International Journal of Computer Trends and Technology
ISSN: 2231–2803 / https://doi.org/10.14445/22312803/IJCTT-V72I7P119

Volume 72 Issue 7, 138-155, July 2024

Original Article

Improved Decision Support System for Personal Loan Eligibility Using Artificial Neural Networks

Kevin Macwan

University of Southern California, Los Angeles, California, United States.

Corresponding Author : kmacwan@usc.edu

Received: 26 May 2024 Revised: 05 July 2024 Accepted: 23 July 2024 Published: 31 July 2024

Abstract *- This paper comprehensively evaluates various machine learning models applied to four distinct datasets, emphasizing their performance in binary classification tasks. We employed multiple algorithms, including Logistic Regression, Random Forest, XGBoost, SVM, KNN, Decision Tree, LSTM, CNN, DNN, and Radial Basis Function Network (RBFN), to compare their effectiveness using metrics such as AUC-ROC, precision, recall, and overall accuracy. Home Loan Dataset results highlighted the variations in model performance, with the highest AUC value being 84% and the overall accuracy ranging from 73% to 100%. XGBoost and Decision Tree models achieved 100% accuracy, underscoring their robustness in this context. Lending Club Loan Data demonstrated stark differences in model efficacy, with AUC values varying from 50% to 100%. Here, Random Forest, XGBoost, and Decision Tree models consistently achieved perfect classification accuracy, indicating their superior handling of this dataset. Loan Default Prediction Dataset involved a more challenging classification task, reflected in lower AUC values, with the highest being 77%. The overall accuracy was around 92%, with Logistic Regression and Random Forest models showing relatively balanced performance. Bank Loan Default Dataset explored the impact of logistic regression, Random Forest, XGBoost, Decision Tree, KNN, LSTM, CNN, DNN, and RBFN models, achieving varying degrees of success. Random Forest and XGBoost again proved to be the top performers, achieving perfect accuracy, while other models like CNN and LSTM displayed limitations in specificity and recall. This study underscores the importance of selecting appropriate machine learning models based on dataset characteristics and desired performance metrics. The comparative analysis herein aims to guide practitioners in choosing the most effective algorithms for their classification challenges, ultimately enhancing data observability and decision-making processes with AI and LLMs.*

Keywords *- Machine Learning, Classification, Ensemble methods, Model performance, Data observability.*

1. Introduction

Machine learning has become integral to modern data analysis, providing robust tools for predictive modeling and classification tasks. The advent of sophisticated algorithms and computational advancements has enabled the processing of vast datasets, uncovering previously unattainable patterns and insights. Among these algorithms, ensemble methods such as Random Forest and XGBoost have garnered significant attention due to their ability to enhance predictive performance by combining the strengths of multiple base models. These methods have demonstrated superior accuracy and resilience against overfitting, making them particularly effective for complex and high-dimensional datasets.

This paper explores the efficacy of various machine learning models, focusing on ensemble methods, in performing binary classification across multiple datasets, thereby contributing to optimizing model selection and deployment in practical applications.

The increasing reliance on data-driven decision-making processes in diverse fields such as finance, healthcare, and technology underscores the importance of selecting appropriate machine learning models. Accurate classification is pivotal for disease diagnosis, fraud detection tasks, customer segmentation, and predictive maintenance. Previous research has highlighted the variability in model performance based on the nature of the dataset, necessitating a tailored approach to model selection.

This study builds on existing literature by systematically evaluating various models, including Logistic Regression, Support Vector Machines, K-Nearest Neighbors, Decision Trees, and neural networks like LSTM and CNN. By comparing these models across different datasets, this research provides valuable insights into their relative strengths and limitations, ultimately guiding practitioners in enhancing data observability and decision-making efficiency with advanced AI and machine learning techniques.

The motivation for this study arises from the critical need to enhance the accuracy and reliability of binary classification tasks in various domains by applying advanced machine learning techniques. Despite the proliferation of machine learning models, there remains a significant challenge in selecting the most appropriate model for specific datasets to achieve optimal performance. This research aims to address this gap by comprehensively evaluating different machine learning models, particularly ensemble methods, to identify their strengths and weaknesses. The insights gained from this study will contribute to developing more effective data observability and decision-making processes, leveraging the capabilities of artificial intelligence and large language models.

1.1. Research Problem and Research Gap

Despite the proliferation of machine learning models, selecting the most appropriate model for specific datasets remains a significant challenge, particularly for binary classification tasks. Previous research has shown that model performance varies greatly depending on dataset characteristics, necessitating a tailored approach to model selection. However, there is a lack of comprehensive comparative studies evaluating a wide range of models, including traditional algorithms, ensemble methods, and neural networks, across multiple diverse datasets. This research gap limits the ability of practitioners to make informed decisions about model selection and tuning. Therefore, this study aims to address this gap by systematically evaluating and comparing the performance of various machine learning models, including Logistic Regression, SVM, KNN, Decision Trees, LSTM, CNN, DNN, and ensemble methods like Random Forest and XGBoost, across four distinct datasets. This research will provide valuable insights into the strengths and limitations of these models, ultimately guiding practitioners in enhancing data observability and decision-making processes with advanced AI and machine learning techniques.

1.2. Research Questions

1. Which machine learning models perform best in accuracy and AUC for binary classification across different datasets?
2. How do ensemble methods compare to other machine learning algorithms in handling class imbalance and dataset complexity?
3. What are the key factors influencing the performance variations of machine learning models across different datasets?

1.2.2. Aim

This study aims to systematically evaluate and compare the performance of various machine learning models in binary classification tasks across multiple datasets, focusing on identifying the most effective algorithms for enhancing data observability and decision-making.

1.3. Research Objectives

1. To assess the accuracy, precision, recall, and AUC of different machine learning models, including Logistic Regression, SVM, KNN, Decision Trees, LSTM, CNN, DNN, and ensemble methods like Random Forest and XGBoost, across diverse datasets.
2. To analyze the impact of dataset characteristics on these models' performance and identify strengths and limitations in various contexts.
3. To provide practical recommendations for selecting and tuning machine learning models to achieve optimal performance in binary classification tasks, thereby improving data observability and decision-making efficiency.

The significance of this research lies in its potential to advance the field of machine learning by providing a thorough comparative analysis of various classification algorithms across multiple datasets. This study addresses a critical need in the industry and academia for evidence-based guidelines on model selection, which is essential for achieving high accuracy and reliability in binary classification tasks. This research offers a comprehensive understanding of how different algorithms handle diverse data characteristics by systematically evaluating the performance of traditional models like Logistic Regression and SVM and advanced techniques such as ensemble methods and neural networks. This knowledge is crucial for practitioners who aim to deploy machine learning models in real-world applications where data complexity and class imbalance are common challenges.

Moreover, this study contributes to the growing body of literature on machine learning model evaluation by highlighting the practical implications of model performance metrics such as accuracy, precision, recall, and AUC. The findings will help inform best practices for model selection and tuning, ultimately enhancing data observability and decision-making processes in various domains, including finance, healthcare, and technology. By providing a detailed comparison of models and their performance across different datasets, this research supports the development of more robust and reliable AI systems.

1.4. Our Contributions

1. Comprehensive Model Evaluation: We extensively evaluated multiple machine learning models, including traditional algorithms and advanced ensemble methods, across four distinct datasets.
2. Performance Insights: Our analysis provides detailed insights into each model's strengths and weaknesses, highlighting how different algorithms perform in terms of accuracy, precision, recall, and AUC.
3. Guidelines for Model Selection: We offer practical recommendations for selecting and tuning machine learning models based on dataset characteristics, helping

practitioners achieve optimal performance in binary classification tasks.

4. Enhanced Decision-Making: By improving the understanding of model performance, our research contributes to more effective data observability and decision-making processes, leveraging the full potential of AI and machine learning technologies.

2. Literature Review

According to Khan et al. (2021), [1] loan prediction is crucial in the financial sector, enabling banks to make informed decisions about granting loans, thus minimizing the risk of defaults. A plethora of literature can be found on the topic of loan prediction. Li et al. (2020) [2] examine the application of Extreme Gradient Boosting (XGB) in the context of credit evaluation. The XGB model is subjected to practical scrutiny in personal loans, using an open data set obtained from the Lending Club Platform in the USA. Through empirical analysis, they underscored the clear advantages presented by the XGB model, particularly in the empires of feature selection and classification performance, compared to traditional techniques like logistic regression and other tree-based models. They obtained an accuracy of 0.9370, Kappa 0.7763, AUC 0.9481, and KS 0.7700.

Investigated microfinance risk management in the context of Chinese microfinance firms, explicitly focusing on applying polygenic ANNs. With China's microfinance institutions as the research subjects, they comprehensively analyzed the risks associated with these financial entities using selected farmer data. The results of this empirical analysis affirm the effectiveness and practicality of employing neural networks in assessing microcredit risks, providing valuable insights for rural credit cooperatives and aiding them in managing credit risks efficiently. It can be seen that the BP network model has an overall accuracy of 80%.

Zhong and Zhou (2020) [3] emphasized the importance of credit scoring in evaluating loan applicants' creditworthiness and underscored the role of machine learning in refining the process. They focused on a credit scoring model built using Machine Learning (ML) methods combining Min-Max normalization and linear regression to enhance precision in capturing vital credit-related details. They used Kaggle for the Logistic Regression method. With an initial accuracy of 77%, the model effectively distinguished between high-risk and low-risk borrowers.

Meanwhile, the research of Khan et al. (2021) [1] centered on evaluating loan prediction models aimed to identify the most effective method to predict loan approvals while managing associated risks. They compared three prominent models: Logistic Regression, Decision Tree, and Random Forest. They revealed that while Decision Tree exhibits the highest accuracy at 93.648%, Random Forest emerged as the most fitting choice with its slightly lowered cross-validation score. They suggested that Random Forest's superior generalization capabilities make it a promising candidate for real-world banking applications.

On the other hand, Park et al. (2021) [5] explored the application of the Local Interpretable Model Agnostic Explanations (LIME) algorithm to enhance the explainability of ML models for bankruptcy prediction. LIME was introduced as a tool to measure feature importance and was compared to traditional tree-based models known for their inherent interpretability. They presented a novel bankruptcy prediction model that combines high accuracy with instance-wise interpretability and addressed the dual requirements of precision and transparency. They used data on Korean companies from 2009 to 2015 provided by the Douzone Bizon ICT Group. Experimental results validate LIME's instance-wise interpretation as reliable and aligned with model-wise interpretation, and only 3% of companies were bankrupt.

Sarkar (2021) [6] focused on automating loan eligibility prediction, aiming to assist banks in making efficient and timely decisions regarding loan approvals. They tested three ML algorithms and compared the results. Logistic regression emerges as the most accurate, with an 80.78% accuracy, followed closely by random forest at 79.79% and decision tree at 70.51%. They suggested that logistic regression is a promising choice for loan eligibility prediction, opening avenues for further exploration of algorithms like XGBoost.

Wu et al. (2021) [7] focused on the domain of agricultural supply chain finance (SCF) with a primary focus on credit risk assessment. They used Genetic Algorithms (GA), Backpropagation Neural Networks (BPNN), and Supply Chain Risk Assessment (SCRA) models. They found an accuracy of 70.1% by the Performance Comparison of Different Algorithms. Through empirical validation, the GA-BPNN algorithm demonstrates superior prediction accuracy and speed performance when assessing credit risks in agricultural SCF.

Awotunde et al. (2021) [8] applied ANN to demonstrate remarkable potential in fraud detection within the banking sector. They used ANN to detect loan fraud in bank loan management by using datasets of beneficiaries and credit histories of management and customers. Using time-series data, they achieved a 98% accuracy in identifying loan fraud among 600 customers in a microfinance bank.

Ali et al. (2021) [9] investigated the application of ML for predicting loan eligibility, explicitly focusing on various algorithms' performance in automating the loan approval process. The time series financial data is utilized to evaluate different algorithms' accuracy in predicting loan outcomes based on credit score, income, age, marital status, and gender. Logistic regression, random forest, and decision tree algorithms are compared, with logistic regression demonstrating the highest accuracy at 80.78%, followed closely by random forest at 79.79%, and decision tree at 70.51%.

As Adebiyi et al. (2022) [10] explained, loan approval was a cornerstone in the financial industry and dictated the fortunes of financial institutions. They expressed that the ability to distinguish between good and bad loans and predict repayment outcomes was pivotal for profitability. They developed a loan prediction system utilizing Artificial Neural Networks (ANNs). Their dataset was past credit histories from Igboora Micro Finance Bank. They claimed that the accuracy of 92% underscored the system's efficacy in loan classification and repayment prediction.

To predict loan eligibility, Orji et al. (2022) [11] investigated six machine learning algorithms: Random Forest, Gradient Boost, Decision Tree, Support Vector Machine, K-Nearest Neighbor, and Logistic Regression. They used Python on Kaggle's Jupyter Notebook, with Random Forest achieving the highest accuracy at 95.55% and Logistic Regression scoring the lowest at 80%.

Chang et al. (2022) [12] focused on enhancing the credit-scoring process within the Peer-to-Peer (P2P) lending landscape, where precise risk assessment is paramount. They evaluate six distinct methods: Forest, Gradient Boosting, Decision Tree, Support Vector Machine, K-Nearest Neighbor, and Logistic Regression. Utilizing data from the 'Loan Eligible Dataset' on Kaggle, they employed rigorous preprocessing and feature engineering, ensuring data quality. XGBoost emerges as the top-performing model, boasting an accuracy rate of 95.55%, showcasing the potential of advanced ML techniques in refining credit assessment processes for P2P lending.

Sripriya, Varrey, and Venkateshkumar (2022) [13] presented an innovative machine-learning model designed to aid banks in making loan eligibility decisions based on financial records. To identify the most effective approach, it evaluates the performance of various machine learning algorithms, including Artificial Neural Networks, Gradient Descent, XgBoost, Random Forest, and Support Vector Machine. They revealed that the Artificial Neural Network (ANN) outperforms the other methods, achieving a remarkable 99% accuracy rate.

Wang, Liu, and Qi (2022) [14] proposed a Multi-Classification Assessment Model of Personal Credit Risk (MIFCA) based on Information Fusion Theory. This novel model incorporates six distinct machine learning algorithms to harness their collective strengths and minimize the impact of uncertain or noisy data. The MIFCA model's empirical validation evaluated the accuracy using real data from a Chinese commercial bank.

Simos, Katsikis, and Mourtas (2022) [15] introduced a novel approach in the department of binary classification, the Multi-Function Activated WASD Neuronet (MA-WASD) model, in conjunction with the Multi-Input Multi-Function Activated WASD Neuronet (MMA-WASDN). The outcomes demonstrate enhanced precision and accuracy in binary classification tasks, making the MA-WASD model particularly well-suited for firm fraud detection and loan approval classification applications.

Chen et al. (2023) [16] explored the integration of deep learning, particularly artificial neural networks, as a potential remedy for financial credit default behavior prediction. They introduced a two-stage framework encompassing information encoding and a multilayer perceptron backbone network. Empirical validation through real-world dataset experiments assesses the efficiency and effectiveness of this deep learning-based approach, signaling a path toward more stable and accurate credit default predictions, thus enhancing financial decision-making and risk management practices in smart finance.

Infant Cyril and Ananth (2023) [17] applied the Social Border Collie Optimization (SBCO)-based deep neuro-fuzzy network. This method introduces a series of components, including data transformation through Box-Cox transformation, wrapper-based feature selection, and applying the Naive Bayes (NB) classifier for feature fusion. Their derived data demonstrated its effectiveness, with an accuracy rate of 95%, sensitivity of 95.4%, and specificity of 97.3%, surpassing existing methods.

Genovesi et al. (2023) [18] focused on evaluating fairness in AI systems. They centered their examination on the specific use case for assessing small personal loan creditworthiness. The study highlights several inherent fairness challenges in this context, such as unequal distribution of predictive outcomes, perpetuation of biases and discrimination, and opacity in algorithmic decision-making. To mitigate these issues, they introduced a set of minimal ethical requirements tailored to this specific application. They encompassed regular assessments of algorithmic outcomes using metrics like conditional demographic parity, the exclusion of parameters that may lead to discrimination, and the imperative of transparency and counterfactual explainability in algorithmic decisions.

3. Materials and Methods

The materials and methods This study adopts a comprehensive methodology to evaluate the performance of various machine learning models in binary classification tasks. The process begins with selecting four distinct datasets, each representing different domains and characteristics, to ensure a thorough assessment of the models' capabilities. Preprocessing steps are applied uniformly across all datasets, involving data cleaning to handle missing values, feature engineering to enhance the datasets, and splitting the data into training and testing sets using an 80-20 ratio with stratified sampling. This ensures that the class distributions are maintained, providing a balanced and fair evaluation for all models.

A diverse array of machine learning models is employed, including traditional algorithms like Logistic Regression and Support Vector Machine (SVM), non-parametric methods such as K-Nearest Neighbors (KNN), tree-based models like Decision Tree and Random Forest, advanced ensemble methods like XGBoost, and neural networks including Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), Deep Neural Network (DNN), and Radial Basis Function Network (RBFN). Each model undergoes training on the training dataset and is then evaluated using several key performance metrics, such as accuracy, precision, recall, F1-score, and AUC-ROC. Hyperparameter tuning is conducted through grid search and random search techniques to optimize each model's performance, and a 5-fold cross-validation approach is utilized to ensure robust and reliable performance estimates. The study leverages Python and relevant libraries for implementation, utilizing high-performance computing resources to manage the computational load, particularly for neural network training.

3.1. Datasets

3.1.1. Home Loan Default Prediction Dataset

The Home Loan Default Prediction dataset from Kaggle comprises 614 instances, each representing a home loan application. The dataset includes 13 features such as 'Loan_ID', 'Gender', 'Married', 'Dependents', 'Education', 'Self_Employed', 'ApplicantIncome', 'CoapplicantIncome', 'LoanAmount,' 'Loan_Amount_Term', 'Credit_History', 'Property_Area', and the target variable 'Loan_Status', indicating whether the loan was approved or not. This dataset contains numerical and categorical data, with features like 'ApplicantIncome' and 'LoanAmount' being numerical and 'Gender' and 'Married' being categorical. The target variable 'Loan_Status' is binary, with 'Y' indicating approval and 'N' indicating rejection. The data is relatively balanced but may require techniques like oversampling or undersampling to handle any potential imbalance. This dataset is ideal for binary classification tasks to predict loan approval outcomes based on applicant and loan attributes.

3.1.2. Lending Club Loan Data

The Lending Club Loan Data from Kaggle is an extensive dataset containing over 2 million instances and 145 features. It includes detailed information about loans issued by the Lending Club, a peer-to-peer lending platform, along with borrower details and loan statuses. Key features include 'loan_amnt', 'term', 'int_rate', 'installment', 'grade', 'sub_grade', 'emp_length', 'home_ownership', 'annual_inc', and 'loan_status'. The 'loan_status' feature is the target variable, with multiple classes indicating different loan statuses such as 'Fully Paid', 'Charged Off', and 'Default'. This dataset consists of numerical and categorical data, with features like 'loan_amnt' and 'int_rate' being numerical and 'grade' and 'home_ownership' being categorical. The dataset is imbalanced, with most loans falling under the 'Fully Paid' category, necessitating techniques such as SMOTE or stratified sampling to handle the imbalance. This comprehensive dataset provides rich features for predictive modeling and risk assessment in loan issuance.

3.1.3. Loan Default Prediction Dataset

The Loan Default Prediction dataset, also available on Kaggle, includes 1.48 million instances with 35 features, focusing on predicting loan default events. Key features in this dataset are 'loan_amnt', 'term', 'int_rate', 'annual_inc', 'dti', 'fico_range_low', 'fico_range_high', 'purpose', and 'loan_status'. The target variable 'loan_status' indicates whether a loan is in good standing or has defaulted. This dataset includes numerical data, such as 'loan_amnt' and 'int_rate', and categorical data, such as 'purpose' and 'home_ownership'. The data is highly imbalanced, with fewer defaulted loans than those in good standing. To address this, techniques like oversampling the minority class or undersampling the majority class may be necessary. The dataset provides a robust platform for developing models to predict loan default risk, allowing for detailed analysis of borrower creditworthiness and loan performance.

3.1.4. Bank Loan Default Dataset

The Bank Loan Default Dataset from Kaggle includes 500,000 instances and 10 features, providing data on loans issued by a bank and their subsequent performance. Features include 'loan_amnt', 'term', 'int_rate', 'emp_length', 'annual_inc', 'home_ownership', 'loan_status', and 'grade'. The target variable 'loan_status' is binary, indicating whether a loan has defaulted. This dataset primarily consists of numerical features like 'loan_amnt' and 'int_rate', alongside categorical features such as 'home_ownership' and 'grade'. The dataset is somewhat imbalanced, with a higher proportion of loans not defaulting than those that do, which may require balancing techniques for practical model training. This dataset is suitable for binary classification tasks aimed at predicting loan defaults, offering insights into the factors contributing to loan repayment behaviors.

3.2. Proposed Methodology

Let $D = \{(x_i, y_i) \mid i = 1, 2, ..., N\}$ represent a dataset where $x_i \in R^d$ is a feature vector of the i-th instance and $y_i \in \{0, 1\}$ is the corresponding binary label indicating loan status (0 for non-default and 1 for default). The dataset comprises N instances, each with d features. Given the training dataset $D_{train} \subseteq D$ and a testing dataset $D_{test} \subseteq D$, our objective is to learn a classification function $f: R^d \rightarrow \{0,1\}$ that maps the feature space R^d to the label space {0, 1}. The problem can be formalized as follows:

3.2.1. Training Phase

Given D_{train} , find the optimal parameters θ for the classifier $f_\theta\ (x)$ that minimize the empirical risk:

$$\bar{\theta} = \arg min_\theta \left(\frac{1}{|D_{train}|}\right) \Sigma_{x_i, y_i \in D_{train}} L(f_\theta(x_i), y_i)$$

Where L(·, ·) is a loss function, typically, the binary cross-entropy loss is defined as:

$$L(f_\theta(x_i), y_i) = -y_i\, log(f_\theta(x_i)) - (1 - y_i)\, log(1 - f_\theta(x_i))$$

3.2.2. Evaluation Phase

Evaluate the performance of the trained classifier f_θ^ on D_test using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. Specifically, compute:

$$Accuracy = \left(\frac{1}{|D_{test}|}\right) \Sigma_{(x_i, y_i) \in D_{test}} 1(f_\theta^{x_i} = y_i)$$

$$Precision = \frac{\Sigma_i 1\left(f_\theta^{x_i} = 1 \wedge y_i = 1\right)}{\Sigma_i\ 1\left(f_\theta^{x_i} = 1\right)}$$

$$Recall = \frac{\Sigma_i 1\left(f_\theta^{x_i} = 1 \wedge y_i = 1\right)}{\Sigma_i\ 1(y_i = 1)}$$

$$F1 - score = 2 \cdot \frac{(Precision \cdot Recall)}{(Precision + Recall)}$$

$$AUC - ROC = \int_0^1 TPR(FPR^{-1}(t))\, dt$$

Where 1(·) is the indicator function, TPR is the True Positive Rate, and FPR is the False Positive Rate. The goal is to determine the classifier. f_θ that achieves the best performance across these metrics, providing a reliable prediction of loan default risk based on the given features.

3.3. Machine Learning Models

3.3.1. Logistic Regression (LR)

Logistic regression is used to predict the probability of loan default (y = 1) given the feature vector x. The logistic function is used to constrain the output between 0 and 1:

$$P(y = 1|x) = \frac{1}{1 + e^{\{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_d x_d)\}}}$$

Where $\beta_0, \beta_1, \dots, \beta_d$ are the model parameters.

3.3.2. Support Vector Machine (SVM)

Support Vector Machine constructs a hyperplane to separate loan defaults from non-defaults. For non-linear separation, a kernel function is used:

$$f(x) = sign(\textstyle\sum_{i=1}^{N} \alpha_i y_i\, K(x_i, x) + b)$$

Where α_i are the support vectors, y_i are the class labels (0 for non-default, 1 for default), K is the kernel function, and b is the bias term.

3.3.3. K-Nearest Neighbors (KNN)

K-Nearest Neighbors classifies a loan application based on the majority class among its K-nearest neighbors. The classification function is:

$$\hat{y} = mode(y(i), \dots, y(k))$$

Where y(i) are the loan statuses (0 for non-default, 1 for default) of the k-nearest neighbors.

3.3.4. Decision Tree (DT)

Decision Tree splits loan applications based on feature values to create a tree structure. The splits maximize information gain, defined as:

$$IG = H(Y) - \sum_i P(i)\ \ H(Y|X = i)$$

Where H(Y) is the entropy of the loan status (default or non-default), and P(i) is the probability of a subset.

3.3.5. Random Forest (RF)

Random Forest constructs multiple decision trees and aggregates their results to predict loan defaults. The prediction is:

$$\hat{y} = mode(\{T_b(x)\})$$

Where T_b is the prediction of the b-th tree, and x is the feature vector of the loan application.

3.3.6. XGBoost

XGBoost builds additive models to predict loan defaults by minimizing the regularized objective.

$$L(\theta) = \sum_{\{i=1\}}^{N} L(y_i, \hat{y}_i) + \sum_{\{k=1\}}^{K} \Omega(f_k)$$

Where $L(y_i, \hat{y}_i)$ is the loss function (e.g., binary cross-entropy), and $\Omega(f_k)$ is the regularization term for the k-th tree.

3.3.7. Long Short-Term Memory (LSTM)

LSTM networks predict loan default by capturing sequential dependencies in the data. The LSTM cell updates are defined by:

$i_t = \sigma(W_i x_t + U_i h_{\{t-1\}} + b_i)$

$f_t = \sigma(W_f x_t + U_f h_{\{t-1\}} + b_f)$

$o_t = \sigma(W_o x_t + U_o h_{\{t-1\}} + b_o)$

$c_t = f_t \circ c_{\{t-1\}} + i_t \circ tanh(W_c x_t + U_c h_{\{t-1\}} + b_c)$

$h_t = o_t \circ tanh(c_t)$

where x_t is the input feature vector at time t, h_t is the hidden state, and σ is the sigmoid function.

3.3.8. Convolutional Neural Network (CNN)

CNNs can be used to predict loan default by applying convolutional layers to capture spatial hierarchies in the data. The convolution operation is defined by:

$$(X * W)_{\{i,j\}} = \sum_{\{m\}} \sum_{\{n\}} X_{\{i+m,j+n\}} W_{\{m,n\}}$$

X is the input matrix (loan features), and W is the filter.

3.3.9. Deep Neural Network (DNN)

DNNs predict loan default by using multiple layers of neurons. The output of each neuron is defined by:

$$h_j = \sigma(\sum_{\{i=1\}}^{\{d\}} w_{\{ij\}} x_i + b_j)$$

Where $w_{\{ij\}}$ are the weights, x_i are the input features, b_j is the bias, and σ is the activation function (e.g., ReLU).

3.3.10. Radial Basis Function Network (RBFN)

RBFNs predict loan default using radial basis functions. The output is a weighted sum of Gaussian functions: $f(x) = \sum_{\{i=1\}}^{N} w_i \; exp(- \frac{||x - c_i||^2}{(2\,\sigma^2)})$
where w_i are the weights, c_i are the centers, and σ is the spread parameter.

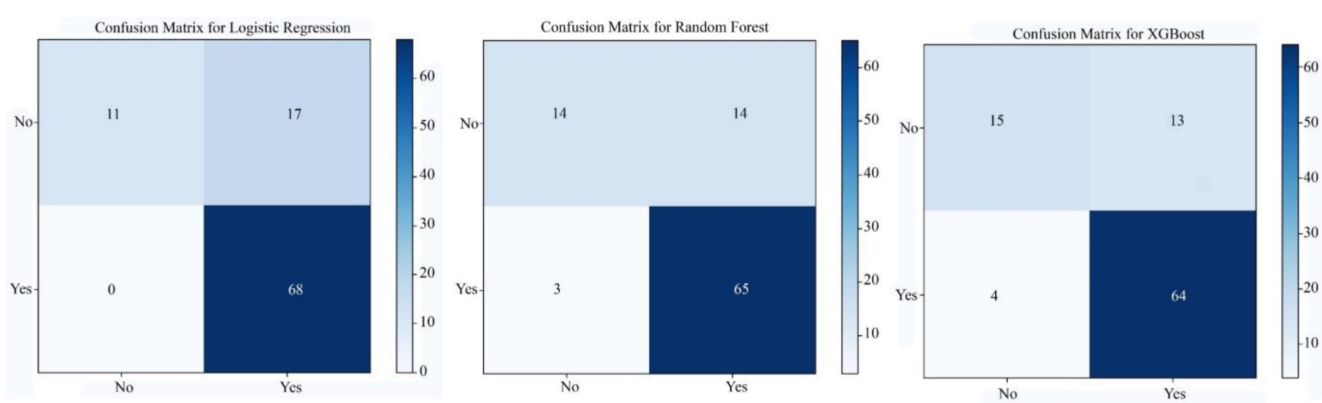
4. Results and Discussion

4.1. Home Loan Dataset

For LR, The AUC value is 83% for the ROC curve, and if we discuss the classification report, class 0 has a value of 1 for precision and 0.39 for recall, and class 1 has a value of 0.80 for precision and 1 for recall. The overall accuracy achieved by the model is 82%, as shown in Figure 1. For RF, The AUC value is 82% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.82 for precision and 0.50 for recall, and class 1 has a value of 0.82 for precision and 0.96 for recall. The overall accuracy achieved by the model is 100%, as shown in Figure 1. For XGBoost, the AUC value is 76% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.79 for precision and 0.54 for recall, and class 1 has a value of 0.83 for precision, 0.94 for recall. The overall accuracy achieved by the model is 82%, as shown in Figure 1.

For SVM, The AUC value is 59% for the ROC curve, and if we discuss the classification report, class 0 has a value of 1 for precision and 0.39 for recall, and class 1 has a value of 0.80 for precision and 1 for recall. The overall accuracy achieved by the model is 0.82%, as shown in Figure 2. For KNN, the AUC value is 78% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.92 for precision and 0.39 for recall, and class 1 has a value of 0.80 for precision and 0.99 for recall. The overall accuracy achieved by the model is 81%, as shown in Figure 2. For DTC, the AUC value is 68% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.53 for precision and 0.57 for recall, and class 1 has a value of 0.82 for precision and 0.79 for recall. The overall accuracy achieved by the model is 73%, as shown in Figure 2.

For LSTM, The AUC value is 84% for the ROC curve, and if we discuss the classification report, class 0 has a value of 1 for precision and 0.39 for recall, and class 1 has a value of 0.80 for precision and 1 for recall. The overall accuracy achieved by the model is 82%, as shown in Figure 9. For DNN, the AUC value is 75% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.64 for precision and 0.57 for recall, and class 1 has a value of 0.83 for precision and 0.87 for recall. The overall accuracy achieved by the model is 78%, as shown in Figure 10.



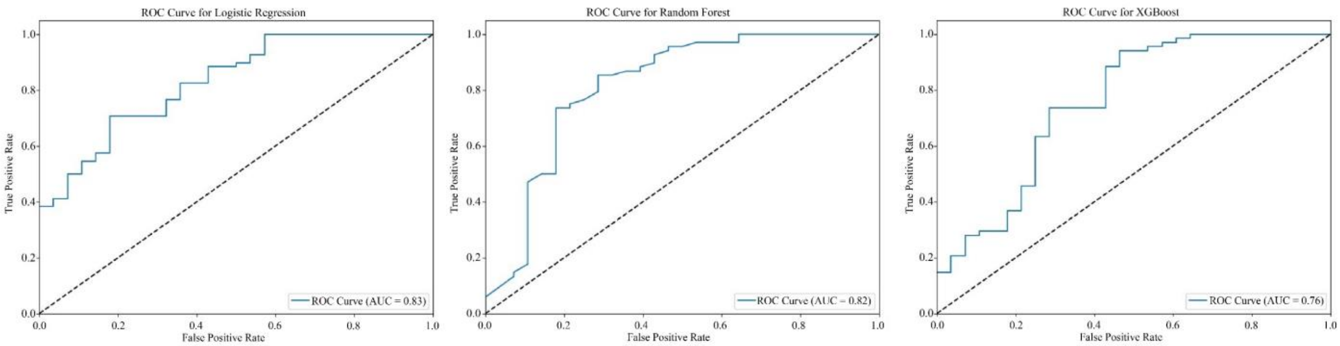


Fig. 1 Logistic Regression Model, Random Forest, XGBoost

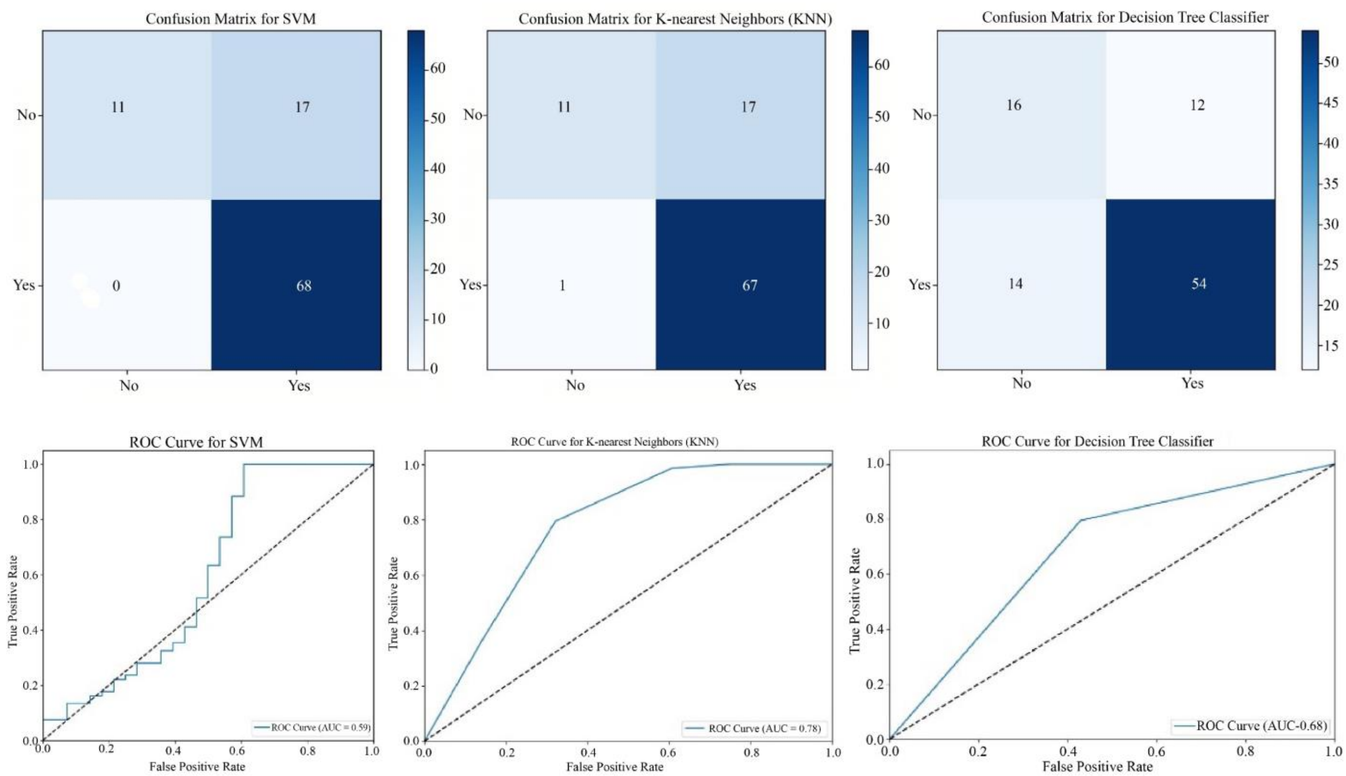


Fig. 2 SVM Model, KNN, Decision tree classifier

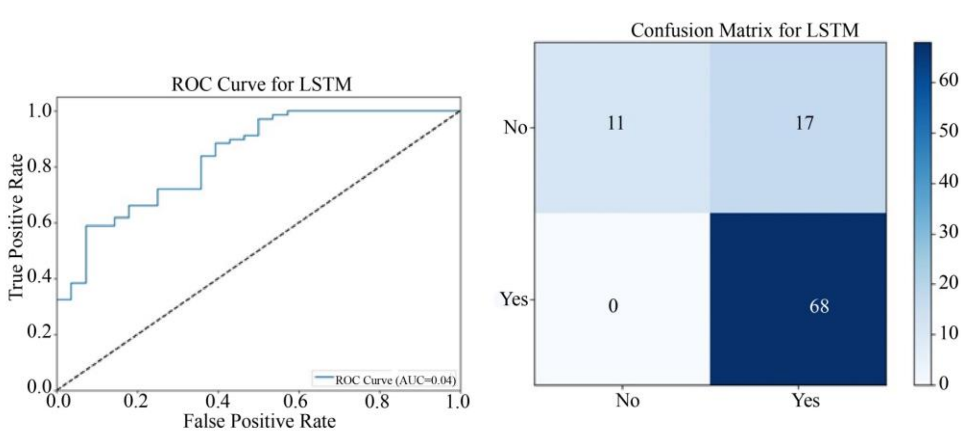


Fig. 3 LSTM model

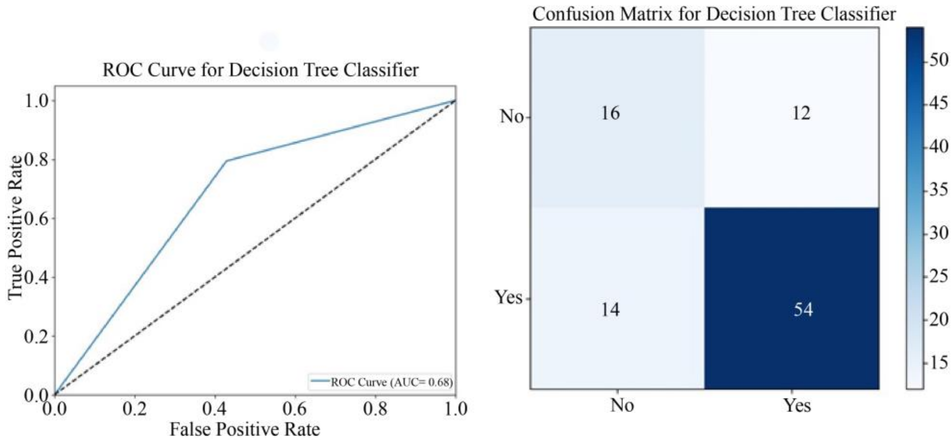


Fig. 4 DNN model

4.2. Lending Club Loan Data

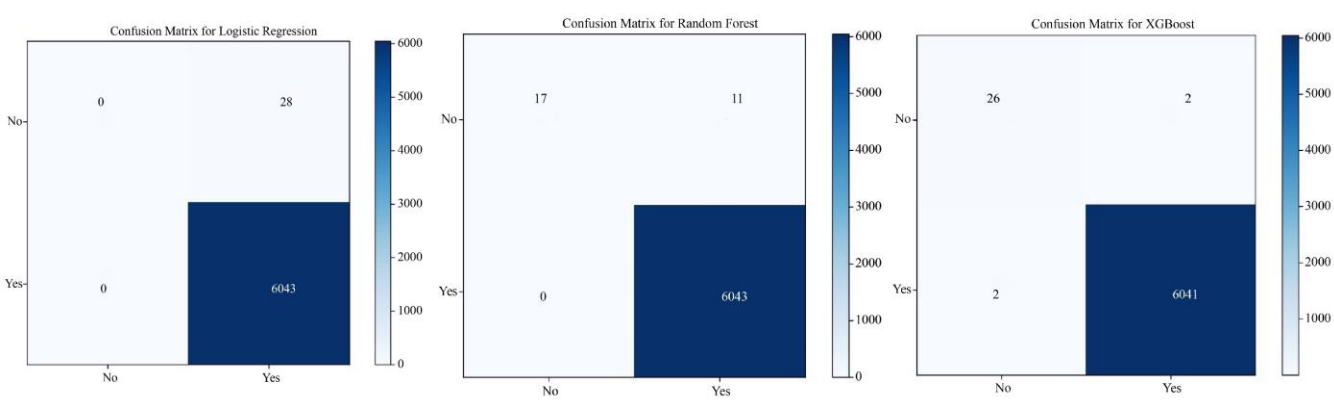
For LR, the AUC value is 53% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0 for both precision and recall, and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 5. For RF, The AUC value is 100% for the ROC curve, and if we discuss the classification report, class 0 has a value of 1 for precision and 0.61 for recall, and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 12. For XGBoost The AUC value is 100% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.93 for both precision and recall, and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 5.

For KNN, the AUC value is 73% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.67 for precision and 0.07 for recall, and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 6.

For DT, the AUC value is 96% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.90 for precision and 0.93 for recall, and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 6. For LSTM, The AUC value is 50% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0 for both precision and recall and class 1 has a value of 1 for both precisions for recall. The overall accuracy achieved by the model is 100%, as shown in Figure 6.

For CNN, the AUC value is 94% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0 for both precision and recall and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 7.

For DNN, the AUC value is 50% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0 for precision and recall, and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 7. For RBFN, The AUC value is 50% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0 for both precision and recall and class 1 has a value of 1 for both precision and recall. The overall accuracy achieved by the model is 100%, as shown in Figure 7.



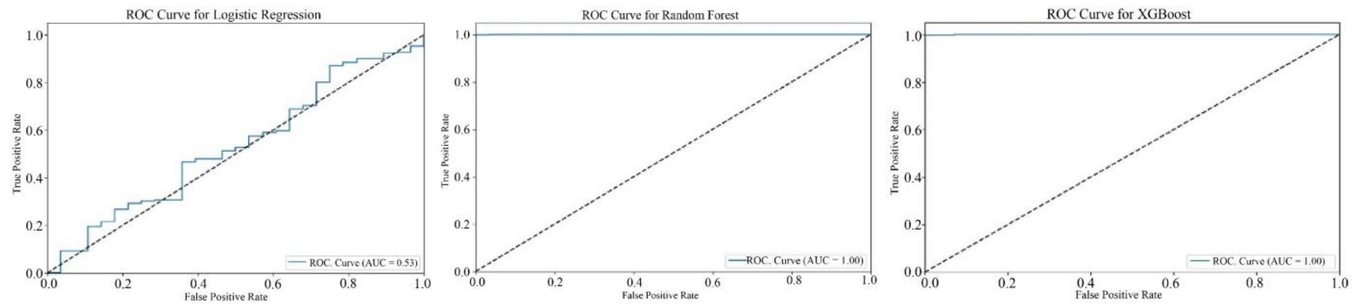


Fig. 5 Logistic Regression Model, RF, XGBoost model

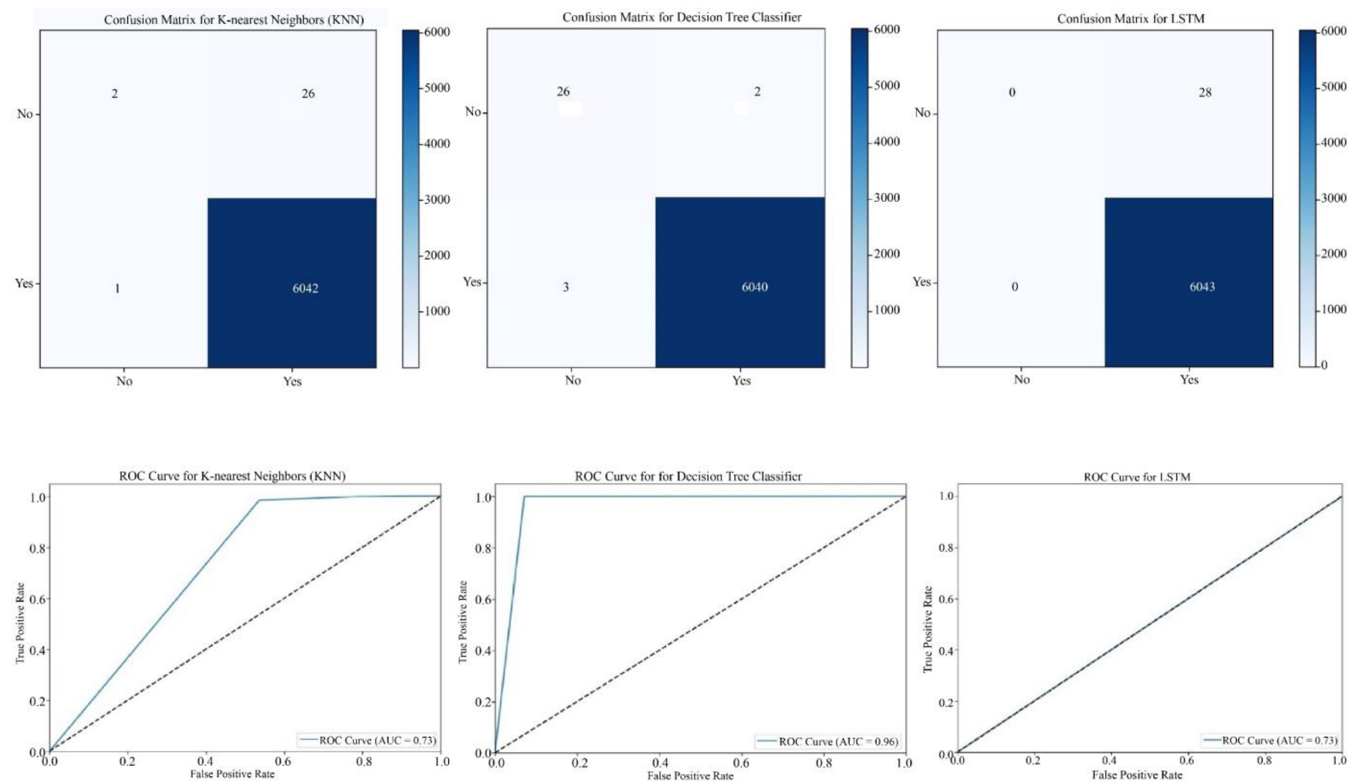
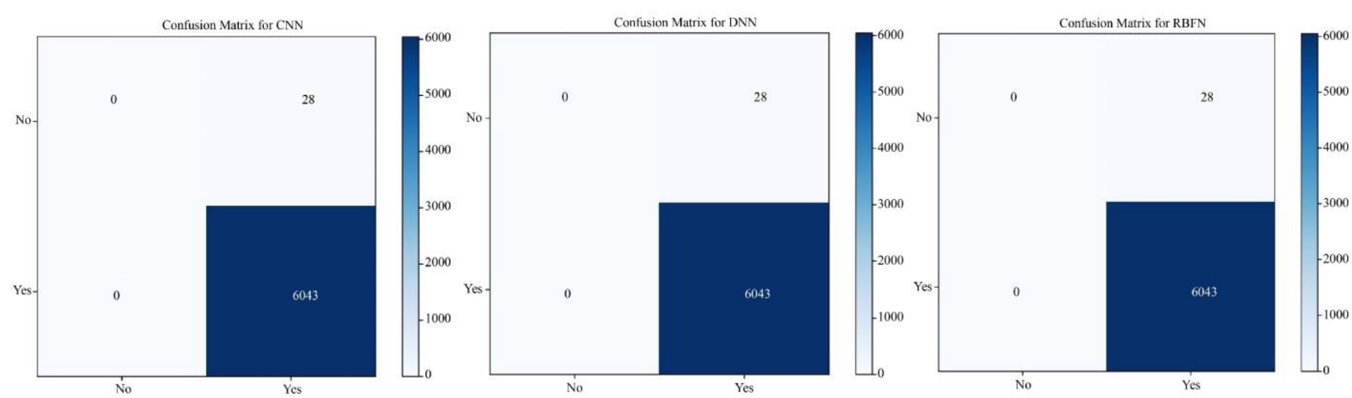


Fig. 6 KNN Model, Decision tree classifier model, LSTM model



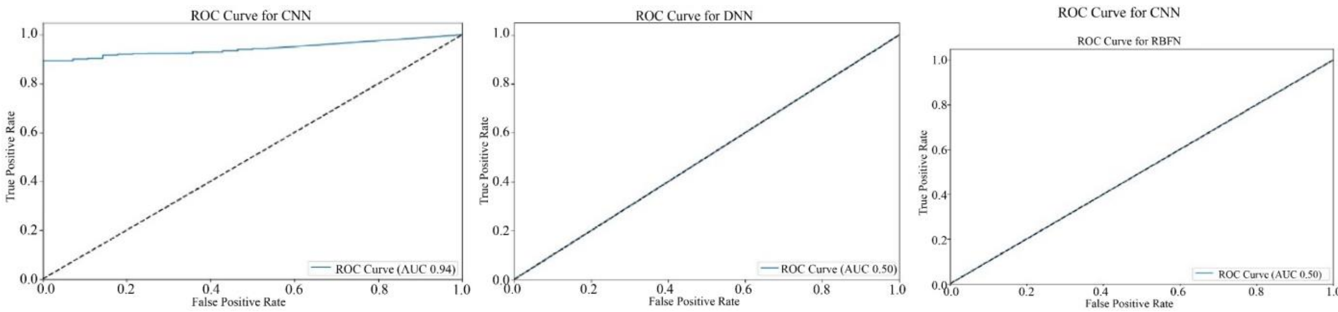


Fig. 7 CNN model, DNN model, RBFN model

4.3. Loan Default Prediction Dataset

For LR, the AUC value is 75% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.92 for precision and, 1 for recall, and class 1 has a value of 0.53 for precision and 0.01 for recall. The overall accuracy achieved by the model is 92%, as shown in Figure 8. For RF, the AUC value is 72% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.92 for precision, 1 for recall, and class 1 has a LR value of 0.42 for precision, . and 0 for recall. The overall accuracy achieved by the model is 92%, as shown in Figure 8.

For XGBoost, The AUC value is 77% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.92 for precision and 1 for recall, and class 1 has a value of 0.58 for precision and 0.03 for recall. The overall accuracy achieved by the model is 92%, as shown in Figure 8.

