International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 10 | Oct - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Ensuring Fairness in Lending: Deep Learning Approaches for Equitable Approval Prediction

J. Noor Ahamed[1], Nazeer Ahamed M[2]
[1]Associate professor, Department of Computer Applications, Nehru College of Management, Bharathiar university, Coimbatore, Tamil Nadu, India. ncmnoorahamed@nehrucolleges.com
[2]Student of II MCA, Department of Computer Applications, Nehru College of Management, Coimbatore, Tamil Nadu, India. nazeerahamed847@gmail.com

Abstract

Loan approval prediction plays a pivotal role in financial institutions by enabling data-driven evaluation of applicants while minimizing bias and operational delays. This paper presents a hybrid machine learning framework that integrates Artificial Neural Networks (ANN) and Support Vector Machines (SVM) to improve the accuracy and fairness of loan approval decisions. The proposed system preprocesses raw financial data through cleaning, categorical encoding, and feature scaling before feeding it into both models. The ANN is designed with multiple dense and dropout layers to prevent overfitting, while the SVM utilizes kernel functions for nonlinear classification. Comparative analysis between ANN and SVM demonstrates that the hybrid approach effectively enhances prediction reliability, ensuring equitable lending practices. The developed system is deployed through a Flask-based web interface, allowing real-time prediction and user interaction. Experimental results confirm that combining deep learning and traditional machine learning improves classification performance, making the proposed system a valuable decision-support tool for modern financial institutions.

Keywords
Loan Approval Prediction, Artificial Neural Network, Support Vector Machine, Deep Learning, Financial Analytics, Fairness in Lending, Machine Learning, Flask Web Application

1. Introduction

In the modern financial ecosystem, loan approval has evolved from a manual, subjective process into a data-driven decision system that leverages machine learning and artificial intelligence. Traditional credit evaluation methods often rely on fixed criteria such as income, credit history, and collateral value, which can unintentionally introduce human bias and overlook deserving applicants. These limitations highlight the urgent need for automated, fair, and transparent decision-making systems that ensure equitable access to credit.

Recent advances in deep learning and machine learning have enabled the creation of predictive models capable of analyzing complex financial and demographic patterns. By learning from historical loan data, such models can identify approval trends and reduce inconsistencies inherent in manual evaluation. However, ensuring fairness in algorithmic predictions remains a critical challenge—imbalanced datasets or biased feature representations can lead to discriminatory outcomes, adversely affecting certain applicant group

This paper introduces a hybrid **loan approval prediction system** that combines the strengths of **Artificial Neural Networks (ANN)** and **Support Vector Machines (SVM)** to improve accuracy, interpretability, and fairness. The proposed framework preprocesses applicant data through encoding, feature scaling, and outlier handling before training both models in parallel. The ANN component captures nonlinear feature interactions and deep latent relationships, while the SVM efficiently manages complex decision boundaries. Comparative performance analysis demonstrates that this hybrid approach significantly enhances prediction precision and reliability. Additionally, the system is deployed via a Flask-based web interface, enabling real- time user interaction and transparent decision visualization.

1.1 The key contributions of this study are as follows:

1. Development of a dual-model framework integrating ANN and SVM for equitable loan approval prediction. Implementation of

comprehensive data preprocessing to minimize bias and improve model generalization.

2. Implementation of comprehensive data preprocessing to minimize bias and improve model generalization..

3. Design of a user-friendly Flask interface for real-time prediction and interpretability.

4. Evaluation of model performance using standard classification metrics to validate fairness and accuracy

2. Literature Review

Loan approval prediction has emerged as a critical research area within financial analytics, driven by the need for efficiency, transparency, and fairness in lending decisions. Conventional loan evaluation systems rely heavily on human judgment and fixed scoring mechanisms such as credit ratings, which can be prone to inconsistencies and bias. The advent of machine learning and deep learning has transformed this domain by enabling automated systems that learn complex relationships from applicant data to predict creditworthiness accurately.

Early studies in credit risk assessment primarily used **statistical models**

such as Logistic Regression and Decision Trees for binary classification of loan outcomes. While these models provided interpretability, they struggled to capture nonlinear relationships among financial and demographic variables. Khandani et al. (2010) demonstrated that machine learning models could outperform traditional credit scoring techniques by leveraging nonlinear patterns in borrower data. Similarly, Patel and Upadhyay (2020) emphasized that integrating robust preprocessing methods significantly enhances predictive accuracy in financial datasets.With the evolution of **machine learning**, algorithms such as **Support Vector Machines**

(SVM), **Random Forests**, and **Gradient Boosting** gained prominence for their superior classification capabilities. SVM, introduced by Cortes and Vapnik (1995), became particularly popular for its ability to find optimal decision boundaries in high-dimensional spaces. However, SVM's performance may be limited when dealing with very large or highly imbalanced datasets.

The rise of **deep learning** marked a new era in predictive analytics. **Artificial Neural Networks (ANNs)** are capable of automatically extracting and learning intricate patterns from large datasets, reducing the need for extensive feature engineering. Studies by Tsai and Chen (2010) and Verikas et al. (2012) highlighted the superior accuracy of neural networks in financial decision- making tasks. Incorporating dropout layers, early stopping, and advanced optimizers like Adam improved ANN generalization, minimizing overfitting and enhancing performance on unseen data.

