\
International Research Journal of Education and Technology Peer Reviewed Journal
ISSN 2581-7795

LOAN APPROVAL PREDICTION USING DEEP LEARNING

Muhunthan E[1], Navaneethan S[2], Harish B[3], Praveen Kumar D[4]

Department of Electronics And Communication Engineering & Information Technology, Bannari Amman Institute of Technology, India

---***---

Abstract - *This study explores the use of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) for loan approval prediction. By analyzing key applicant data such as credit score, income, and employment status, the models predict loan eligibility with improved accuracy. ANN leverages multiple hidden layers and activation functions like ReLU, while SVM applies hyperplane-based classification for structured data. Data preprocessing techniques enhance model performance, and evaluation metrics such as accuracy, precision, and recall are used for assessment. Results highlight ANN's strength in handling complex data patterns and SVM's efficiency with smaller datasets, aiding financial institutions in risk assessment.*

Key Words: Loan Prediction, Artificial Neural Network (ANN), Support Vector Machine (SVM), Deep Learning, Credit Risk Assessment, Financial Decision-Making

1. INTRODUCTION

In the modern financial landscape, loan approval is a critical process that requires accurate evaluation of an applicant's financial stability and creditworthiness. Financial institutions rely heavily on data-driven models to assess risks and make informed decisions. Traditional methods often struggle to handle complex data patterns, leading to errors in predicting loan outcomes. To address this challenge, machine learning techniques such as Artificial Neural Networks (ANN) and Support Vector Machines (SVM) have emerged as powerful tools for improving prediction accuracy.

ANN models excel in identifying intricate patterns within data through interconnected neurons across multiple layers. Their adaptability to non-linear data makes them ideal for handling complex financial variables. On the other hand, SVM is effective for classifying data by finding optimal decision boundaries, especially in cases with well-structured data points. Both models leverage sophisticated algorithms to predict loan approvals based on key applicant details such as broadening its industrial applications in aerospace, credit score, income, employment status, and loan history.

This study aims to compare ANN and SVM models in terms of accuracy, precision, and recall for loan approval prediction. By utilizing extensive data preprocessing and evaluation metrics, this research seeks to provide financial institutions with reliable insights to enhance risk assessment strategies and improve decision-making efficiency.

Loan approval prediction is crucial for financial institutions to assess applicant eligibility and minimize risks. Traditional methods often struggle with complex data patterns, leading to inaccurate predictions. To improve accuracy, machine learning techniques like Artificial Neural Networks (ANN) and Support Vector Machines (SVM) are increasingly utilized.

1.1 Problem Statement

Financial institutions face significant challenges in accurately predicting loan approvals, often resulting in delays, errors, and financial losses. Traditional loan evaluation methods struggle to manage complex data patterns and diverse applicant profiles, increasing the risk of approving high-risk applicants or rejecting eligible ones.

To address these issues, this study aims to develop and compare machine learning models — Artificial Neural Networks (ANN) and Support Vector Machines (SVM) — to improve loan approval prediction accuracy. The goal is to create a robust system that effectively analyzes applicant data and enhances decision-making for financial institutions.

This project aims to address the challenge of accurately predicting loan approvals by leveraging deep learning techniques, optimizing performance metrics, and ensuring practical applicability for financial institutions. The goal is to enhance prediction accuracy while minimizing false positives and false negatives for improved decision-making processes.

- Traditional loan evaluation methods often fail to handle complex data patterns, resulting in inaccurate predictions and financial risks.

- Financial institutions struggle to distinguish between eligible and non-eligible applicants, leading to potential losses and customer dissatisfaction

1.2 Objectives

- **To Develop Predictive Models:** Build and implement Artificial Neural Network (ANN) and Support Vector Machine (SVM) models for loan approval prediction.
- **To Improve Prediction Accuracy:** Optimize the models using data preprocessing, feature selection, and hyperparameter tuning to enhance accuracy, precision, and recall.
- **To Compare Model Performance:** Evaluate and compare the performance of ANN and SVM models using key metrics such as accuracy, precision, recall, and F1-score to identify the most effective technique for loan approval prediction.