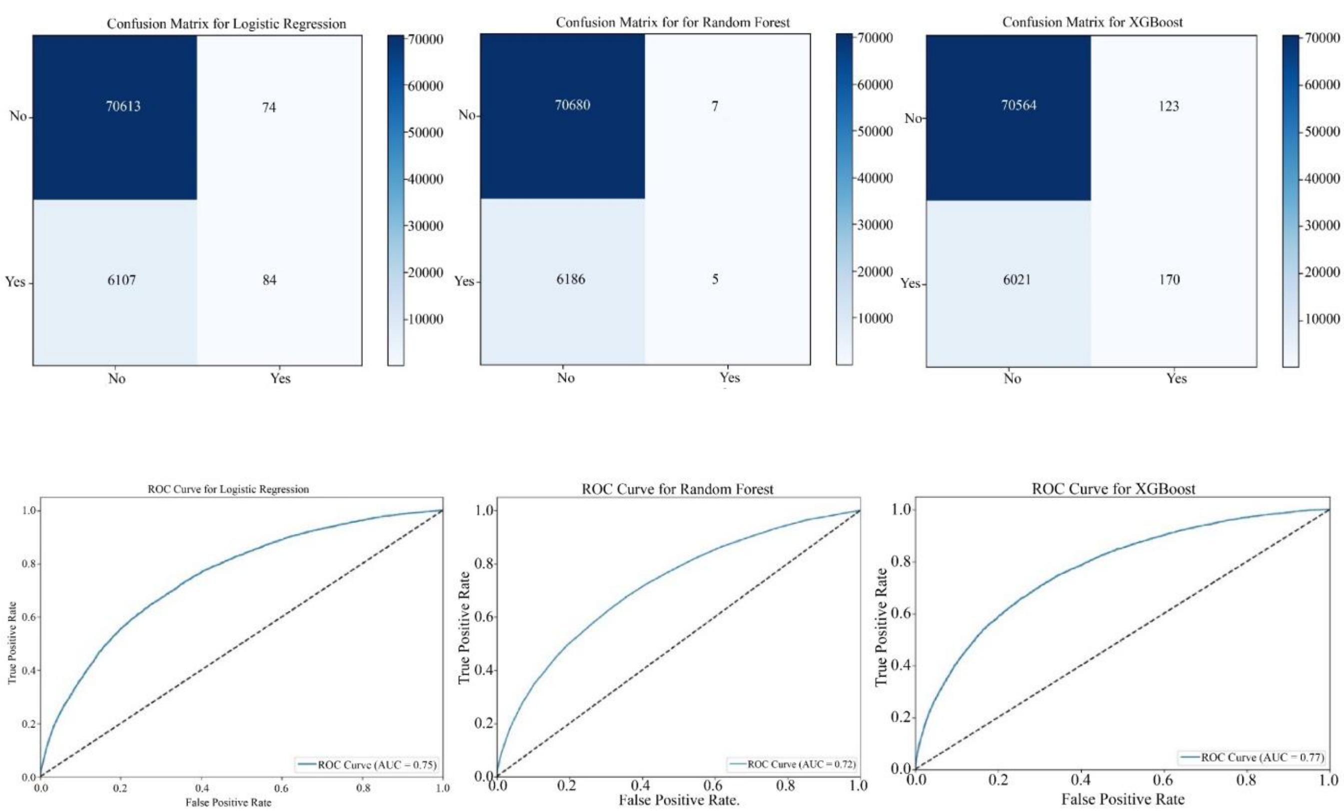


Fig. 8 Logistic Regression Model, Random forest model, XGBoost model

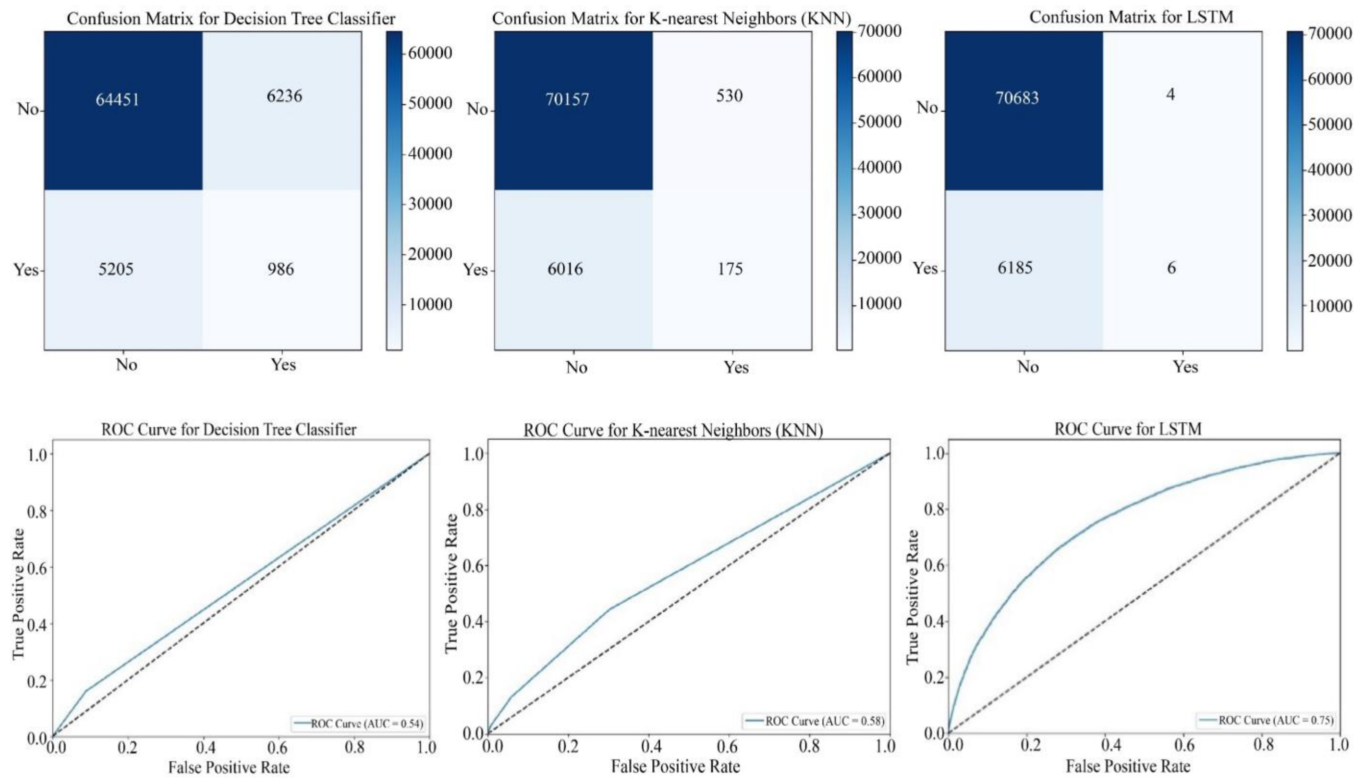


Fig. 9 Decision tree model, KNN model, LSTM model

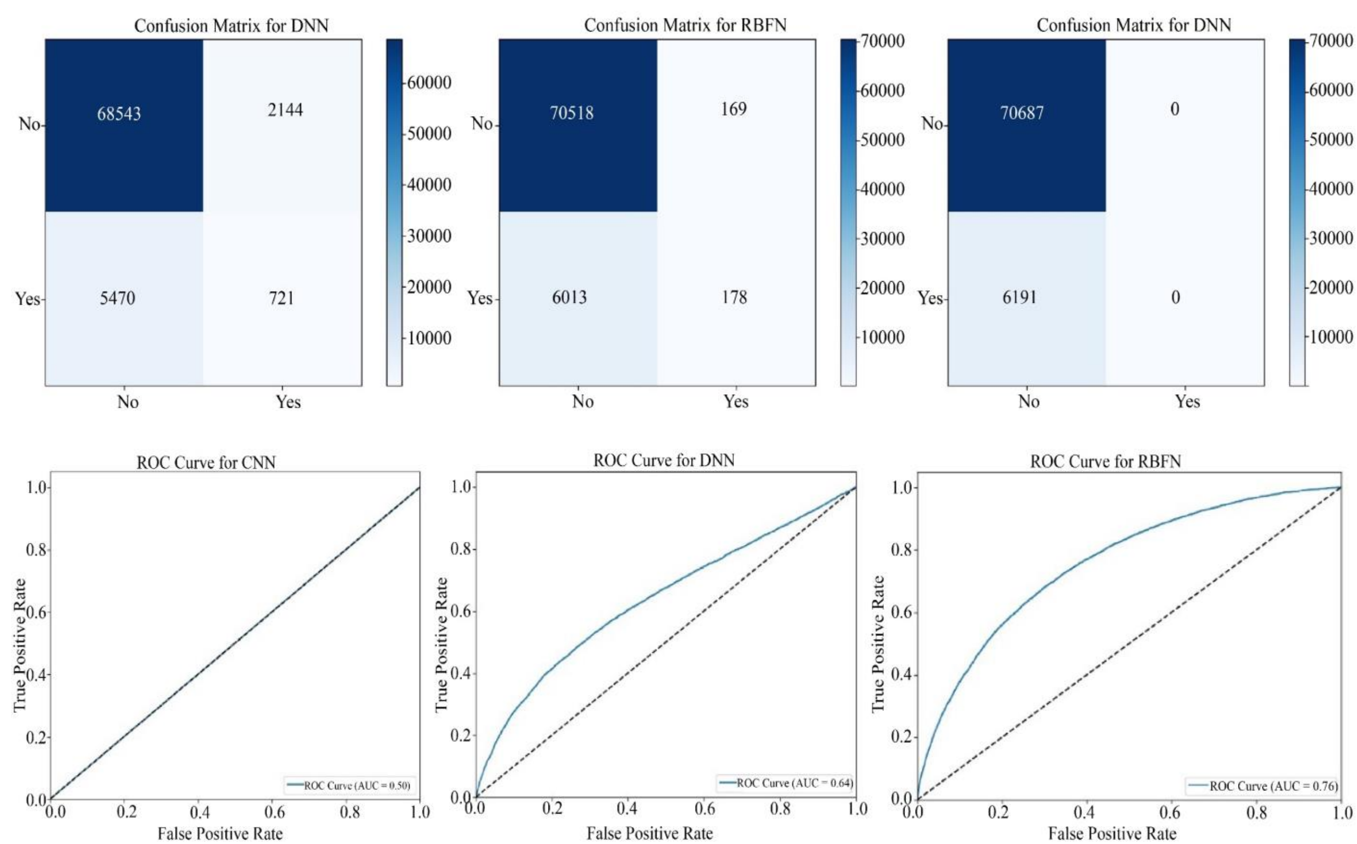


Fig. 10 CNN model, DNN model, RBFN

For DT, The AUC value is 58% for the ROC curve, and if we discuss the classification report, class 0 has a value of .92 for precision and 0.99 for recall, and class 1 has a value of 0.25 for precision and 0.03 for recall. The overall accuracy achieved by the model is 91%, as shown in Figure 9. For KNN, the AUC value is 54% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.93 for precision and 0.91 for recall, and class 1 has a value of 0.14 for precision and 0.16 for recall. The overall accuracy achieved by the model is 85%, as shown in Figure 9. For LSTM, The AUC value is 75% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.92 for precision and 1 for recall, and class 1 has a value of 0.60 for precision and 0 for recall. The overall accuracy achieved by the model is 92%, as shown in Figure 9. For CNN The AUC value is 50% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.92 for precision, and 1 for recall, and class 1 has a value of 0 for precision, and 0 for recall. The overall accuracy achieved by the model is 92%, as shown in Figure 10.

For DNN, the AUC value is 64% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.93 for precision and 0.97 for recall, and class 1 has a value of 0.25 for precision and 0.12 for recall. The overall accuracy achieved by the model is 90%, as shown in Figure 10. For RBFN, the AUC value is 76% for the ROC curve, and if we discuss the classification report, class 0 has a value of 0.92 for precision, 1 for recall, and class 1 has a value of 0.51 for precision, and 0.03 for recall. The overall accuracy achieved by the model is 92%, as shown in Figure 10.

4.4. Bank Loan Default Dataset

In the logistic regression model, it was predicted that two outcomes for binary classification could be found. Its random state is 42, which means 42 is the starting point of this model, and the maxiter is 1000, meaning the model will keep trying to improve itself 1000 times. This Figure shows the ROC curve value of the True Positive (TP) and True Negative (TN) rate combination at the Y and X axes, respectively. It achieved an AUC value of 69% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 has a value of 0.75 for precision and 0.01 for recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154960 in TP, 3 in TN, 1 in FP, and 579 in FN. This confusion matrix shows that the model is not perfect but performs well, as shown in Figure 11.

In the Random Forest model, Its random state is 42, which means 42 is the starting point of this model. This Figure shows the ROC curve value of the TP and TN rate combination at the Y and X axes, respectively. It achieved an AUC value of 100% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 has a value of 0.1 for precision and 0.99 for recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154961 in TP, 578 in TN, 0 in FP, and 4 in FN. This confusion matrix showed that the model performs very well and has achieved near-perfect performance for the given dataset and problem, as shown in Figure 11.

In the XGBoost model, it was proposed that possible results be found confined to accuracy rate. Its random state is 42, learning rate 0.1, n estimators 100, max depth 6. This Figure shows the ROC curve value of the TP and TN rate combination at the Y and X axes, respectively. It achieved an AUC value of 100% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 also has a value of 1 for precision and recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154961 in TP, 582 in TN, 0 in FP, and 0 in FN. This confusion matrix showed that the model performs exceptionally well, achieving a perfect classification for the given dataset and problem, as shown in Figure 11.

The random state of the decision tree model is 42, which is the starting point of this model. This Figure shows the ROC curve value of the TP and TN rate combination at the Y and X axes, respectively. It achieved an AUC value of 100% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 also has a value of 1 for precision and recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154961 in TP, 582 in TN, 0 in FP, and 0 in FN. This confusion matrix shows that the model performs very well and has achieved a perfect classification for the given dataset and problem, as shown in Figure 11.

The K-Nearest Neighbors model was designed to determine classification outcomes. Its N Neighbor is 5, which means 5 is the starting point of this model. This Figure shows the ROC curve value of the TP and TN rate combination at the Y and X axes, respectively. It achieved an AUC value of 96% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 has a value of 0.96 for precision and 0.70 for recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154944 in TP, 408 in TN, 17 in FP, and 174 in FN. This confusion matrix showed that the model is imperfect for classifying the given dataset and problem, as shown in Figure 12.

In the LSTM model, it was predicted that possible outcomes would be found in combination at the Y and X axes, respectively. It achieved an AUC value of 80% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 has a value of 0 for both precision and recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154961 in TP, 0 in TN, 0 in FP, and 582 in FN. This confusion matrix showed that the model is achieving high sensitivity but has low specificity, indicating it may be biased towards classifying samples as positive, potentially leading to false positives, as shown in Figure 12.

In the CNN model, it was predicted that possible outcomes would be found. This Figure shows the ROC curve value of the TP and TN rate combination at the Y and X axes, respectively. It achieved an AUC value of 50% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 has a value of 0 for both precision and recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154961 in TP, 0 in TN, 0 in FP, and 582 in FN. The model is not well-balanced as it has high sensitivity but low specificity, which may not be suitable for all applications, as shown in Figure 12.

In the DNN model, it was predicted that possible outcomes would be found. This Figure shows the ROC curve value of the TP and TN rate combination at the Y and X axes, respectively. It achieved an AUC value of 100% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 has a value of 0.94 for precision and 1 for recall. The overall accuracy achieved by the model is 100%. Further, we discussed the confusion matrix for this model, which has 154925 in TP, 580 in TN, 36 in FP, and 2 in FN. The model is performing well with a good balance between true positives and true negatives, indicating it has a reasonable ability to classify both positive and negative instances, as shown in Figure 13.

The RBFN is a classification task model comprising an input layer, a hidden layer with RBF activation functions, and an output layer. The number of clusters in the hidden layer is adjustable to fit the dataset. It excels at capturing complex patterns in data for enhanced classification. This Figure shows the ROC curve value of the TP and TN rate combination at the Y-axis and X-axis, respectively. It achieved an AUC value of 100% for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 also has 1 value for precision and recall. Further, we discussed the confusion matrix for this model, which has 154961 in TP, 582 in TN, 0 in FP, and 0 in FN. The model performs well, with a good balance and 100% accuracy in this scenario, as shown in Figure 13.