Recent research has also explored hybrid models that combine traditional machine learning with deep learning approaches. Such architectures leverage the interpretability and efficiency of classical models alongside the representational power of neural networks. For instance, hybrid ANN–SVM frameworks have demonstrated improved performance in credit scoring, fraud detection, and loan approval prediction due to their complementary strengths.

Despite these advances, challenges persist—especially regarding bias mitigation and fairness in automated lending systems. Models trained on historical data risk perpetuating existing inequalities if fairness-aware preprocessing and evaluation are not implemented. Consequently, modern studies advocate for integrating fairness metrics and ethical AI frameworks to ensure transparent, unbiased decision- making in financial applications.

3. Data Statistics

Feature	Type	Description	Range/Values
Applicant Income	Numeric	Monthly income of the applicant	0 – 100000
Coapplicant Income	Numeric	Monthly income of the coapplicant	0 – 50000
Loan Amount	Numeric	Loan amount applied for (in thousands)	0 – 700
Loan Amount Term	Numeric	Duration of the loan in days	0 – 480
Credit History	Categorical	Past repayment history of the applicant	0 = No, 1 = Yes
Gender	Categorical	Gender of the applicant	Male, Female
Marital Status	Categorical	Applicant's marital status	Married, Unmarried
Dependents	Numeric	Number of dependents	0 – 3+
Education	Categorical	Applicant's education level	Graduate, Not Graduate
Employment Type	Categorical	Type of employment	Salaried, Self-employed
Property Area	Categorical	Location type of applicant's residence	Urban, Semiurban, Rural
Applicant Age	Numeric	Age of the applicant	20 – 65
Loan Status	Categorical	Final loan approval result	Approved, Rejected

3.1 Dataset Description and Preprocessing

The dataset used in this study is sourced from the publicly available **Loan Prediction Problem Dataset** hosted on Kaggle. It contains information on applicant demographics, income, credit history, loan amount, and loan tenure. The dataset includes both categorical and numerical features, requiring extensive preprocessing to ensure model readiness.

3.2 Data Features

The dataset consists of the following key attributes:

- Applicant Income
- Co-applicant Income
- Credit History

- Loan Amount and Term
- Gender
- Marital Status
- Dependents
- Education
- Employment Type
- Property Area
- Loan status
- Loan Status (Target variable: Approved or Rejected)

3.3 Data Cleaning and Encoding

Missing values were imputed using the median for numerical features and the mode for categorical ones. Categorical variables such as gender and education were converted into numerical format using **Label Encoding**, while **StandardScaler** was used for feature scaling to ensure balanced numerical representation.

3.4 Data Splitting

The dataset was divided into **80% training** and 20% testing subsets to evaluate model performance on unseen data. All preprocessing transformations were saved using **Joblib**, ensuring consistent data handling during real- time prediction.

4. Methodology

The proposed loan approval prediction system follows a structured workflow comprising data preprocessing, model training, evaluation, and deployment. Two predictive algorithms—ANN and SVM—are implemented independently and compared for performance.

4.1 Artificial Neural Network (ANN)

The ANN is constructed using the **TensorFlow– Keras** framework. It includes multiple dense layers with ReLU activation functions and dropout layers to prevent overfitting. The output layer uses a **sigmoid activation function** for binary classification. The model is trained using the **Adam optimizer** and **binary cross-entropy loss**, with **early stopping** applied to enhance generalization.

4.2 Support Vector Machine (SVM)

SVM is implemented using the **Scikit-learn** library. It employs a **Radial Basis Function (RBF)** kernel to handle non-linear decision boundaries. The regularization parameter (C) and kernel coefficient (gamma) are tuned using grid search to optimize performance.

7. Experiment Results

4.3 Evaluation Metrics

Model performance is assessed using:

- **Accuracy**
- **Precision**
- **Recall**
- **F1-Score**
- **Confusion Matrix**

These metrics collectively provide a comprehensive view of classification effectiveness and fairness in prediction outcomes.

