Loan Approval Prediction Using Neural Networks and CART

Muskan Sharma, Ravnoor Ubhi, Tsai-Hsun Lu, Ujwal Mistry, Utkarsh Singh, Vibhav Vohra

1. Introduction

The ability to accurately predict loan approval outcomes is a critical challenge for financial institutions aiming to minimize credit risk and improve decision-making efficiency. This project investigates whether **Machine Learning (ML)** techniques can effectively predict **"Loan Approval Status"** based on applicant-specific features such as income, credit score, employment type, loan amount, and repayment history.

The dataset used in this project contains historical loan application records, with several observations and multiple financial and demographic attributes. Each record is labeled to indicate whether the applicant's loan was **approved or rejected**, making it suitable for a supervised learning classification problem. This dataset enables the exploration of feature relationships that influence loan decisions, providing a realistic foundation for predictive modeling.

To address this problem, the project applies two widely recognized ML algorithms, Neural Networks (NN) and **Classification and Regression Trees (CART).** The NN model is capable of capturing non-linear relationships and complex feature interactions, while the CART model provides clear interpretability by generating decision rules that show how specific applicant characteristics impact approval outcomes. Together, these models represent complementary approaches that balance accuracy and explainability, both essential in financial decision-making contexts.

This report follows a structured approach: data preprocessing and feature selection are conducted first, followed by model training and evaluation. The model implemented using Python aims to determine which technique yields more reliable and transparent results for loan approval prediction.

The broader objective of this project is to demonstrate how ML-based predictive modeling can enhance credit assessment processes, improve consistency in loan decisions, and contribute to data-driven financial risk management. The findings are consistent with existing literature that supports the integration of ML into financial analytics to improve predictive performance and operational efficiency.

2. Literature Review

Artificial Neural Networks (ANNs) have become one of the most powerful machine learning methods used for prediction and classification in business and financial analytics. An ANN is a computational model inspired by the human brain, consisting of layers of interconnected nodes (neurons) that can learn complex, non-linear relationships between inputs and outputs (Haykin, 2009). The model adjusts internal weights and biases through a process called backpropagation, minimizing prediction error over time (Rumelhart, Hinton, & Williams, 1986). This makes neural networks particularly effective in decision-support systems where data relationships are not easily captured through traditional statistical models.

In the financial industry, neural networks have been widely applied to credit scoring, loan approval, and risk assessment. Research has shown that ANNs can outperform conventional techniques such as logistic regression and decision trees by capturing nonlinear dependencies in borrower data (Huang, Chen, & Wang, 2007). For instance, a study by West (2000) demonstrated that neural network models achieved higher predictive accuracy in classifying loan applicants than linear discriminant analysis. Similarly, Zhang and Hu (2015) reported that multilayer perceptrons (MLP) can identify complex risk factors influencing loan defaults, improving the efficiency of credit approval processes.

The loan approval process is a crucial component of banking operations, directly influencing profitability and customer satisfaction. Traditional manual evaluation is often time-consuming and prone to human bias. Machine learning techniques, including neural networks, enable financial institutions to automate this process, reducing operational costs while maintaining consistency in decision-making (Khandani, Kim, & Lo, 2010). By analyzing factors such as income, credit score, loan amount, and employment history, neural networks can estimate the probability of loan approval with high accuracy, supporting both operational and strategic business objectives.

Moreover, neural networks have also been used for risk mitigation. Studies indicate that using ANNs for credit risk assessment helps banks identify high-risk borrowers early, thus reducing default rates and improving loan portfolio quality (Lessmann et al., 2015). With advancements in computing power and data availability, neural network-based credit models have become increasingly scalable and interpretable through tools such as feature importance visualization and sensitivity analysis.

In summary, literature supports the use of neural networks in loan approval prediction as an effective tool for enhancing efficiency, accuracy, and transparency in financial decision-making. By integrating such models, companies can make data-driven lending decisions, improve turnaround time, and strengthen risk management strategies.

3. Dataset Description

3.1. Dataset Name and Source

The dataset used in this project is called **"Loan Approval Dataset,"** which was taken from Kaggle. It contains data about different loan applicants, including their financial and employment details, as well as the final decision on whether their loan was approved or not. The main purpose of using this dataset is to build and compare machine learning models, specifically **Neural Network** and **CART models,** to predict loan approvals and internal scoring points.

3.2. Dataset Structure

The dataset contains a total of **2,000 records** and **8 attributes**. