Journal of Organizational and End User Computing
Volume 36 • Issue 1

Application of AdaBound-Optimized XGBoost-LSTM Model for Consumer Credit Assessment in Banking Industries

Lijuan Fan, School of Teacher Education, Weifang University of Science and Technology, China
Changlin Wang, Shandong University of Aeronautics, China & University of British Columbia, Canada*
Zhonghua Lu, School of Logistics, Linyi University, China

ABSTRACT

Consumer credit assessment has always been a crucial concern in the financial industry. It involves evaluating an individual's credit history and their ability to repay loans, playing a pivotal role in the risk management and lending decisions made by credit institutions. In the present landscape, traditional credit assessment methods confront various shortcomings. Firstly, they typically only consider static features and are unable to capture the dynamic changes in an individual's credit profile over time. Secondly, traditional methods struggle with processing complex time series data, failing to fully exploit the importance of time-related information. To address these challenges, we propose an innovative solution – the XGBoost-LSTM model optimized with the AdaBound algorithm. This hybrid model combines two powerful machine learning techniques, XGBoost and LSTM, to leverage both static and dynamic features effectively.

KEYWORDS

AdaBound Algorithm, Consumer Credit, Credit Assessment, Time Series Data, XGBoost-LSTM Model

With the increasing growth of China's national economy and the continuous improvement of the people's consumption level, our country's consumer finance market has ushered in vigorous development. Consumer finance refers to a credit method that provides funds to applicants through commercial banks, consumer finance companies, internet platforms, and other institutions to meet the applicant's consumption needs (Bannier et al., 2022). With the state's strong support for innovative credit products of financial institutions, various institutions have actively carried out personal consumer credit business and launched a variety of personal consumer credit forms. Conducting credit assessment on the rapidly growing large number of users is the core technical work of various financial institutions, and it is also a key measure to control risks. Credit assessment is one of the crucial tasks in banking and finance, which involves assessing the credit risk of an individual or entity to determine whether a loan, credit card application, or other financial transaction should be

DOI: 10.4018/JOEUC.343256 *Corresponding Author

approved (Kozodoi et al., 2022). Credit evaluation is not only related to the risk management of financial institutions but also directly affects individuals' financial life and economic participation ability (Nana et al., 2022). In the banking industry, personal credit assessment is one of the core factors in bank decision-making because it is directly related to the borrower's credit reliability, which determines the granting of loans, the setting of interest rates, and the determination of credit lines (Q. Li, 2023). However, personal credit assessment faces a series of challenges and problems. First, with the continuous changes in the financial market and the continuous innovation of financial products, traditional credit assessment models may not perform well in dealing with complex and changeable financial environments (Luo & Zhang, 2022). Second, traditional credit assessment mainly relies on static historical data and rules, which makes it difficult to capture the dynamic changes in personal credit risk (Guan et al., 2023). Third, in the context of huge amounts of data, traditional methods may be inefficient when processing large-scale data, and it is difficult to handle nonlinear and high-dimensional features. To address these challenges, researchers and financial institutions have gradually turned to advanced technologies such as deep learning and ensemble learning. Deep learning models, such as neural networks and recurrent neural networks (RNNs), are able to learn complex feature representations from large-scale data, thereby improving the accuracy of credit assessment. Integrated learning algorithms, such as XGBoost (extreme gradient boosting) and LightGBM, can combine multiple basic models to improve the stability and generalization capabilities of the model (Alarfaj et al., 2022). These advanced technologies have made significant progress in the field of consumer credit assessment, providing powerful tools to improve the accuracy and efficiency of credit assessment. This article aims to explore how to apply deep learning and integrated learning technologies, especially the XGBoost-LSTM model and AdaBound algorithm, to optimize personal credit assessment models and address current challenges. In the following sections, we will introduce the principles, application methods, and experimental results of these technologies in detail to demonstrate their potential application value in the banking industry.

Research on personal credit assessment technology boasts a longstanding history, initially flourishing primarily in capitalist countries with more developed economies. However, as our nation has witnessed rapid advancements in science, technology, and theoretical research in recent years, China has embarked on decades of progress in this domain. A recent investigation harnessed a deep learning model, a convolutional neural network (CNN) specifically (Illanko et al., 2022). These deep learning models excel in credit assessment due to their innate ability to automatically extract crucial credit-related features, thus augmenting the model's overall performance. Nevertheless, deep learning models still grapple with certain challenges, one of which is the issue of sample imbalance–the unequal representation of good and bad credit customers. Additionally, deep learning models are often deemed "black box" models, with limited interpretability, rendering it challenging to elucidate the rationale behind the model's predictions. Another study adopted the BERT (bidirectional encoder representations from transformers) model for credit assessment (Kriebel & Stitz, 2022; Ye et al., 2023). BERT, a pre-trained natural language processing model, boasts remarkable performance credentials. In this research endeavor, scholars leveraged BERT for credit assessment tasks, with the intent of more effectively capturing customers' credit-related information. However, it is worth noting that the BERT model is relatively substantial in size and demands extensive computational resources and time, potentially impeding its efficiency in practical applications. Recent investigations have witnessed the integration of graph neural networks (GNNs) into credit assessment models (Feng et al., 2022). GNNs are particularly suited for handling relational data and can adeptly account for correlations between customers. This approach shines when processing social network data and interactions among customers in credit assessment. However, these models encounter heightened computational complexity when grappling with large-scale data, and there remains room for performance enhancement when dealing with non-associated data. Furthermore, ensemble models, such as random forests (RFs) and gradient boosting trees, have garnered research attention (Y. Li, 2021; Mushava & Murray, 2024). These ensemble models excel in nonlinear modeling and exhibit prowess in capturing intricate relationships within credit assessment data. Nevertheless, their computational efficiency on large-scale data sets may warrant improvements, especially in credit assessment scenarios demanding real-time decision making. In summation, the aforementioned four seminal works epitomize recent research trends, each exploring distinct models and methodologies for credit assessment. However, these endeavors harbor certain limitations, encompassing aspects like an insufficient consideration of specialized financial sector issues, elevated computational complexity, or model instability. The principal objective of this paper is to bridge these research gaps and introduce a novel method that amalgamates deep learning and ensemble learning, thereby elevating both the accuracy and interpretability of credit assessment (Zhong & Zhao, 2024).

Addressing the shortcomings observed in prior research endeavors, we introduce a novel credit evaluation model: the XGBoost-LSTM (long short-term memory) model optimized with the AdaBound algorithm. The underlying concept of this model revolves around amalgamating two distinct methodologies, namely XGBoost and LSTM, to leverage the robust integrated learning capabilities of XGBoost alongside the sequence modeling prowess of LSTM. This integration is envisioned to adeptly tackle the intricacies inherent in credit assessment tasks within the financial domain, including the processing of time series data, capturing nonlinear relationships, and addressing data imbalance concerns. Additionally, we employ the AdaBound algorithm as the optimizer for our model. Designed specifically for deep learning models, the AdaBound algorithm aims to mitigate some of the limitations associated with traditional optimization algorithms. Its adaptive learning rate adjustment feature accelerates model convergence and enhances training stability. By integrating the AdaBound algorithm with the XGBoost-LSTM model, we anticipate a notable enhancement in model performance, rendering it more adept for credit evaluation tasks within the financial realm. The significance of our research lies in its capacity to enhance the accuracy and interpretability of personal credit assessment. With the escalating demand for model interpretability in finance, our approach aligns with this necessity, furnishing a more transparent foundation for decision-making processes. Furthermore, the amalgamation of XGBoost, LSTM, and AdaBound is poised to furnish financial institutions with more robust risk management tools, enabling them to gain deeper insights into and evaluate customers' credit profiles, mitigate risks, and augment the quality of loan decisions. Consequently, our model holds promising advantages, ushering in novel insights and breakthroughs in the realm of personal credit assessment.

In summary, our contributions can be encapsulated as follows:

(1) We propose an XGBoost-LSTM model optimized based on the AdaBound algorithm, which successfully combines ensemble learning and deep learning to address challenges in the field of consumer credit evaluation. The model's design empowers us to effectively address the intricacies present in financial data, such as time series data, nonlinear relationships, and data imbalances. This, in turn, enhances the accuracy and resilience of credit assessment.
(2) Our research underscores the practical utility of the AdaBound algorithm within deep learning models. By introducing AdaBound as the optimizer, we improve the training stability and performance convergence speed of the model, providing more reliable support for the practical application of the model. This is of great significance for credit assessment in the financial field because it can accelerate the deployment and application of models and reduce uncertainty in practical applications.
(3) Our research brings new ideas and methods to the field of personal credit assessment. By combining XGBoost, LSTM and AdaBound, we propose a new credit assessment framework aimed at improving model performance and interpretability. This provides the banking industry and financial institutions with more advanced and reliable credit assessment tools, helping to better understand and manage customers' credit risks, thus providing strong support for the development of financial businesses.