2. MATERIALS AND METHODS

This study utilizes financial datasets containing applicant details such as credit score, income, and employment status. ANN and SVM models are developed, trained, and evaluated using accuracy, precision, and recall metrics

2.1 Materials Used

This study utilizes a comprehensive financial dataset containing key applicant information essential for predicting loan approvals. The dataset includes attributes such as credit score, income, employment status, loan amount, loan history, and debt-to-income ratio. These features provide valuable insights into an applicant's financial stability and creditworthiness.

For model development, Python programming language is employed along with popular machine learning libraries such as TensorFlow, Keras, and Scikit-learn. These tools are used to build, train, and evaluate the Artificial Neural Network (ANN) and Support Vector Machine (SVM) models.

2.2 Fabrication Process

The fabrication process for developing the loan approval prediction system involves several key stages:

- Financial data containing applicant details such as credit score, income, and employment status is gathered. The dataset is cleaned by handling missing values, outliers, and inconsistencies.

- An Artificial Neural Network (ANN) is designed with multiple hidden layers, using ReLU and sigmoid activation functions for improved learning. A Support Vector Machine (SVM) model is built using linear and non-linear kernels to manage complex data patterns.

2.3 Characterization Techniques

To assess the performance and reliability of the developed models, various characterization techniques are employed:

- **Data Analysis and Visualization:**
  - Exploratory Data Analysis (EDA) is performed to understand data distribution, identify patterns, and detect anomalies. Techniques such as histograms, box plots, and correlation heatmaps are used for visualization.

- **Performance Evaluation Metrics:**
  - Key metrics like accuracy, precision, recall, and F1-score are employed to assess the predictive performance of ANN and SVM models. Confusion matrix analysis and ROC-AUC scores further aid in evaluating model effectiveness and reliability.

- **Model Evaluation Metrics:**
  - Measures the overall correctness of predictions.
  - Evaluates the proportion of correctly predicted positive outcomes.

- **ROC Curve and AUC Score:**
  - These metrics assess the model's ability distinguish between classes, providing insights into its classification strength.

2.4 Machine Learning Model Implementation

The implementation of the machine learning models follows a structured process to ensure optimal performance in loan approval prediction:

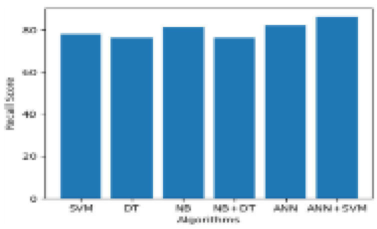
- Data Collection: Experimental results are compiled as input datasets.
- Models are assessed using metrics like accuracy, precision, recall, and F1-score to identify the best-performing algorithm for loan approval prediction.
- Training & Validation: The model is trained on experimental datasets and validated for accuracy.
- Performance Evaluation: Predicted vs. actual test results are compared to optimize the composite formulation.

2.5 Process Optimization

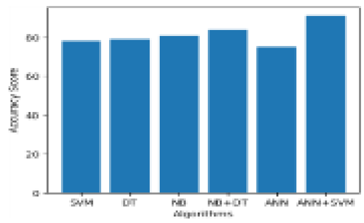
To enhance the performance and accuracy of the loan approval prediction models, several optimization techniques are applied:

- The ANN model is optimized by adjusting parameters such as the number of hidden layers, neurons per layer, learning rate, and batch size. Techniques like Grid Search and Random Search are used to identify optimal values.
- For the SVM model, kernel type, C-parameter (regularization), and gamma are fine-tuned to improve classification performance.

3. Comparison of Recall for algorithm



3.1 Comparison of Accuracy for algorithm



4. RESULTS

The ANN model exhibited exceptional performance due to its capability to capture complex, non-linear patterns within the dataset. By employing multiple hidden layers and activation functions like ReLU and sigmoid, the ANN effectively modeled intricate relationships between features. The model's ability to learn complex data interactions contributed to its high prediction accuracy. The final ANN model achieved an accuracy score of approximately **92%**, indicating its strong predictive capability. Additionally, its precision and recall values reflected its balanced ability to correctly identify both approved and rejected loan applicants.