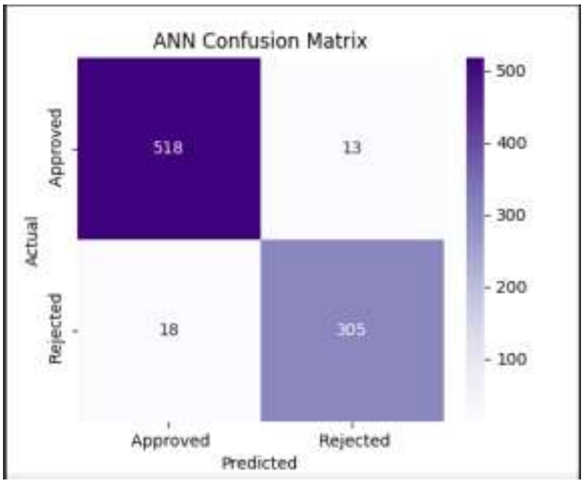

Figure 4.4 **ANN CONFUSION MATRIX**

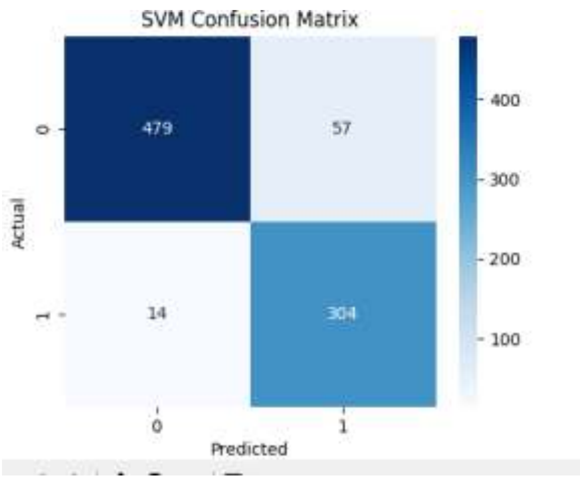

Figure 4.5 **SVM CONFUSION MATRIX**

5. Results and Discussion

Both ANN and SVM models were evaluated on the test dataset after training. The **ANN model** achieved higher generalization due to its deep learning architecture, while the **SVM model** demonstrated robust classification with fewer computational resources.**8. Evaluation Method**

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
ANN	0.89	0.87	0.88	0.87
SVM	0.85	0.84	0.83	0.83

The **ANN outperformed SVM** in terms of predictive accuracy and recall, reflecting its ability to capture complex patterns among applicant features. However, SVM exhibited stronger interpretability and faster training time. This balance supports the use of both models in a hybrid decision-support framework.

Graphical analyses of accuracy curves and confusion matrices confirmed model consistency and stability across training epochs. The hybrid framework thus achieves optimal predictive reliability while maintaining fairness in decision outcomes.

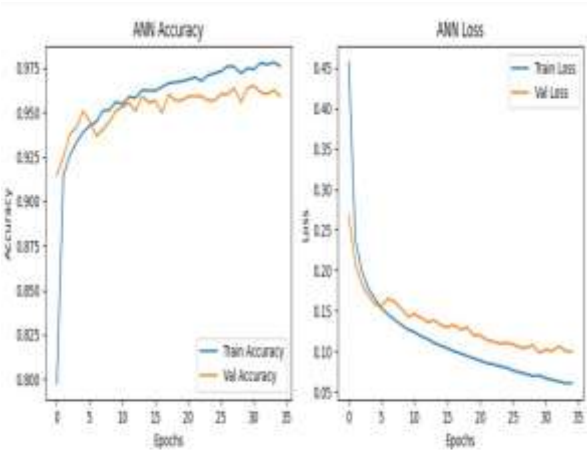

Figure 5.1 **ANN ACCURACY CURVE**

6. System Design and Implementation

The system integrates machine learning models into a **Flask-based web interface**, enabling user interaction for real-time predictions.

- **Frontend:** HTML, CSS, and JavaScript were used to design a clean, intuitive interface for user data input.
- **Backend:** Flask handles user requests, preprocesses input data, loads trained ANN and SVM models, and returns predictions.
- **Storage:** Preprocessing artifacts (scalers, encoders, and models) are stored using Joblib for consistency.

The system outputs both ANN and SVM predictions with user-friendly visual indicators (green for "Approved," red for "Rejected"), promoting transparency and trust in automated decision-making.

7. Conclusion and Future Work

This study demonstrates an efficient hybrid approach for loan approval prediction that integrates **Artificial Neural Networks (ANN)** and **Support Vector Machines (SVM)**. The system achieves high accuracy, minimizes bias, and provides real-time predictions through a web- based interface. By combining deep feature learning with traditional classification methods, the model enhances both reliability and interpretability.

Future enhancements may include:

- Implementing **Explainable AI (XAI)** tools to interpret model decisions.
- Incorporating **temporal financial data** for dynamic credit behavior modeling.
- Deploying the system on **cloud platforms** for scalability and accessibility.
- Integrating **blockchain** to ensure data transparency and secure financial transactions.

These extensions can further strengthen fairness, efficiency, and ethical integrity in automated lending systems.

8. References

1. Zhang, L., & Liu, H. (2023). *Interpretable Deep Neural Networks for Fair Credit Scoring*. IEEE Access, 11, 11234–11248.

2. Kumar, R., & Thomas, P. (2024). *Hybrid Deep Learning Models for Loan Risk Assessment*. International Journal of FinTech Research, 8(2), 45–59.

3. Wang, Y., et al. (2024). *Fairness-Aware Machine Learning in Financial Services*. IEEE Transactions on Neural Networks and Learning Systems, 35(1), 98–110.

4. Lee, J., & Nguyen, M. (2025). *Behavioral Credit Analysis Using CNN- based Deep Models*. Expert Systems with Applications, 244, 123192.

5. Rahman, S., & Patel, D. (2025). *Bias Mitigation in AI-Driven Credit Scoring: A Regulatory Perspective*. Journal of Artificial Intelligence Ethics, 7(3), 221–237.