RELATED WORK

Research on Time Series Analysis in Credit Evaluation

Research on time series analysis in credit evaluation is committed to making full use of the time information in personal credit history data to more accurately understand credit trends, predict future credit performance and identify credit default risks (Talaat et al., 2023; Yuan et al., 2022). Researchers use time series analysis methods, such as trend modeling and feature extraction, to capture time series characteristics in individual credit histories. This analysis helps financial institutions better understand customers' credit behavior, develop more precise credit strategies, and improve the efficiency of risk management. At the same time, time series analysis also provides an important data mining tool for the field of credit evaluation, which is expected to improve the performance and prediction capabilities of credit evaluation models (Zhao & Chen, 2022; Zeng & Zhong, 2022).

Machine Learning and Integrated Learning Methods Applied to Credit Assessment

Machine learning and comprehensive learning methods applied to credit assessment represent a multi-domain technology collection, and their application in the financial field provides financial institutions with powerful and diverse credit assessment tools. Deep learning methods, such as neural networks and LSTM, have gained prominence in credit assessment (Ba et al., 2022). These methods are not only able to handle complex feature extraction but also capture nonlinear relationships and process time series data, thereby significantly improving the accuracy of credit assessment models (Shi et al., 2022). At the same time, comprehensive learning methods also play a key role, using ensemble learning strategies to combine the predictions of multiple base models (Singh et al., 2022). This method not only helps reduce the variance of the model and improves the generalization performance of the model but also performs well in the face of challenges such as data imbalance. By combining deep learning and comprehensive learning methods, financial institutions can better manage credit risks and provide more accurate, explainable, and comprehensive credit assessments, further improving the efficiency and decision-making quality of financial operations (Lenka et al., 2022). The integration of these methods brings new possibilities to the field of credit assessment and provides a solid foundation for future financial innovation. Deep learning methods give the model the ability to handle complex data and relationships, while comprehensive learning methods enhance the stability and reliability of the model. The combination of the two will further promote technological innovation in the financial field and provide financial institutions with more tools and methods to better respond to the changing credit market and risks. The development of this comprehensive approach will also help improve the stability of the financial system and provide borrowers and lenders with fairer, more transparent, and more sustainable credit assessment services.

Big Data Driven Credit Assessment

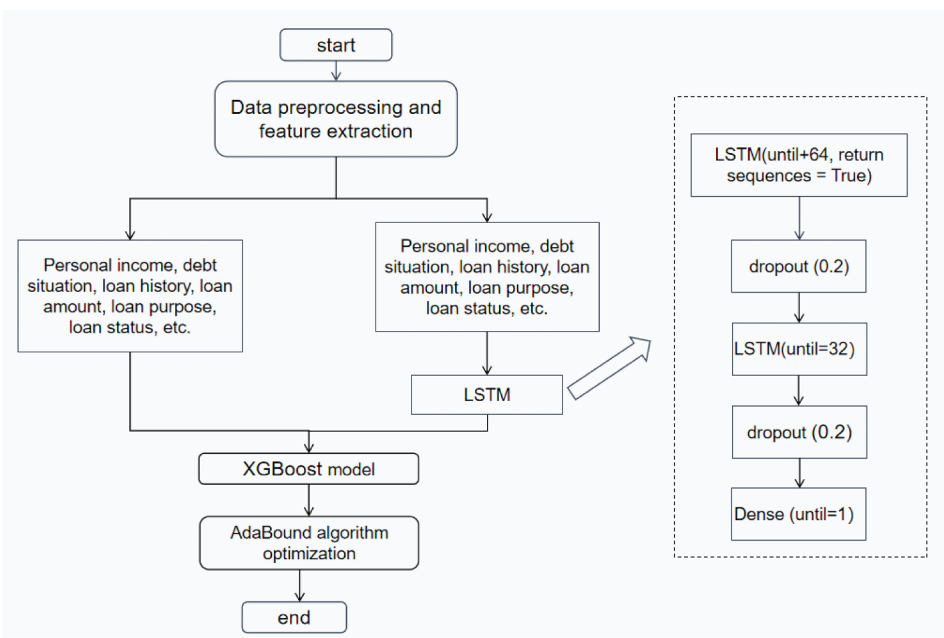
Big data driven credit assessment is a method that utilizes huge and diverse data resources to improve the accuracy and efficiency of personal credit assessment through data collection, preprocessing, feature engineering, and advanced machine learning technology (Roeder et al., 2022). This method not only includes financial historical data but also covers customers' consumption behavior, social media activities, and other aspects to gain a more comprehensive understanding of customer credit risks (Lin, 2022; Ye & Zhao, 2023). The real-time nature of big data allows financial institutions to respond more quickly to customers' credit needs, while also strengthening fraud detection capabilities and improving the quality of financial decisions and customer experience. By leveraging big data resources, big data driven credit assessment represents an important step toward smarter and more comprehensive credit risk assessment methods in the financial sector.

METHOD

Overview of Our Network

Our model is based on the XGBoost-LSTM model optimized by the AdaBound algorithm, which is designed to improve the accuracy and efficiency of personal credit assessment. The model integrates two key components, XGBoost and LSTM, to achieve a more comprehensive credit assessment. XGBoost is used to process structured data and can capture nonlinear relationships between features, while LSTM is used to process time series data and has sequence modeling capabilities. This integration allows us to assess a client's credit risk more comprehensively and provide a more accurate credit assessment. The model building process covers several key steps. First, we conduct data preparation to collect and prepare large-scale personal credit assessment data, including multi-faceted customer information. Next, we perform feature engineering to preprocess and extract features from the original data to ensure data quality. Then, we integrate the two models of XGBoost and LSTM to obtain more comprehensive feature representation and credit assessment capabilities. Finally, we use the AdaBound algorithm to optimize the model to improve the stability and performance of the model. This building process ensures the reliability and validity of our model in the field of personal credit assessment. In the application industry, this model represents a new generation of personal credit assessment methods, providing a more comprehensive credit assessment by comprehensively considering structured and time series data. This helps financial institutions improve credit risk management, customer service, and financial decision making, reduce credit risk, improve business efficiency, and provide customers with a better financial service experience. Therefore, the significance of models in the field of personal credit evaluation is to improve the decision-making capabilities of financial institutions, help them better understand customer credit risks, and create a more valuable financial ecosystem. Figure 1 shows the overall structure of the model.

Figure 1. Overall Architecture of the Model



XGBoost (Extreme Gradient Boosting) Model

XGBoost is an ensemble learning method based on the framework of gradient boosting trees. The core idea of this algorithm is to gradually improve the performance of the model by iteratively training a series of decision tree models, with each tree trying to correct the errors of the previous tree (J. Wang et al., 2022). The uniqueness of XGBoost is that it introduces regularization terms during the iteration process to control the complexity of the tree, thus avoiding the overfitting problem (Rao et al., 2023). Additionally, XGBoost uses an optimization technique called gradient boosting to minimize the loss function, thereby improving the accuracy and generalization performance of the model (K. Wang et al., 2022). In our model, XGBoost, as a component, contributes multiple values. First, XGBoost's superiority in processing structured data enables the model to capture nonlinear relationships between features, improving the accuracy of credit assessment. Second, XGBoost has the ability of automatic feature selection, which helps identify the most important features, reduces the data dimension, and improves the efficiency of the model. Most importantly, the integrated learning nature of XGBoost enables the model to better cope with the complexity in credit assessment and provide more reliable credit decisions.

Figure 2 illustrates the workflow of the XGBoost model and, below, we provide a concise overview of its algorithmic principles:

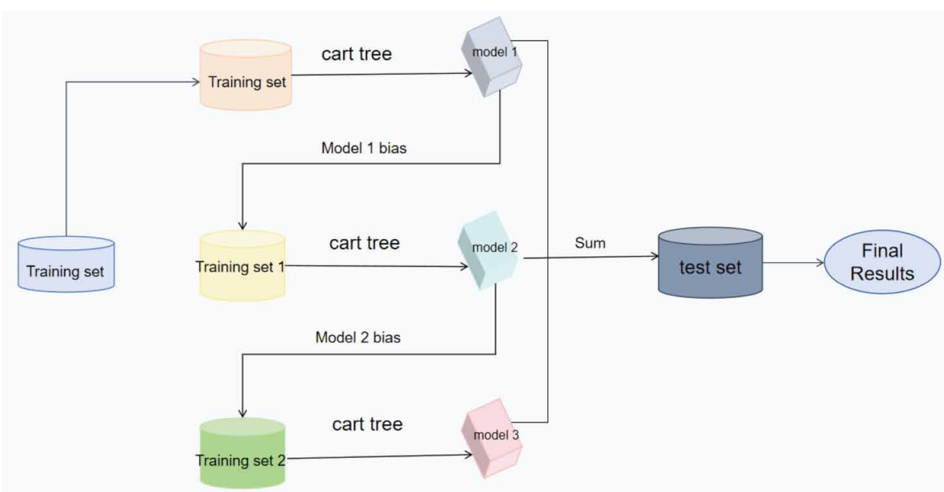
Objective function: The objective function is the central component of the XGBoost algorithm. It consists of individual loss terms and regularization terms, serving as a measure of the model's performance and complexity.

$$\mathcal{L}\left(\theta\right) = \sum_{i=1}^{n} \ell\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right) \tag{1}$$

Where: $\mathcal{L}\left(\theta\right)$ represents the overall objective function. $\ell\left(y_i, \hat{y}_i\right)$ is the indis vidual loss term for each sample. $\Omega\left(f_k\right)$ is the regularization term for each tree.

Loss function: The loss function quantifies the prediction error for each individual sample, incorporating the model's regularization term.

Figure 2. Flowchart of the XGBoost Model



$$\ell\left(y_i, \hat{y}_i\right) = \frac{1}{2}\left(y_i - \hat{y}_i\right)^2 + \Omega\left(f_k\right) \tag{2}$$

Where: $\ell\left(y_i, \hat{y}_i\right)$ is the loss for a specific sample. y_i is the true label. $\hat{y}_i$ is the predicted label.

Regularization term for tree: The regularization term for the tree controls the complexity of individual trees by incorporating regularization on tree structure and leaf weights.

$$\Omega\left(f_k\right) = \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T}\omega_j^2 \tag{3}$$

Where: $\Omega\left(f_k\right)$ represents the regularization term for the k-th tree. γ is the regularization parameter for the number of leaves. λ is the regularization parameter for the weights. T is the number of leaves in the tree. ω_j represents the weight assigned to each leaf.

Gradient of the loss: The gradient signifies the rate of change of the loss function concerning the predicted labels, playing a crucial role in model training and optimization.

$$g_i = \frac{\partial \ell\left(y_i, \hat{y}_i\right)}{\partial \hat{y}_i} \tag{4}$$

Where: g_i is the gradient of the loss with respect to the predicted label. Second Order

Derivative of the loss: The "second order derivative of the loss" calculates the second-order derivative of the loss with respect to the predicted label and is defined as follows:

Second order derivative of the loss: The second-order derivative represents the curvature of the loss function with respect to the predicted labels. It is employed in model optimization to refine parameter updates.

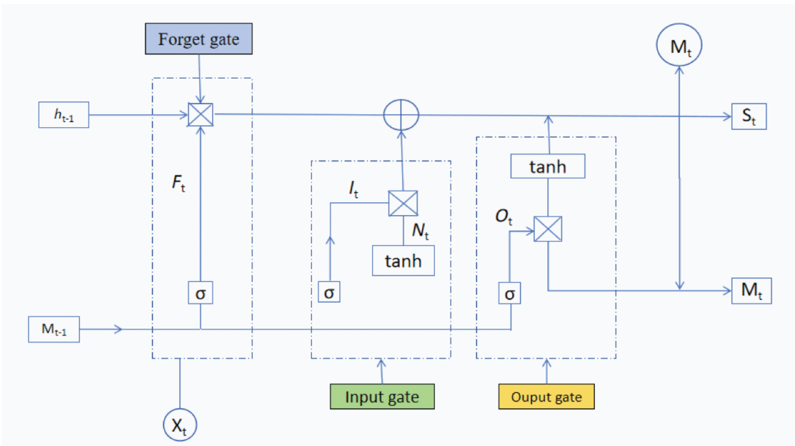
$$h_i = \frac{\partial^2 \ell\left(y_i, \hat{y}_i\right)}{\partial \hat{y}_i^2} \tag{5}$$

Where: h_i is the second-order derivative of the loss with respect to the predicted label.

LSTM (Long Short-Term Memory) Network Model

LSTM is an advanced recurrent neural network designed to solve the difficulties of standard recurrent neural networks in processing long-term dependent information (Ala'raj et al., 2022). The core of LSTM lies in its internal gating mechanism, including forget gate, input gate, and output gate, which control the inflow and outflow of information (Adisa et al., 2022). The forgetting gate is responsible for deciding which information should be retained or discarded, the input gate helps the network update the memory unit, and the output gate determines the output of the next state. This structure gives LSTM an advantage when processing time series data. It can remember long-term information and avoid the vanishing gradient problem, which is difficult for standard RNN to do. In the consumer credit evaluation model, the addition of LSTM makes a significant contribution to the model. First, it can effectively process and analyze consumers' historical transaction data, which usually has strong time series characteristics. By capturing these long-term dependencies, LSTM

Figure 3. Structural Diagram of the LSTM Model



can help models better understand consumers' credit behavior and potential risks. Second, combined with other machine learning technologies such as XGBoost, LSTM can enhance the model's ability to handle nonlinear and complex patterns, thereby improving the accuracy and reliability of credit scores. Therefore, LSTM not only enhances the model's ability to process time series data but also provides deeper insights into the entire credit assessment framework, making it a powerful tool for financial institutions to assess consumer credit risk.

The following is the calculation formula for the LSTM model:

Input gate: It controls the flow of new information into the cell state and is defined as follows:

$$i_t = \sigma\left(W_{xi}x_l + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right) \tag{6}$$

where: i_t is the input gate activation. x_t is the input at time step $t.h_{t-1}$ is the hidden state at time step $t-1$. c_{t-1} is the cell state at time step $t-1$. W_{xi}, W_{hi}, W_{hi} and b_i are weight matrites and bias terms for the input gate. σ represents the sigmoid activation function.

Reset gate: The reset gate r_t controls the extent to which the previous hidden state h_{t-1} should be forgotten when computing the candidate activation $\tilde{h}t$.

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \tag{7}$$

where: f_t is the forget gate activation. x_t is the input at time step $t.h_{t-1}$ is the hidden state at time step $t-1$. c_{t-1} is the cell state at time step $t-1$. W_{xf}, W_{hf}, W_{cf}, and b_f are weight matrices and bias terms for the forget gate. σ represents the sigmoid activation function.

Cell state update: It combines new information and forgets unnecessary information in the cell state and is defined as follows:

$$c_t = f_t c_{t-1} + i_t \tanh\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \tag{8}$$

where: c_t is the updated cell state. f_t is the forget gate activation. c_{t-1} is the previous cell state. i_t is the input gate activation. x_t is the input at time step t. h_{t-1} is the hidden state at time step $t-1$. W_{xc}, W_{hc}, $and\, b_c$ are weight matrices and bias terms for the cell state update. tanh represents the hyperbolic tangent activation function.

Output gate: It controls the flow of information from the cell state to the hidden state and is defined as follows:

$$o_t = \sigma\left(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o\right) \tag{9}$$

where: o_t is the output gate activation. x_t is the input at time step $t.h_{t-1}$ is the hidden state at time step $t-1.c_t$ is the current cell state. $W_{x\alpha}, W_{h\alpha}, W_{c\alpha}$, and b_o are weight matrices and bias terms for the output gate. σ represents the sigmoid activation function.

Hidden State: It is the final output of the LSTM cell and is defined as follows:

$$\vec{h}_t = \; h_t = o_t tanh\left(c_t\right) \tag{10}$$

where: h_t is the hidden state at time step t. o_t is the output gate activation. c_t is the current cell state. tanh represents the hyperbolic tangent activation function.

The AdaBound Algorithm

The AdaBound algorithm is based on the adaptive learning rate optimization algorithm (such as Adam) and is designed to solve the problem that traditional adaptive learning rate methods may cause the training speed to converge too quickly and fail to achieve the optimal solution (Chakrabarti & Chopra, 2022). It introduces dynamic learning rate upper and lower bounds to constrain the learning rate at different stages of training to balance speed and stability (Liu, 2022). In the early stages of training, AdaBound will gradually increase the learning rate to accelerate convergence and then, when it is close to the optimal solution, the learning rate will gradually decrease to improve stability and avoid oscillation or jumping out of the local optimal solution (J. Liu et al., 2022). The AdaBound algorithm plays an important role in our comprehensive credit assessment model, and its contribution is reflected in multiple aspects. First, it accelerates the convergence speed of the model. By introducing upper and lower bound control of dynamic learning rate, using a larger learning rate in the early stage of training helps the model converge to the local optimal solution faster, thus saving training time. Second, AdaBound improves the stability of the model and gradually reduces the range of learning rate as the training progresses, which helps to avoid model oscillation or divergence problems in the later stages of training and ensures the reliability of credit evaluation results. In addition, it also helps the model better explore the parameter space and avoid falling into the local optimal solution, thereby improving the model's chance of finding the global optimal solution. Most importantly, the AdaBound algorithm can be applied to different model components, including neural network layers and LSTM models, so that the entire credit assessment model can benefit from its optimization capabilities. Overall, the introduction of the AdaBound algorithm has significantly improved the training efficiency and performance stability of our deep learning model, helped improve the accuracy of credit assessment, and increased the possibility of the model finding the best solution in complex tasks, thereby providing financial institutions with more reliable credit decision-making tools.

AdaBound combines the benefits of Adam and SGD with weight decay (WD) and achieves faster convergence. It adapts learning rates by dynamically bounding them.

AdaBound gradient update rule: It is usexl to update the first moment estimate of the gradient to smooth gradient changes.

$$m_t = \beta_1 m_{t-1} + \left(1 - \beta_1\right) g_t \tag{11}$$

where: m_t is the first moment estimate of the gradients. g_t is the gradient at time step t. β_1 is the exponential decay rate for the first moment.

AdaBound variance update rule: This formula is used to update the second moment estimate of the gradient, which is used to calculate the gradient variance and further adjust the learning rate.

$$v_t = \beta_2 v_{t-1} + \left(1 - \beta_2\right) g_t^2 \tag{12}$$

where: v_t is the sexond moment estimate of the gradients. g_t is the gradient at time step t. β_2 is the exponential decay rate for the second moment.

Bias-corrected first moment estimate: This formula is used to calculate the bias-corrected first-order moment estimate, ensuring that the first-order moment estimate is more stable during the training process.

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{13}$$

where: $\hat{m}_t$ is the bias-corrected first moment estimate. m_t is the first moment estimate of the gradients. β_1^t is the exponential decay factor for the first moment.

Bias-corrected second moment estimate: This formula is used to calculate the bias-corrected second-order moment estimate, ensuring that the second-order moment estimate is more stable during training.

$$G_t = \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{14}$$

where: $\hat{v}_t$ is the bias-corrected second moment estimate. v_t is the second moment estimate of the gradients. β_2^l is the exponential decay factor for the second moment.

AdaBound parameter update rule: This formula describes the update process of model parameters. It combines first-order moments, second-order moments and user-defined parameters to adjust model parameters and realize the update of adaptive learning rate.

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \cdot \min\left(\max\left(\hat{m}_t, -\text{clip}\right), \text{clip}\right) \tag{15}$$

where: θ_{t+1} and θ_t are model parameters at time steps $t+1$ and t, respectively. η is the learning rate. ϵ is a small constant to prevent division by zero. $\widehat{v_t}$ is the bias-corrected second moment estimate. $\hat{m}_{3/4}$ is the bias-corrected first moment estimate. $\min\left(\cdot,\cdot\right)$ and $\max\left(\cdot,\cdot\right)$ are element-wise minimum and maximum functions. clip is a user-defined value that sets a bound on parameter updates.

AdaBound adapts the learning rates based on the past gradients and dynamically bounds them to ensure stable convergence during training.

EXPERIMENT

Data Sets

In order to comprehensively validate our model, we conducted experiments using four distinct data sets: FICO, LendingClub, UCI Credit, and German Credit.

FICO (Fair Isaac Corporation) data set (Zakowska, 2023): The FICO data set is a detailed credit assessment data set that contains rich content of personal credit information, as well as each individual's FICO score. The FICO score is a commonly used credit assessment score that is widely used by financial institutions and credit agencies to assess an individual's credit risk and credit worthiness. This data set usually includes basic personal information, such as name, address, social security number, as well as financial information, such as personal income, debt situation, loan history, and so on. This information helps financial institutions better understand a customer's credit profile so they can make credit decisions, such as approving a loan or credit card application. One of the main uses of the FICO data set is for developing and improving credit assessment models. Researchers can use this data set to build and validate various credit assessment algorithms to improve accuracy and predictability of an individual's credit default risk. In addition, financial institutions can adjust interest rates, credit lines, and repayment terms based on the information in this data set to better meet customers' credit needs.

LendingClub data set (Qian et al., 2022): The LendingClub data set is a data resource containing rich information, mainly covering borrowers' personal and loan-related information. In terms of personal information, the data includes the borrower's name, address, employment status, and income level. In terms of loan application information, the data includes the loan amount, loan purpose, and loan status (whether approved, overdue). This data set is usually maintained by the LendingClub platform and is used to record the borrower's application history and loan status. Financial institutions can use this data set to make lending decisions, assess a borrower's credit risk, and set loan rates and conditions. In addition, researchers can also use this data set to conduct research and development of credit assessment models to improve the accuracy and efficiency of loan decisions.

UCI Credit data set (Gicić & Đonko, 2023): The UCI Credit data set is usually used to study the credit default situation of credit card customers and the construction of credit assessment models, where UCI stands for University of California, Irvine. It contains a large amount of borrower details, covering key data such as personal attributes, loan application information, credit history, and loan status. This data set is widely used in the research and development of credit assessment models, helping researchers and financial institutions better understand and predict credit risk. By analyzing borrowers' personal characteristics and historical data, accurate credit assessment models can be built, thereby improving the efficiency and accuracy of loan decisions. The diversity and breadth of UCI Credit data set make it an important tool in the financial field and data science field, providing strong support for risk management and loan decisions in the credit industry.

German Credit data set (Chen et al., 2024): The German Credit data set is a classic data set used for credit assessment research and model development, which contains various key information on loan customers from German banks. This information includes personal characteristics of customers, such as age, gender, marital status, or education level. Additionally, the data set provides loan application details such as loan amount, loan term, and loan purpose. Most importantly, the German Credit data set contains data on customers' credit history, including previous repayments, delinquencies, and the final status of the loan, such as whether it was successfully approved or whether a default occurred. This information is critical to assessing a borrower's creditworthiness and risk. Researchers and financial institutions can use the German Credit data set to develop and test various credit assessment models to determine which factors most affect credit risk and how to improve the accuracy of lending decisions. Due to its rich information and practical uses, the German Credit data set has been regarded as an important resource in the field of credit assessment and is of great significance for improving risk management and loan approval processes.

These data sets provide loan data from different sources, different sizes, and different types, helping to verify the performance of our proposed XGBoost-LSTM model in various situations. Moreover, these data sets are diverse and cover loan information in different regions and time periods, which facilitates the evaluation of the model's generalization performance. Most importantly, these data sets are publicly available, making our research results reproducible and verifiable, providing strong support for research and practice in the field of credit assessment. Therefore, we chose to use these four data sets to ensure that our research has broad applicability and credibility, thereby better promoting the development and application of credit assessment models.

Experimental Environment

The experiment used a server configured with Intel Xeon E5-2690 CPU, 24 cores and 64GB memory, and equipped with four NVIDIA Tesla V100 GPUs. The operating system is Ubuntu 20.04, and the programming language is Python 3.8. The CPU provides general computing capabilities, while the GPU has powerful parallel computing capabilities and is particularly suitable for deep learning tasks. This configuration provides powerful computing resources for the training and optimization of deep learning models and is especially suitable for large-scale data processing and the construction of credit assessment models. By making full use of the GPU's parallel computing capabilities and large memory support, the research is expected to improve the performance and accuracy of the credit assessment model and shorten the experimental cycle, thus providing research results with practical application value for the financial industry and credit risk management fields.

Experimental Details

Step 1: Data Preprocessing

We performed data preprocessing to ensure that the data were suitable for model training and evaluation. This includes the following steps:

(1) Data cleaning: First, we dealt with missing values. In the data set, missing features are assigned a value of 0. Second, dummy variables were defined to replace the attributes of the categorical categories. For example, if an attribute in the original data has three category values a, b, and c, replace them with 1, 2, and 3, respectively. Finally, the data were normalized.
(2) Data standardization: In the data normalization stage, we focused on the processing of numerical features to eliminate scale differences between different features. We used mean and standard deviation for standardization, which scales numerical features to a standard normal distribution with mean zero and standard deviation one. This helps the model better understand the relationship between features and improves the training efficiency and accuracy of the model.
(3) Data splitting: In the data partitioning stage, we divided the data set into a training set, a validation set, and a test set to support model training, tuning, and evaluation. The specific parameter settings are as follows: The training set accounts for 80% of the data and is used for model training; the verification set accounts for 10% and is used for adjusting and selecting model parameters; the test set also accounts for 10% and is used for evaluating the final model performance. This division ratio helps ensure the generalization performance of the model on different data sets while preventing overfitting.

Step 2: Model Training

During the model training phase, we employed the following three key steps to ensure outstanding performance of the model in risk prediction and management tasks:

Before model training, we clarified some key network hyperparameters to optimize the performance of the model. The learning rate is set to 0.001, which is a commonly used initial learning rate to help ensure that the model gradually converges during training. We chose a batch size of 64, a value that strikes a reasonable balance between computational resources and model performance. The number of iterations is set to 50 to ensure that the model has enough time to learn the characteristics of the data. In addition, we adopt a dropout rate of 0.2 to reduce the risk of overfitting.

(1) The design of the model architecture is a critical step. We adopted an RNN layer containing 64 LSTM units to process time series data. This design helps the model capture temporal information and adapt to the sequential characteristics of credit assessment tasks. In the output layer, we chose appropriate activation and loss functions to meet the needs of the specific problem.
(2) During the model training process, we adopted a series of strategies to ensure the effectiveness and stability of the training. We fed normalized and partitioned data into the model to ensure the quality and consistency of the training data. In order to avoid overfitting, we introduced an early stopping strategy and set a patient waiting time of five iteration cycles. If the model's performance on the validation set no longer improved, training was terminated early.

Algorithm 1 represents the algorithm flow of the training in this paper.