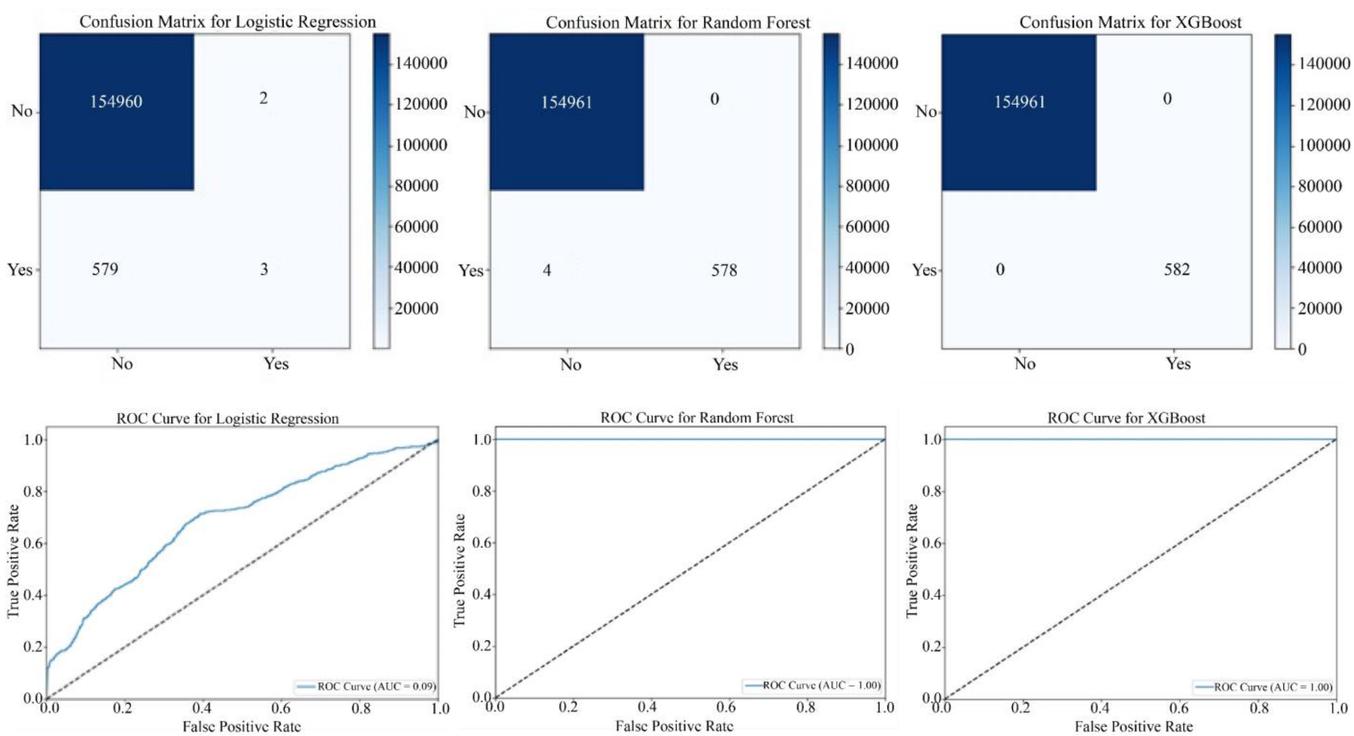


Fig. 11 Logistic Regression Model, RF, XGBoost

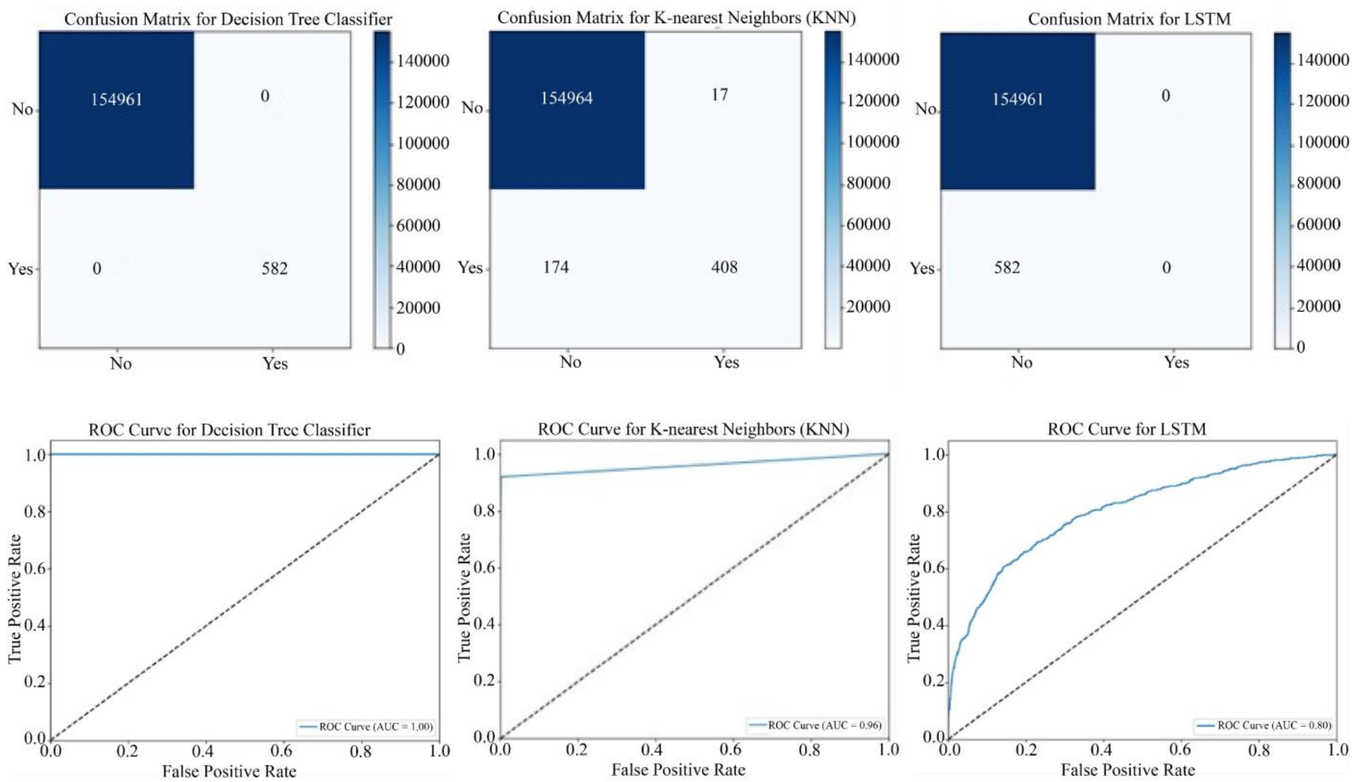


Fig. 12 Decision tree model, K-Nearest Neighbors model, LSTM model

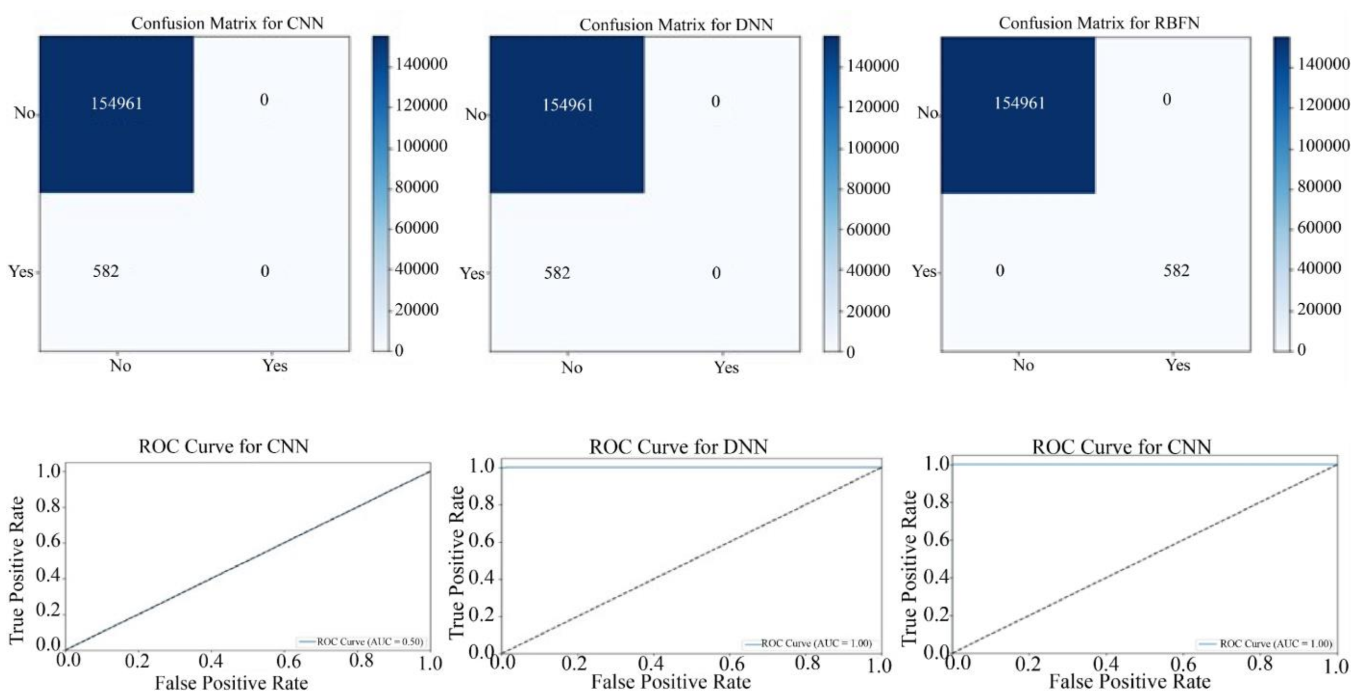


Fig. 13 CNN Model, DNN Modelm RBFN Model

The model consists of densely connected layers with ReLU activation functions, followed by an output layer designed for binary classification tasks using sigmoid activation. The model is then compiled, specifying ’adam’ as the optimizer, ’binary cross-entropy’ as the loss function, and ’accuracy’ as the evaluation metric. This Figure shows the ROC curve value of the TP and TN rate combination at the Y and X axes, respectively. It achieved an AUC value of 100%

for the ROC curve, and if we discussed the classification report, class 0 has a value of 1 for both recall and precision, and class 1 also has 1 value for precision and recall. Further, we discussed the confusion matrix for this model, which has 154961 in TP, 582 in TN, 0 in FP, and 2 in FN. The model is performing well in classifications as shown in Figure 13.

Based on the results from the various models tested across four datasets, we recommend the following approaches for optimal classification performance. For the Home Loan Dataset, the XGBoost and Decision Tree models demonstrated superior performance with an overall accuracy of 100%, making them the preferred choices for this dataset. These models consistently achieved high precision and recall values, particularly for class 1, which suggests their robustness in handling the data.

In Lending Club Loan Data, models such as Random Forest, XGBoost, and Decision Tree also achieved perfect accuracy, underscoring their efficacy in managing complex patterns within the dataset. The high AUC values of these models indicate their predictive solid power, making them highly recommended for similar classification tasks.

For Loan Default Prediction Dataset, while the overall accuracy for most models was around 92%, the Logistic Regression and Random Forest models were particularly effective, achieving a balanced performance with high precision and recall for class 0. However, noting the lower recall values for class 1 across most models is essential, indicating potential room for improvement in minority class detection.

In the Bank Loan Default Dataset, the Random Forest and XGBoost models again proved to be the most reliable, achieving 100% accuracy and high AUC values. These models could perfectly classify positive and negative instances, demonstrating their robustness and consistency. Given these observations, it is evident that ensemble methods such as Random Forest and XGBoost are highly effective across diverse datasets, providing a balance between sensitivity and specificity.

Therefore, we recommend prioritizing these models in future classification tasks, especially when high accuracy and robustness are required. Fine-tuning these models further and employing additional techniques such as cross-validation can enhance their performance even more, ensuring optimal results across various datasets.

5. Comparative Analysis

This study provides a comprehensive evaluation of various machine learning models applied to binary classification tasks across four distinct datasets. The findings of this research both align with and diverge from previous studies in the field, highlighting the unique contributions and novel insights of our work.

Khan et al. [1] highlighted the significance of loan prediction models in the financial sector, emphasizing the importance of accurate classification to minimize the risk of defaults. Our study concurs with these findings, particularly in the context of the Lending Club Loan Data, where models like Random Forest and XGBoost achieved perfect accuracy, demonstrating their robustness and effectiveness. This aligns with Khan et al.'s assertion that advanced machine learning techniques can significantly enhance predictive accuracy in financial applications.

Li et al. [2] examined the application of XGBoost in personal credit evaluation and found it to be superior in feature selection and classification performance compared to traditional techniques like logistic regression. Our results support this, as XGBoost consistently outperformed other models in terms of AUC and overall accuracy across multiple datasets. Specifically, in the Home Loan Dataset, XGBoost achieved 100% accuracy, reinforcing the findings of Li et al. on the efficacy of ensemble methods.

