91.43**	**82.68**	**91.02**
Without LSTM (no sequential learning)	87.02	77.48	87.61
Without DT segmentation	88.14	78.92	88.32
Without XGBoost (LSTM + DT + MLP)	85.81	75.31	85.93

Based on the results, it is clear that all three parts are essential for the model to work well. The significance of temporal pattern learning is shown by the performance decline (~4.4% fall in AUC) that occurs when the LSTM is removed. Similarly, removing the DT segmentation reduces interpretability and results in an AUC drop of approximately 2.7%. The biggest hit to performance comes from swapping out XGBoost for a basic multi-layer perceptron.

4.6.4. Discussion and Insights

The findings show that TriNet-FinRisk successfully integrates the advantages of many modeling paradigms to represent the complex nature of financial risk. Temporal learning via LSTM detects critical patterns that static models overlook, such as repayment delays, missing payments, and payback consistency. Segment-based features from DTs improve interpretability by letting the model learn borrower-specific rules (for example, low-income borrowers with short loan periods tend to default at greater rates).

Boosting with XGBoost efficiently and effectively deals with non-linearities and interactions between features. One important measure in financial applications is cost-effectiveness, and the hybrid method also excels in this area. TriNet-FinRisk is well-suited for use in high-stakes settings, such as retail credit scoring, microfinance, or investment screening, due to its ability to drastically decrease the number of false approvals. The SHAP value representation of the proposed TriNet-FinRisk is depicted in Figure 3.

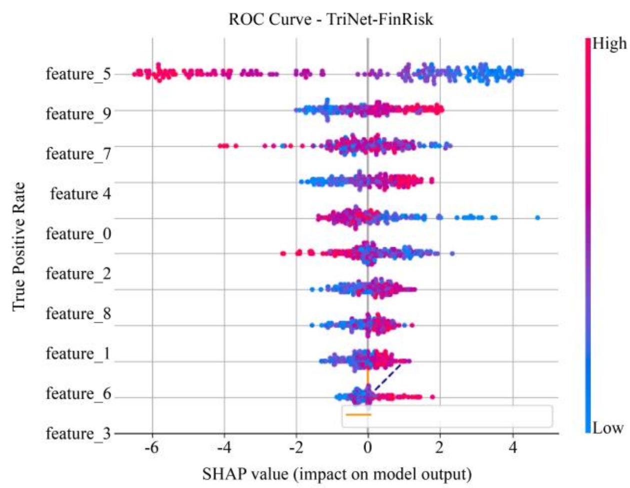

Fig. 3 SHAP value (impact of model output)

Compliance in the financial services industry is of the utmost importance, and the explainability modules (such as SHAP values and decision rules) guarantee that the model outputs are visible and audit-friendly. The proposed model is also superior to existing techniques as it includes both learning of the temporal sequence and explainable AI to offer prediction capabilities and transparency, unlike models that solely rely on the prediction capability with a lack of explainability.

5. Conclusion and Future Work

This research presented TriNet-FinRisk, a hybrid artificial intelligence framework that combines DTs, LSTM networks, and XGBoost networks for financial risk forecasting. The model enhances predictive accuracy and interpretability in credit risk assessment by effectively combining rule-based segmentation, temporal sequence learning, and gradient-boosted decision modeling. Empirical findings on the LendingClub loan dataset show that the suggested architecture outperforms conventional machine learning architectures, standalone LSTM architectures, and flat learning architectures. With an AUC of 91.02% and an F1-score of 82.68%, TriNet-FinRisk outperformed all other assessment measures. The model's ability to be used in important financial decisions was shown through ROC analysis and SHAP-based feature interpretation, which highlighted its strength and clarity. The ablation research highlighted how important each part of the model is, especially how structured segmentation and temporal learning help improve its overall performance. By using Explainable AI (XAI) methodologies, decision-makers may have a better understanding of how features contribute, which in turn helps with automated credit assessment reliability and regulatory compliance. Although TriNet-FinRisk has shown effectiveness in static scenarios for forecasting credit risk, there are numerous intriguing avenues that might be explored in future research, including real-time risk prediction and federated learning integration.