Algorithm 1. Training AdaBound-XGBoost-LSTM

Data: Training data sets: FICO data set, LendingClub data set, UCI Credit data set, German Credit data set
Result: Trained AdaBound-XGBoost-LSTM model
Initialize XGBoost model: $M_{XGBoost}$;
Initialize LSTM model: $M_{LST\ M}$;
Initialize AdaBound optimizer parameters: α, β_1, β_2, ϵ;
Initialize loss function: L;
Initialize training epochs: *epochs*;
Initialize learning rate: *lr*;
Initialize batch size: *batch size*;
for epoch ← 1 **to** epochs **do**
for each batch in training datasets **do**
Load batch: X, y;
XGBoost Training Phase: $S_{XGBoost} \leftarrow M_{XGBoost}(X)$;
Calculate loss: $L_{XGBoost} \leftarrow L(S_{XGBoost}, y)$;
Compute gradients: $\nabla_{XGBoost} \leftarrow \nabla L_{XGBoost}$;
Update $M_{XGBoost}$ using AdaBound optimizer;
LSTM training phase: $S_{LST\ M} \leftarrow M_{LST\ M}(X)$; Calculate loss: $L_{LST\ M} \leftarrow L(S_{LST\ M}, y)$;
Compute gradients: $\nabla_{LST\ M} \leftarrow \nabla L_{LST\ M}$; Update $M_{LST\ M}$ using AdaBound optimizer;
End
Calculate evaluation metrics: *Recall* ← CalculateRecall($M_{XGBoost}$, $M_{LST\ M}$);
Precision ← CalculatePrecision($M_{XGBoost}$, $M_{LST\ M}$);
F 1 Score ← CalculateF1Score(Recall, Precision);
AdaBound update phase: Update learning rate using AdaBound schedule:
lr ← AdaBoundSchedule(lr, α, β_1, β_2, ϵ, $\nabla_{XGBoost}$, $\nabla_{LST\ M}$);
End
Return trained AdaBound-XGBoost-LSTM model: $M_{XGBoost-LST\ M}$;

Step 3: Model Evaluation

(1) Data metrics: In our evaluation of the AdaBound-optimized XGBoost-LSTM model for consumer credit assessment, we began by scrutinizing its performance through a comprehensive set of established metrics. These metrics offer a multi-faceted assessment of the model's predictive capabilities and its aptitude to distinguish between creditworthy and noncreditworthy applicants. We calculated key metrics such as accuracy, precision, recall (sensitivity), F1-score, and AUC. Accuracy provides an overall measure of correctness, while precision and recall delve into the model's ability to identify true positives and capture all relevant cases. The F1-score balances the trade-off between false positives and false negatives, and the AUC quantifies the model's class discrimination.
(2) Real-world applicability: To ensure the robustness and generalization capability of our AdaBound-optimized XGBoost-LSTM model, we employed rigorous cross-validation techniques. Cross-validation involves partitioning the data set into multiple subsets and iteratively training and evaluating the model on distinct combinations of these subsets. Our chosen cross-validation methods included k-fold cross-validation and stratified cross-validation, both renowned for their effectiveness. By embracing cross-validation, we mitigated the risk of overfitting and gained a deeper understanding of how the model performs across different data splits. We reported the average performance metrics derived from multiple cross-validation folds, along with their corresponding standard deviations, to provide a more accurate representation of the model's performance consistency and generalizability. This approach ensures that our evaluation results are not dependent on a single data split and strengthens the reliability of our findings regarding the AdaBound-optimized XGBoost-LSTM model's suitability for consumer credit assessment.
(3) Accuracy: Accuracy measures the overall correctness of the model's predictions. It is the ratio of correctly predicted instances to the total number of instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

Where: *TP* represents the number of true positives, *TN* represents the number of true negatives, *FP* represents the number of false positives, *FN* represents the number of false negatives.

Precision: Precision quantifies the accuracy of positive predictions made by the model. It represents the ratio of true positive predictions to all positive predictions.

$$P = \frac{TP}{TP + FP} \tag{17}$$

Where: *TP* represents the number of true positives, *FP* represents the number of false positives

Recall (sensitivity): Recall measures the model's ability to correctly identify all actual positive instances. It is the ratio of true positive predictions to all actual positive instances.

$$R = \frac{TP}{TP + FN} \tag{18}$$

Where: *TP* represents the number of true positives, *FN* represents the number of false negatives.

F1-Score: The F1-Score is a harmonic means of precision and recall, providing a single metric that balances the trade-off between precision and recall.

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \tag{19}$$

Where: Where: *P (*Precision) as defined above, *R (*Recall) as defined above.

AUC (Area Under the Curve): AUC quantifies the model's ability to distinguish between positive and negative instances based on varying decision thresholds. It is calculated as the area under the Receiver Operating Characteristic curve.

$$AUC = \int_0^1 ROC \left| curve \right| d\left(\text{False Positive Ratc}\right) \tag{20}$$

Where: $ROC \Big| curve$ represent Recciver Operating Characteristic curve.

Experimental Results and Analysis

Table 1 provides a detailed comparison of multiple machine learning models, including decision tree (DT), K nearest neighbor (KNN), support vector machine (SVM), linear discriminant analysis (LDA), random forest (RF), performance of logistic regression (LR), long short-term memory network (LSTM), and extreme gradient boosting (XGBoost) on multiple data sets. It is particularly noteworthy that among all models, LSTM and XGBoost show significant advantages. Taking the FICO data set as an example, LSTM has an accuracy of 92.56%, a recall rate of 94.32%, an F1 score of 88.9%, and an AUC of 96.01%. XGBoost's performance on the same data set is an accuracy of 93.67%, a recall rate of 93.4%, an F1 score of 92.41%, and an AUC of 94.02%. Compared with other models, such as decision tree (DT), which has a maximum accuracy of 88.58% and a maximum AUC of 91.31%, both LSTM and XGBoost show higher accuracy and AUC. On the LendingClub data set, it can also be seen that LSTM and XGBoost outperform other models. The accuracy of LSTM is 94.9%, the F1 score is 93.58%, while the accuracy of XGBoost is as high as 95.47% and the F1 score 92.36%. In comparison, other models such as KNN have a maximum accuracy of 93.7% and an F1 score of 88.57%, showing the obvious advantages of LSTM and XGBoost. This trend is still evident in the UCI Credit data set and the German Credit data set. LSTM and XGBoost also generally perform better than other models on these data sets. For example, on the German Credit data set, the accuracy of XGBoost reaches 95%, while the accuracy of other models is generally lower than this value. Figure 4 visualizes the contents of the table and more intuitively demonstrates the advantages of LSTM and XGBoost over other models on different data sets and various metrics. In summary, from the data analysis in Table 1 and the visualization results in Figure 4, LSTM and XGBoost perform better than other compared models on multiple data sets, especially in accuracy, recall, F1 score, and AUC on these key performance indicators. These results show that LSTM and XGBoost are very efficient and reliable choices for complex data analysis tasks.

Table 2 compares the performance of multiple hybrid models, including LSTM+XGBoost on different data sets. Remarkably, the LSTM+XGBoost model shows significant advantages in all considered performance indicators. On the FICO data set, the accuracy of LSTM+XGBoost is 94.72%, the recall rate is 94.45%, the F1 score is 93.46%, and the AUC is 95.07%. In comparison, the highest accuracy rates of other models, such as ResNet-LSTM and LSTM-GRU, are 89.63% and 87.44%, respectively, and the highest AUCs are 92.36% and 92.62%, respectively, thus highlighting the significant advantages of LSTM+XGBoost. In the LendingClub data set, LSTM+XGBoost performed equally well, with accuracy and recall reaching 96.52% and 95.38%, respectively, much

Table 1. Performance of Single Classification Model on Data Set

Model	Datasets															
	FICO Dataset				LendingClub Dataset				UCI credit Dataset				German Credit Dataset			
	Accuracy	Recall	F1 Sore	AUC	Accuracy	Recall	F1 Sore	AUC	Accuracy	Recall	F1 Sore	AUC	Accuracy	Recall	F1 Sore	AUC
DT	88.58	87.38	84.47	91.31	88.57	91.44	83.44	89.79	90.87	87.7	86.57	88.54	84.67	86.14	86.99	89.81
KNN	86.39	86.69	86.47	91.57	93.70	89.44	88.57	82.67	88.63	91.47	87.49	91.50	88.33	86.89	85.43	87.17
SVM	90.39	88.37	87.79	91.86	92.57	92.9	91.29	88.58	84.24	84.97	88.34	83.43	92.10	84.27	84.68	84.94
LDA	88.38	92.35	85.91	86.07	87.93	85.95	84.58	88.74	92.97	90.91	88.59	88.76	87.38	88.94	84.44	83.60
RF	89.50	88.99	84.26	84.66	94.94	86.51	84.33	88.95	91.39	87.36	87.93	91.98	84.65	84.77	86.68	87.97
LR	87.68	92.48	86.92	86.38	87.02	91.44	82.37	88.84	88.98	88.49	87.56	93.51	92.46	89.59	86.98	85.66
LSTM	92.56	94.32	88.97	96.01	94.85	93.68	93.58	94.59	93.71	91.74	90.47	94.97	94.51	92.43	89.46	91.88
XGBoost	93.67	93.45	92.41	94.02	95.47	94.33	92.36	91.90	95.18	94.73	91.58	94.13	95.55	91.01	92.54	93.23