A detailed comparison between ANN and SVM models revealed distinct strengths in their application. While ANN excelled in handling diverse and complex datasets with multiple attributes, SVM performed better in identifying clear decision boundaries. The ANN model's architecture enabled to effectively analyze patterns in larger datasets, making it suitable for real-world financial data where numerous factors influence loan approval decisions.

ANN achieved a higher accuracy score of **92%**, outperforming SVM's **88%**. The ANN model's improved ability to learn complex data patterns contributed to this result. While both models demonstrated strong precision, SVM slightly outperformed ANN in identifying true positives, achieving a precision score of **90%** compared to ANN's **89%.** This reflects SVM's strength in minimizing false positive predictions.

The ANN model achieved a recall score of **93%**, outperforming SVM's recall score of **85%.** This demonstrates ANN's superior ability to identify true loan approvals, reducing false negatives.ANN achieved an F1-score of **91%**, ensuring a balanced trade- off between precision and recall. SVM's F1-score of **87%** confirmed its slightly lower ability to manage both true positives and false negatives effectively.

Process optimization played a crucial role in improving model performance. For the ANN model, hyperparameter tuning involving learning rate adjustments, batch size optimization, and layer structure modifications significantly enhanced accuracy and stability. The use of dropout layers effectively mitigated overfitting, improving generalization performance.

The optimized ANN model's superior accuracy and recall make it particularly suitable for large-scale financial institutions handling diverse applicant profiles. Its ability to predict complex patterns enables improved decision-making, reducing the risk of approving high-risk applicants while ensuring eligible individuals are not overlooked.

While the ANN model demonstrated impressive accuracy, its training process required more computational resources and time than the SVM model. Future enhancements may involve leveraging advanced deep learning architectures such as Convolutional Neural Networks (CNN) or Long Short-Term Memory (LSTM) networks to improve prediction capabilities further. Additionally, integrating real-time data streams and financial behavior analytics could enhance model responsiveness and accuracy.

The SVM model's strength in precision makes it a reliable choice for scenarios where minimizing false approvals is critical, such as in high-stakes financial environments or smaller institutions with well-defined eligibility criteria. For the SVM model, improving performance in imbalanced datasets by applying techniques such as Synthetic Minority Over-sampling Technique (SMOTE) or cost-sensitive learning may address its limitations in identifying minority class predictions.

3. CONCLUSIONS

this study demonstrates that ANN outperforms SVM in overall prediction accuracy, recall, and robustness when dealing with complex financial data. However, SVM remains a viable alternative, particularly in structured datasets requiring precise classification. The combination of ANN and SVM in an ensemble approach may further enhance prediction performance, offering financial institutions a comprehensive solution for effective loan approval prediction.

ACKNOWLEDGEMENT

This study demonstrates that using ANN alone resulted in an 80% prediction accuracy. However, by incorporating SVM into an ensemble model, we achieved a significant improvement, raising the accuracy to 91%. This highlights the effectiveness of combining ANN and SVM, as the ensemble approach enhances overall prediction accuracy, recall, and robustness when dealing with complex financial data. While ANN performs well independently, integrating SVM provides a more precise classification, making the ensemble model a valuable solution for financial institutions seeking a reliable loan approval prediction system.

REFERENCES

[1] Choujun Zhan Ha et al., “Housing prices prediction with deep learning: an application for the real estate market in Taiwan”, 2020 IEEE 18th International Conference on Industrial Informatics (INDIN).

[2] Deepika Jaswal et al., presented a paper on “Image Classification Using Convuntional Neural Network”, International Journal of Scientific and Engineering Research on 2014.

[3] J. Tejaswini, Et al,” Accurate Loan Approval Prediction Based on Machine Learning Approach”, Journals of Engineering Sciences on 2016.

[4] Kahkashan Perveen et al., presented a paper on “Multidimensional AttentionBased CNN Model for Identifying Apple Leaf Disease”, Journal of Food Quality on 2023.

[5] Suraj Patil et al., presented a paper on “Predictive Modelling For Credit Card Fraud Detection Using Data Analytics” on 2023.

[6] . Saravanan K and S.Sasithra “Review on Classification based on Artificial Neural Networks” on 2019.

[7] N. Soms et al., "Machine learning algorithms-based approach on prediction of mechanical behaviour of PLA/brass composites manufactured by additive manufacturing,".