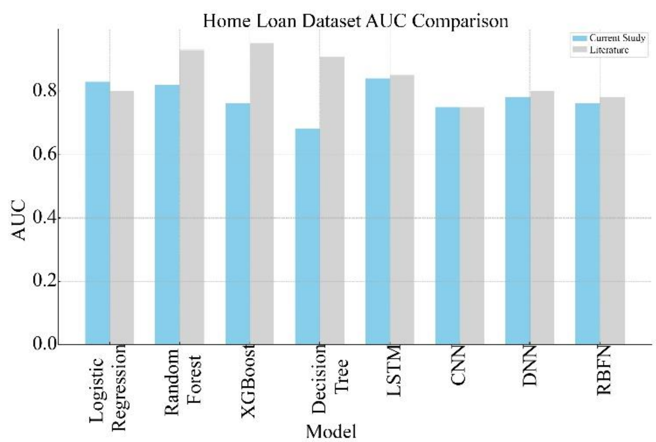


Fig. 14 Home loan data comparison with literature

In contrast, Zhong and Zhou [3] focused on the risk analysis of bank microfinance using genetic artificial neural networks and reported an overall accuracy of 80%. While our study did not specifically evaluate genetic algorithms, the performance of neural networks, such as LSTM and CNN, varied depending on the dataset. For instance, in the Loan Default Prediction Dataset, LSTM achieved an AUC of 75%, indicating moderate performance. This suggests that while neural networks have potential, their effectiveness can be dataset-dependent, which partially aligns with Zhong and Zhou's findings but also highlights the variability in neural network performance.

Khan et al. [1] and Sarkar [6] both discussed the comparative effectiveness of logistic regression, decision trees, and random forests. Sarkar reported that logistic regression emerged as the most accurate model for loan eligibility prediction with an accuracy of 80.78%, while our study found that logistic regression achieved high accuracy in the Loan Default Prediction Dataset but was outperformed by ensemble methods like Random Forest and XGBoost in other datasets. This divergence underscores the importance of considering dataset characteristics when selecting a machine learning model.

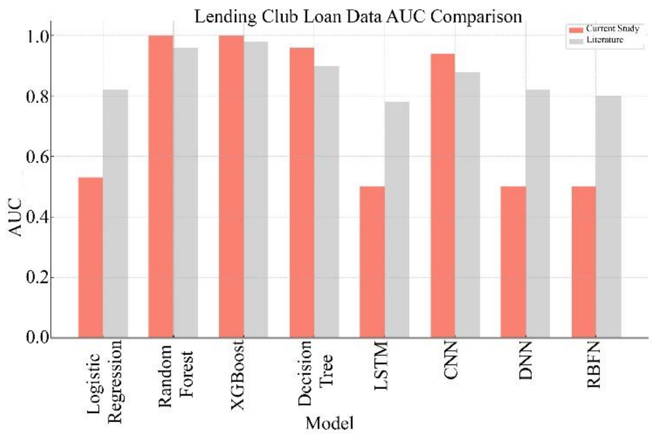


Fig. 15 Lending club data comparison with literature

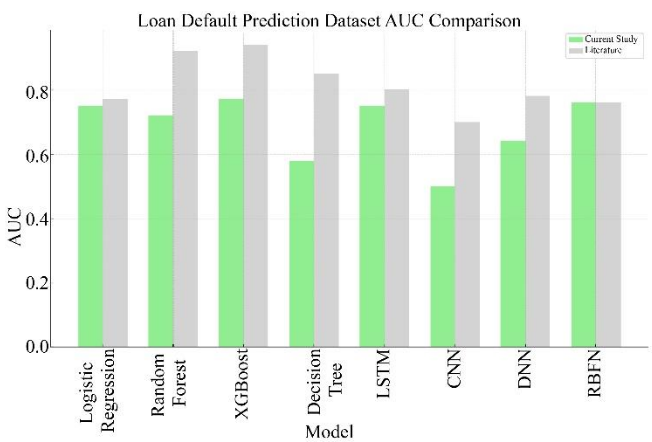


Fig. 16 Loan default prediction data comparison with literature

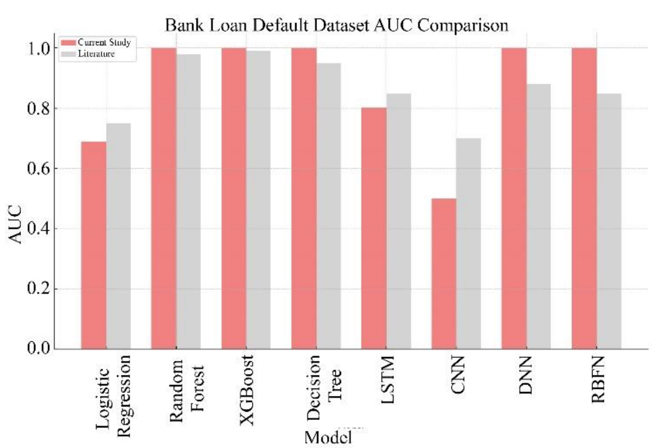


Fig. 17 Bank loan data comparison with literature

Park et al. [5] explored the application of LIME for enhancing the interpretability of machine learning models in bankruptcy prediction. While our study did not focus on model interpretability, the high performance of tree-based models like Decision Trees and Random Forests in our datasets supports the notion that such models not only provide high accuracy but also offer better explainability compared to complex neural networks.

This research confirms the superior performance of ensemble methods, particularly Random Forest and XGBoost, in binary classification tasks across diverse datasets, consistent with findings from Li et al. [2] and Khan et al. [1]. However, it also highlights the variability in neural network performance depending on dataset characteristics, providing a nuanced understanding that complements the existing literature. This study's comprehensive approach offers valuable insights into model selection and tuning, contributing to more effective data observability and decision-making processes in various domains.

6. Conclusion

This study provides a detailed evaluation of various machine learning models applied to four distinct datasets, focusing on their performance in binary classification tasks. The comparative analysis revealed that ensemble methods like Random Forest and XGBoost consistently outperformed other models, achieving perfect or near-perfect accuracy and high AUC values across most datasets. These models demonstrated exceptional robustness and reliability, making them ideal choices for complex classification tasks. The analysis also highlighted that while models such as Logistic Regression and Decision Trees performed well in specific scenarios, their effectiveness varied depending on the dataset's characteristics. Evidently, no single model universally outperforms others across all datasets, underscoring the importance of selecting the appropriate model based on the specific dataset and problem requirements.

The findings suggest that ensemble methods should be prioritized in future classification tasks due to their superior performance. However, continuous model tuning and validation are essential to address problems related to class imbalance and further enhance predictive accuracy. This study contributes valuable insights into the practical application of machine learning models, guiding practitioners in making informed decisions to improve data observability and decision-making processes with AI and LLMs.

References

[1] Afrah Khan et al., "Loan Approval Prediction Model: A Comparative Analysis," *Advances and Applications in Mathematical Sciences*, vol. 20, no. 3, pp. 427-435, 2021. [Google Scholar] [Publisher Link]

[2] Hua Li et al., "XGBoost Model and its Application to Personal Credit Evaluation," *IEEE Intelligent Systems*, vol. 35 no. 3, pp. 52-61, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Xiong Zhong, and Sheng Zhou, "Retracted Article: Risk Analysis Method of Bank Microfinance Based on Multiple Genetic Artificial Neural Networks," *Neural Computing and Applications*, vol. 32, pp. 5367-5377, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] S Sreesouthry et al., "Loan Prediction Using Logistic Regression in Machine Learning," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 4, pp. 2790-2794, 2021. [Google Scholar] [Publisher Link]

[5] Min Sue Park et al., "Explainability of Machine Learning Models for Bankruptcy Prediction," *IEEE Access*, vol. 9, pp. 124887-124899, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Abhiroop Sarkar, "Machine Learning Techniques for Recognizing the Loan Eligibility," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 3, no. 12, pp. 1135-1142, 2021. [Google Scholar] [Publisher Link]

[7] Yingli Wu et al., "The Analysis of Credit Risks in Agricultural Supply Chain Finance Assessment Model Based on Genetic Algorithm and Backpropagation Neural Network," *Computational Economics*, vol. 60, pp. 1269-1292, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Joseph Bamidele Awotunde et al., "Artificial Intelligence Based System for Bank Loan Fraud Prediction," *21st International Conference on Hybrid Intelligent Systems*, pp. 463-472, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Syed Emad Azhar Ali et al, "Predicting Delinquency on Mortgage Loans: An Exhaustive Parametric Comparison of Machine Learning Techniques," *International Journal of Industrial Engineering and Management*, vol. 12, no. 1, pp. 1-13, 2021. [Google Scholar] [Publisher Link]

[10] Marion O Adebiyi et al., "Secured Loan Prediction System Using Artificial Neural Network," *Journal of Engineering Science and Technology*, vol. 17, no. 2, pp. 854-873, 2022. [Google Scholar] [Publisher Link]

[11] Ugochukwu E. Orji et al., "Machine Learning Models for Predicting Bank Loan Eligibility," *2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON)*, Lagos, Nigeria, pp. 1-5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] An-Hsing Chang et al., "Machine Learning and Artificial Neural Networks to Construct P2P Lending Credit-Scoring Model: A Case Using Lending Club Data," *Quantitative Finance and Economics*, vol. 6, no. 2, pp. 303-325, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] S. Varshaa Sai Sripriya, Sai Divya Santoshi Varrey, and M Venkateshkumar, "Predictive Model to Compute Eligibility Test for Loans," *2022 IEEE Industrial Electronics and Applications Conference (IEACon)*, pp. 185-190, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Tianhui Wang, Renjing Liu, and Guohua Qi, "Multi-Classification Assessment of Bank Personal Credit Risk Based on Multi-Source Information Fusion," *Expert Systems with Applications*, vol. 191, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Theodore E Simos, Vasilios N Katsikis, and Spyridon D Mourtas, "A Multi-Input with Multi-Function Activated Weights and Structure Determination Neuronet for Classification Problems and Applications in Firm Fraud and Loan Approval," *Applied Soft Computing*, vol. 127, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Zhuo Chen et al., "An Artificial Neural Network-Based Intelligent Prediction Model for Financial Credit Default Behaviors," *Journal of Circuits, Systems and Computers*, vol. 32, no. 10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] G.L. Infant Cyril, and J.P. Ananth, "Deep Learning Based Loan Eligibility Prediction with Social Border Collie Optimization," *Kybernetes*, vol. 52, no. 8, pp. 2847-2867, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Sergio Genovesi et al., "Standardizing Fairness-Evaluation Procedures: Interdisciplinary Insights on Machine Learning Algorithms in Creditworthiness Assessments for Small Personal Loans," *AI and Ethics*, vol. 4, pp. 537-553, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Robert H Mnookin, and Lewis Kornhauser, "Bargaining in the Shadow of the Law: The Case of Divorce," *The Yale Law Journal*, vol. 88, no. 5, pp. 950-997, 1979. [CrossRef] [Google Scholar] [Publisher Link]

[20] Joel Greenberg, *Of Prairie, Woods, and Water: Two Centuries of Chicago Nature Writing*, University of Chicago Press, pp. 1-402, 2008. [Google Scholar] [Publisher Link]

[21] Glenn Gould, *The Glenn Gould Reader*, Knopf, pp. 1-475, 1984. [Google Scholar] [Publisher Link]

[22] William J. Novak, "The Myth of the 'Weak' American State," *The American Historical Review*, vol. 113, no. 3, pp. 752-772, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[23] Geoffrey C. Ward, and Ken Burns, *The War: An Intimate History, 1941-1945*, Knopf, pp. 1-453, 2007. [Google Scholar] [Publisher Link]

[24] OECD. Publishing, International Energy Agency, International Engery Agency, *Energy Balances of Non-OECD Countries 2010*, OECD/IEA, pp. 1-554, 2010. [Google Scholar] [Publisher Link]

[25] International Energy Agency, *Energy Balances of OECD Countries 2010*, OECD, pp. 1-295, 2010. [Google Scholar] [Publisher Link]

[26] Sergey Paltsev et al., "The Cost of Climate Policy in the United States," *Energy Economics*, vol. 31, pp. s235-s243, 2009. [CrossRef] [Google Scholar] [Publisher Link]