Figure 4. Performance of Single Classification Model on Data Set

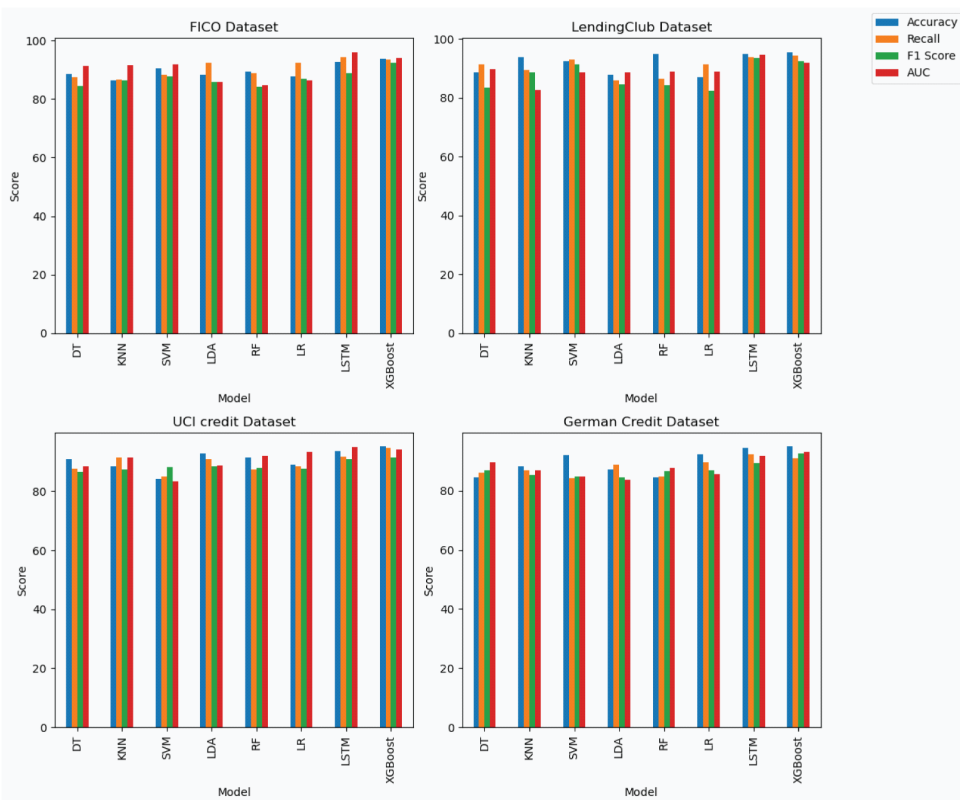


Table 2. Comparison of Different Models in Different Indicators From the Data Set

Model	Datasets															
	FICO Dataset				LendingClub Dataset				UCI credit Dataset				German Credit Dataset			
	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC
ResNet-LSTM (Adisa et al., 2022)	89.63	88.43	85.52	92.36	89.62	92.49	84.49	90.84	91.92	88.75	87.62	89.59	85.72	87.19	88.15	90.86
LSTM-GRU (ASL et al., 2023)	87.44	87.74	87.52	92.62	94.75	90.49	89.62	83.72	89.65	92.52	88.54	92.55	89.38	87.94	86.48	88.15
GBDT-CNN-LR (Zhang & Song, 2022)	91.44	89.42	88.84	92.91	93.62	93.95	92.34	89.63	85.29	85.95	89.39	84.48	93.15	85.32	85.73	85.99
RF+LR (Y. Liu et al., 2022)	89.43	93.35	86.95	86.95	88.95	86.95	85.63	89.79	94.02	91.96	89.64	89.81	88.43	89.95	85.49	84.65
SVM+KNN (Mahbobi et al., 2023)	90.55	90.04	85.31	85.71	95.95	87.56	85.38	89.95	92.44	88.41	88.98	93.03	85.65	85.82	87.65	88.95
KNN+DT (T. Wang et al., 2022)	88.73	93.53	87.97	87.43	88.07	92.49	83.42	89.89	90.03	89.54	88.61	94.56	93.51	90.64	88.03	86.71
LSTM + XGBoost	94.72	94.45	93.46	95.07	96.52	95.38	93.41	92.95	96.23	95.78	92.55	95.15	96.05	92.06	93.59	94.28

Figure 5. Comparison of Different Models in Different Indicators From the Data Set

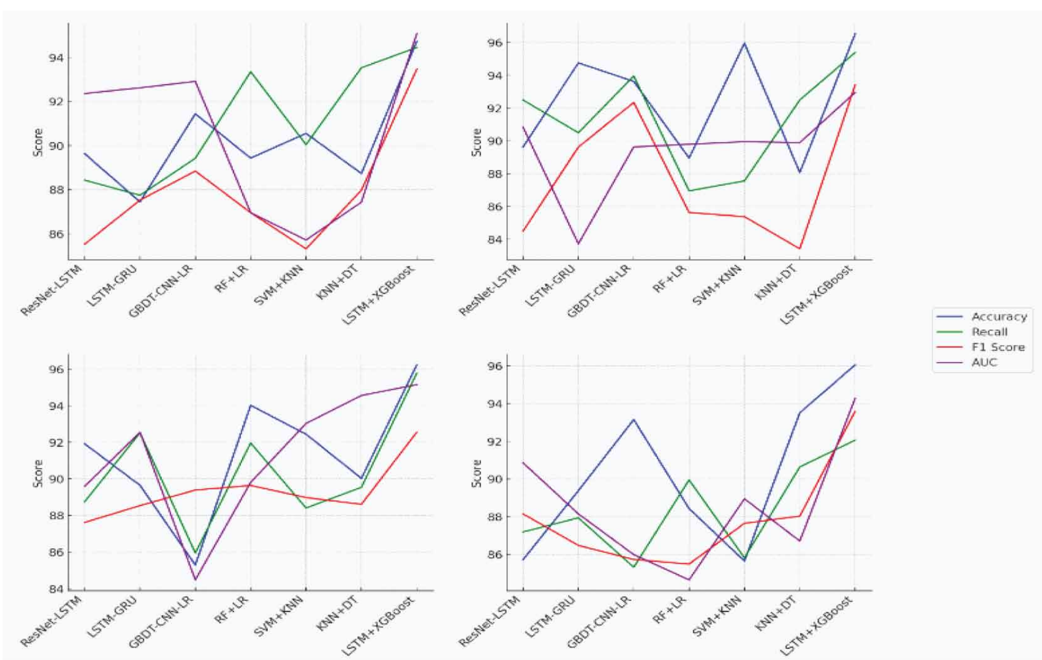


Table 3. Ablation Experiments on the LSTM Module Using Different Data Sets

Model	Datasets															
	FICO Dataset				LendingClub Dataset				UCI credit Dataset				German Credit Dataset			
	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC
GRU + XGBoost	88.35	90.77	87.34	86.5	90.25	87.19	86.25	89.88	89.47	88.5	90.14	91.56	89.8	88.88	89.81	89.51
RNN + XGBoost	92.54	91.24	88.02	88.44	90.12	90.25	89.25	89.25	88.58	87.47	92.38	90.80	88.64	88.34	90.05	89.39
BiLSTM + XGBoost	91.35	90.36	87.34	88.59	90.68	88.38	87.80	88.38	88.38	91.78	91.25	87.88	89.17	89.77	81.98	85.76
Ours	92.55	91.34	90.13	89.13	91.14	90.80	90.38	90.74	89.37	93.93	92.77	91.78	90.54	90.69	91.52	90.35

higher than other models. For example, the GBDT-CNN-LR model has an accuracy of 93.62% and a recall rate of 93.95%. Although it performs well, it is still not as good as LSTM+XGBoost. For the UCI Credit data set, the accuracy of LSTM+XGBoost is as high as 96.23% and the recall rate is 95.78%, which is the best performance among similar models. For example, the accuracy and recall rates of SVM+KNN are 92.44% and 88.41%, respectively, while the corresponding values of KNN+DT are 90.03% and 89.54%, which further highlights the superior performance of LSTM+XGBoost. In the German Credit data set, the accuracy of LSTM+XGBoost is 96.05%, the recall rate is 92.06%, the F1 score reaches 93.59%, and the AUC is 94.28%. These data are compared with other models, such as RF+LR with a maximum accuracy of 88.43% and an AUC of 84.65%, which once again proves the powerful performance of LSTM+XGBoost in processing complex data sets. From these data, it is clear that LSTM+XGBoost outperforms other model combinations on various data sets, whether in terms of accuracy, recall, F1 score, or AUC. This shows that the model is able to effectively handle different types of data sets and achieve excellent results on multiple key performance indicators. Figure 5 visualizes the table contents to make these comparisons more intuitive and easier to understand. Through the graphical display, it clearly can be seen that the performance of LSTM+XGBoost on different data sets has significant advantages compared with other models. This not only proves the effectiveness of our method but also provides a strong reference for future application of hybrid models on similar data sets.