Acknowledgement

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R793), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

[1] Nesrine Gafsi, "Machine Learning Approaches to Credit Risk: Comparative Evidence from Participation and Conventional Banks in the UK," *Journal of Risk and Financial Management*, vol. 18, no. 7, pp. 1-18, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[2] Foued Saâdaoui, and Hana Rabbouch, "Financial Forecasting Improvement with LSTM-ARFIMA Hybrid Models and Non-Gaussian Distributions," *Technological Forecasting and Social Change*, vol. 206, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Cheng Zhang, Nilam Nur Amir Sjarif, and Roslina Ibrahim, "Deep Learning Models for Price Forecasting of Financial Time Series: A Review of Recent Advancements: 2020–2022," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 14, no. 1, pp. 1-33, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Xuemin Huang et al., "A Hybrid ARIMA-LSTM-XGBoost Model with Linear Regression Stacking for Transformer Oil Temperature Prediction," *Energies*, vol. 18, no. 6, pp. 1-21, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[5] Dominik Stempień, and Robert Ślepaczuk, "Hybrid Models for Financial Forecasting: Combining Econometric, Machine Learning, and Deep Learning Models," *Arxiv Preprint*, pp. 1-30, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[6] Victor Chang et al., "Credit Risk Prediction Using Machine Learning and Deep Learning: A Study on Credit Card Customers," *Risks*, vol. 12, no. 11, pp. 1-33, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Chen Zhang, and Xinmiao Zhou, "Forecasting Credit Risk of SMEs in Supply Chain Finance Using Bayesian Optimization and XGBoost," *Mathematical Problems in Engineering*, vol. 2023, no. 1, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Chang Yu et al., "Gradient Boosting Decision Tree with LSTM for Investment Prediction," *Arxiv Preprint*, pp. 1-6, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[9] Zhuangwei Shi et al., "Attention-Based CNN-LSTM and XGBoost Hybrid Model for Stock Prediction," *Arxiv Preprint*, pp. 1-7, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Hassan Oukhouya et al., "Forecasting International Stock Market Trends: XGBoost, LSTM, LSTM-XGBoost, and Backtesting XGBoost Models," *Statistics, Optimization & Information Computing*, vol. 12, no. 1, pp. 200-209, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Hui Zhang, and Weihua Zhang, "Advancing Enterprise Risk Management with Deep Learning: A Predictive Approach Using the XGBoost-CNN-BiLSTM Model," *PloS One*, vol. 20, no. 4, pp. 1-23, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[12] Xinyu Sun, Jiayu Liu, and Yan Zhang, "Enhancing Credit Risk Prediction through an Ensemble of Explainable Model," *Journal of Systems Science and Systems Engineering*, pp. 1-22, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[13] Harsh Pathak Shreya, "Explainable Artificial Intelligence Credit Risk Assessment Using Machine Learning," *Arxiv Preprint*, pp. 1-15, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[14] Sijing Yu, "Advancing Stock Market Return Forecasting with LSTM Models and Financial Indicators," *Advances in Economics, Management and Political Sciences*, vol. 122, pp. 137-144, 2024. [CrossRef] [Publisher Link]

[15] Sima Siami-Namini, and Akbar Siami Namin, "Forecasting Economics and Financial Time Series: ARIMA vs. LSTM," *Arxiv Preprint*, pp. 1-19, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[16] Yi Chen, Youzhong Dong, and Wen Liu, "Prediction of Credit Default Based on the XGBoost Model," *Applied and Computational Engineering*, vol. 96, pp. 85-92, 2024. [CrossRef] [Publisher Link]

[17] Yuwen Zeng et al., "Integrated RFE-XGBoost Credit Risk Prediction for SMEs Using Multi-Source Heterogeneous Big Data," *Proceedings of the 3rd International Conference on Big Data Economy and Digital Management*, Ningbo, China, pp. 1-10, 2024. [CrossRef] [Publisher Link]

[18] Mengran Zhu et al., “Ensemble Methodology: Innovations in Credit Default Prediction Using LightGBM, XGBoost, and LocalEnsemble,” *2024 IEEE 4th International Conference on Electronic Technology, Communication and Information*, Changchun, China, pp. 421-426, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Hakan Pabuccu, and Adrian Barbu, “Feature Selection with Annealing for Forecasting Financial Time Series,” *Financial Innovation*, vol. 10, pp. 1-26, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Shiqi Yang et al., “Interpretable Credit Default Prediction with Ensemble Learning and SHAP,” *arXiv Preprint*, pp. 1-5, 2025. [CrossRef] [Google Scholar] [Publisher Link]