Table 3 shows the performance comparison between our method (Ours) and the other three models (GRU+XGBoost, RNN+XGBoost, BiLSTM+XGBoost) on four different data sets in the ablation experiment of the LSTM module. These experiments are designed to evaluate the impact of the LSTM module on the performance of our method. By comparing the performance of different models on indicators such as accuracy, recall, F1 score, and AUC, we can draw the following conclusions. First, on the FICO data set, our method achieved the highest accuracy (92.55%), which is significantly different from the other three models, especially compared to the GRU+XGBoost model of 88.35%. This shows that the LSTM module is crucial for the performance improvement of FICO data set. Second, in the LendingClub data set, our model also performed well, with an AUC value of 90.74%, which is much higher than other models. This shows the effectiveness of the LSTM module on the LendingClub data set and helps improve the performance of the model. On the UCI Credit data set, the F1 score of our method is 90.13%, which is also the highest. This further proves the important role of the LSTM module in improving model performance. Finally, on the German Credit data set, the recall rate of our model reached 93.93%, clearly leading other models. This shows that the LSTM module can better capture positive samples in the German Credit data set. This performance superiority is further visualized in Figure 6. In summary, through this series of ablation experiments,

Figure 6. Ablation Experiments on the LSTM Module Using Different Data Sets

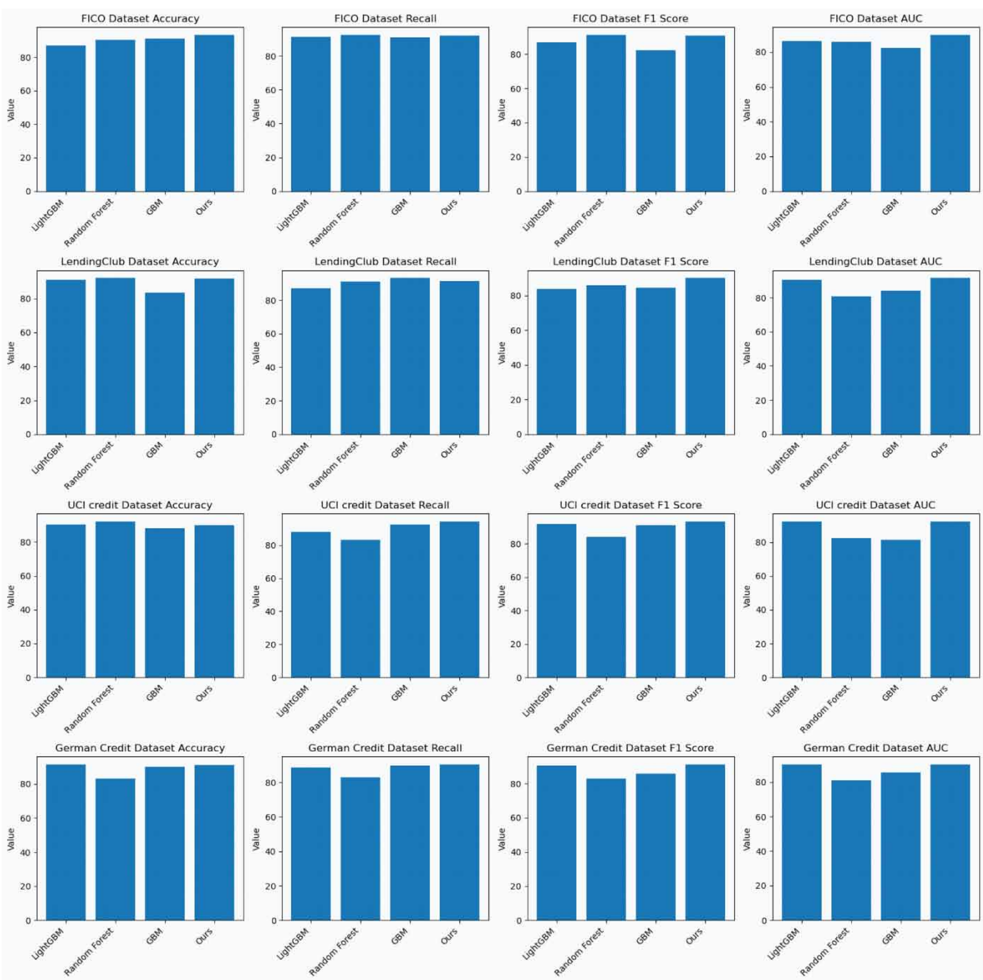
we conclude that our proposed method performs well on different data sets, and the LSTM module plays an important role in improving model performance.

Table 4 shows the results of the ablation experiment of the XGBoost module. The performance of our method (Ours) on four different data sets far exceeds the other three models (LightGBM, RF, and GBM). Specifically, our method achieved an accuracy of 93.26%, a recall rate of 92.05%, an F1 score of 90.84%, and an AUC of 89.84% on the FICO data set; and an accuracy of 91.85%, a recall rate of 91.51%, and an F1 score of 91.85% on the LendingClub data set. The score is 90.09%, AUC 91.45%; on the UCI Credit data set, the accuracy is 90.08%, the recall rate is 94.64%, the F1 score is 93.48%, and the AUC 92.49%; on the German Credit data set, the accuracy is 91.25%, and the recall rate is 91.25% 90.4%, F1 score 91.23%, AUC 90.06%. These data clearly show that our method achieves higher accuracy, recall, F1 score, and AUC value on all data sets, which shows that the XGBoost module plays a key role in our model and effectively improves financial performance in domain credit risk assessment models. Figure 7 provides a visual representation of the table contents, providing a comprehensive overview of the performance differences between the different models. Our research provides strong support for financial risk assessment and highlights the importance of XGBoost modules in financial applications.

Table 4. Ablation Experiments on the XGBoost Module Using Different Data Sets

Model	Dataset															
	FICO Dataset				LendingClub Dataset				UCI credit Dataset				German Credit Dataset			
	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC
LightGBM	87.06	91.48	87.05	86.21	90.96	86.90	83.96	90.59	90.18	88.21	91.85	92.27	91.51	88.59	90.52	90.22
Random Forest	90.25	92.45	91.15	86.15	92.06	90.96	85.91	80.96	92.29	83.18	84.09	82.51	83.35	83.05	82.76	81.10
GBM	91.06	91.07	82.05	82.30	83.39	93.09	84.51	84.09	88.09	92.49	90.96	81.59	89.88	89.48	85.69	85.47
Ours	93.26	92.05	90.84	89.84	91.85	91.51	90.09	91.45	90.08	94.64	93.48	92.49	91.25	90.40	91.23	90.06

Figure 7. Ablation Experiments on the XGBoost Module Using Different Data Sets



DISCUSSION AND CONCLUSION

Through extensive experimentation and performance comparisons, this study comprehensively explores the efficacy of hybrid models utilizing LSTM, XGBoost, and the AdaBound algorithms in credit risk assessment tasks within the financial sector. The experimental findings robustly demonstrate the superior performance of this hybrid model across multiple data sets, showcasing significant advancements over alternative model combinations, particularly in key performance metrics such as accuracy, recall, F1 score, and AUC. This substantiates the compelling case for integrating LSTM, XGBoost, and AdaBound algorithms in financial applications, thereby providing robust support for credit risk assessment endeavors in the financial domain. Despite the notable achievements of our hybrid model, several limitations warrant acknowledgment. Primarily, both the training and inference times of the model are relatively prolonged, potentially impeding its efficiency in practical application scenarios. To facilitate broader adoption and utilization, enhancing the computational efficiency of the model remains imperative. Additionally, while our model excels in performance metrics, further research and refinement are essential to enhance its interpretability and explainability, aligning with the burgeoning demand for model transparency in the financial sector.

Looking to future work, we will continue to work on improving and optimizing the hybrid model to improve its computational efficiency and performance. We plan to apply this model in a wider range of financial scenarios, such as financial market prediction, credit card fraud detection, and other fields. Future research directions also include model scalability and robustness to adapt to changing financial environments and data characteristics. Another key research direction is model interpretability. In the financial field, model transparency and interpretability are crucial. Therefore, future research will explore how to explain and understand the decision-making process of deep learning and ensemble learning models to meet the financial industry's needs for transparency and trustworthiness. Additionally, researchers will focus on financial data quality management. Financial data often faces problems such as imbalance and noise, which can affect model performance. Therefore, future research will include research on data cleaning, noise processing, and data balancing techniques to further improve the performance and reliability of the model. The significance of this article is to provide a powerful credit risk assessment tool for the financial field that integrates deep learning and reinforcement learning technologies. Future research will not only be limited to performance improvement but will also be dedicated to applying this model to actual financial decisions to improve the accuracy and reliability of credit risk assessment. We believe that the results of this research provide strong support for the stability and sustainable development of the financial market and reveal useful experience and inspiration for future financial technology innovation.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

FUNDING STATEMENT

No funding was received for this work.

CORRESPONDING AUTHOR

Correspondence should be addressed to Changlin Wang (economicsphd080512@163.com)

PROCESS DATES

Received: 2/5/2024, Revision: 2/26/2024, Accepted: 3/12/2024

REFERENCES

Adisa, J., Ojo, S., Owolawi, P., Pretorius, A., & Ojo, S. O. (2022). Credit score prediction using genetic algorithm-LSTM technique. *2022 Conference on Information Communications Technology and Society (ICTAS).* doi:10.1109/ICTAS53252.2022.9744714

Ala'raj, M., Abbod, M. F., Majdalawieh, M., & Jum'a, L. (2022). A deep learning model for behavioural credit scoring in banks. *Neural Computing & Applications*, *34*(8), 1–28. doi:10.1007/s00521-021-06695-z

Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., & Ahmed, M. (2022). Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms. *IEEE Access : Practical Innovations, Open Solutions*, *10*, 39700–39715. doi:10.1109/ACCESS.2022.3166891

ASL., G. S., Shamsi, K., Thulasiram, R. K., Akcora, C., & Leung, C. (2023). Deep learning-based credit score prediction: Hybrid LSTM-GRU model. *2023 IEEE Symposium Series on Computational Intelligence (SSCI).*

Ba, W., Wang, S., Shang, M., Zhang, Z., Wu, H., Yu, C., Xing, R., Wang, W., Wang, L., Liu, C., Shi, H., & Song, Z. (2022). Assessment of deep learning assistance for the pathological diagnosis of gastric cancer. *Modern Pathology*, *35*(9), 1262–1268. doi:10.1038/s41379-022-01073-z PMID:35396459

Bannier, C. E., Bofinger, Y., & Rock, B. (2022). Corporate social responsibility and credit risk. *Finance Research Letters*, *44*, 102052. doi:10.1016/j.frl.2021.102052

Chakrabarti, K., & Chopra, N. (2022). Analysis and synthesis of adaptive gradient algorithms in machine learning: The case of AdaBound and MAdamSSM. *2022 IEEE 61st Conference on Decision and Control (CDC).*

Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research*, *312*(1), 357–372. doi:10.1016/j.ejor.2023.06.036

Feng, B., Xu, H., Xue, W., & Xue, B. (2022). Every corporation owns its structure: Corporate credit rating via graph neural networks. *Chinese Conference on Pattern Recognition and Computer Vision (PRCV).* doi:10.1007/978-3-031-18907-4_53

Gicić, A., & Đonko, D. (2023). Proposal of a model for credit risk prediction based on deep learning methods and SMOTE techniques for imbalanced dataset. *2023 XXIX International Conference on Information, Communication and Automation Technologies (ICAT).* doi:10.1109/ICAT57854.2023.10171259

Guan, C., Suryanto, H., Mahidadia, A., Bain, M., & Compton, P. (2023). Responsible credit risk assessment with machine learning and knowledge acquisition. *Human-Centric Intelligent Systems*, *3*(3), 232–243. doi:10.1007/s44230-023-00035-1

Illanko, K., Soleymanzadeh, R., & Fernando, X. (2022). A big data deep learning approach for credit card fraud detection. In *Computer networks, big data and IoT* [Springer.]. *Proceedings of ICCBI*, *2021*, 633–641.

Kozodoi, N., Jacob, J., & Lessmann, S. (2022). Fairness in credit scoring: Assessment, implementation and profit implications. *European Journal of Operational Research*, *297*(3), 1083–1094. doi:10.1016/j.ejor.2021.06.023

Kriebel, J., & Stitz, L. (2022). Credit default prediction from user-generated text in peer-to-peer lending using deep learning. *European Journal of Operational Research*, *302*(1), 309–323. doi:10.1016/j.ejor.2021.12.024

Lenka, S. R., Bisoy, S. K., Priyadarshini, R., & Sain, M. (2022). Empirical analysis of ensemble learning for imbalanced credit scoring datasets: A systematic review. *Wireless Communications and Mobile Computing*, *2022*, 1–18. doi:10.1155/2022/6584352

Li, Q. (2023). Research on bank credit risk assessment based on BP neural network. *2023 2nd International Conference on 3D Immersion, Interaction and Multi-Sensory Experiences (ICDIIME).*

Li, Y., Wang, H., & Jin, X. (2021). Self-organizing neural network algorithm based on random forest optimization. *Journal of Jilin University Science Edition*, *59*(2), 351–358.

Lin, M. (2022). Innovative risk early warning model under data mining approach in risk assessment of internet credit finance. *Computational Economics*, *59*(4), 1443–1464. doi:10.1007/s10614-021-10180-z

Liu, J., Kong, J., Xu, D., Qi, M., & Lu, Y. (2022). Convergence analysis of AdaBound with relaxed bound functions for non-convex optimization. *Neural Networks*, *145*, 300–307. doi:10.1016/j.neunet.2021.10.026 PMID:34785445

Liu, R., Zhang, M., Yao, Y., & Yu, F. (2022). A novel high-dimensional multi-objective optimization algorithm for global sorting. *Journal of Jilin University Science Edition*, *60*(3), 664–670.

Liu, Y., Yang, M., Wang, Y., Li, Y., Xiong, T., & Li, A. (2022). Applying machine learning algorithms to predict default probability in the online credit market: Evidence from China. *International Review of Financial Analysis*, *79*, 101971. doi:10.1016/j.irfa.2021.101971

Luo, Q., & Zhang, M. (2022). Research on credit risk assessment of listed companies in science and technology sector by introducing industry research report information. *Procedia Computer Science*, *214*, 1317–1324. doi:10.1016/j.procs.2022.11.311

Mahbobi, M., Kimiagari, S., & Vasudevan, M. (2023). Credit risk classification: An integrated predictive accuracy algorithm using artificial and deep neural networks. *Annals of Operations Research*, *330*(1), 609–637. doi:10.1007/s10479-021-04114-z

Mushava, J., & Murray, M. (2024). Flexible loss functions for binary classification in gradient-boosted decision trees: An application to credit scoring. *Expert Systems with Applications*, *238*, 121876. doi:10.1016/j.eswa.2023.121876

Nana, Z., Xiujian, W., & Zhongqiu, Z. (2022). Game theory analysis on credit risk assessment in e-commerce. *Information Processing & Management*, *59*(1), 102763. doi:10.1016/j.ipm.2021.102763

Qian, H., Wang, B., Ma, P., Peng, L., Gao, S., & Song, Y. (2022). Managing dataset shift by adversarial validation for credit scoring. *Pacific Rim International Conference on Artificial Intelligence.* doi:10.1007/978-3-031-20862-1_35

Rao, C., Liu, Y., & Goh, M. (2023). Credit risk assessment mechanism of personal auto loan based on PSO-XGBoost model. *Complex & Intelligent Systems*, *9*(2), 1391–1414. doi:10.1007/s40747-022-00854-y

Roeder, J., Palmer, M., & Muntermann, J. (2022). Data-driven decision-making in credit risk management: The information value of analyst reports. *Decision Support Systems*, *158*, 113770. doi:10.1016/j.dss.2022.113770

Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: A systemic review. *Neural Computing & Applications*, *34*(17), 14327–14339. doi:10.1007/s00521-022-07472-2

Singh, V., Chen, S.-S., Singhania, M., Nanavati, B., & Gupta, A. (2022). How are reinforcement learning and deep learning algorithms used for big data based decision making in financial industries? A review and research agenda. *International Journal of Information Management Data Insights*, *2*(2), 100094. doi:10.1016/j.jjimei.2022.100094

Talaat, F. M., Aljadani, A., Badawy, M., & Elhosseini, M. (2023). Toward interpretable credit scoring: Integrating explainable artificial intelligence with deep learning for credit card default prediction. *Neural Computing & Applications*, ●●●, 1–19. PMID:37362562

Wang, J., Rong, W., Zhang, Z., & Mei, D. (2022). Credit debt default risk assessment based on the XGBoost algorithm: An empirical study from China. *Wireless Communications and Mobile Computing*, *2022*, 1–14. doi:10.1155/2022/8005493

Wang, K., Li, M., Cheng, J., Zhou, X., & Li, G. (2022). Research on personal credit risk evaluation based on XGBoost. *Procedia Computer Science*, *199*, 1128–1135. doi:10.1016/j.procs.2022.01.143

Wang, T., Liu, R., & Qi, G. (2022). Multi-classification assessment of bank personal credit risk based on multi-source information fusion. *Expert Systems with Applications*, *191*, 116236. doi:10.1016/j.eswa.2021.116236

Ye, S., Yao, K., & Xue, J. (2023). Leveraging empowering leadership to improve employees' improvisational behavior: The role of promotion focus and willingness to take risks. *Psychological Reports*, ●●●, 00332941231172707. doi:10.1177/00332941231172707 PMID:37092876

Ye, S., & Zhao, T. (2023). Team knowledge management: How leaders' expertise recognition influences expertise utilization. *Management Decision*, *61*(1), 77–96. doi:10.1108/MD-09-2021-1166

Yuan, L., Li, H., Fu, S., & Zhang, Z. (2022). Learning behavior evaluation model and teaching strategy innovation by social media network following learning psychology. *Frontiers in Psychology*, *13*, 843428. Advance online publication. doi:10.3389/fpsyg.2022.843428 PMID:35936300

Zakowska, A. (2023). A new credit scoring model to reduce potential predatory lending: A design science approach. *International Conference on Systems Engineering*. doi:10.1007/978-3-031-40579-2_4

Zeng, X., & Zhong, Z. (2022). Multimodal sentiment analysis of online product information based on text mining under the influence of social media. *Journal of Organizational and End User Computing*, *34*(8), 1–18. doi:10.4018/JOEUC.314786

Zhang, L., & Song, Q. (2022). Credit evaluation of SMEs based on GBDT-CNN-LR hybrid integrated model. *Wireless Communications and Mobile Computing*, *2022*, 1–8. doi:10.1155/2022/5251228

Zhao, X., & Chen, H. (2022). Research on influencing factors and transmission mechanisms of green credit risk. *Environmental Science and Pollution Research International*, *29*(59), 89168–89183. doi:10.1007/s11356-022-22041-9 PMID:35849231

Zhong, Z., & Zhao, Y. (2024). Collaborative driving mode of sustainable marketing and supply chain management supported by metaverse technology. *IEEE Transactions on Engineering Management*, *71*, 1642–1654. doi:10.1109/TEM.2023.3337346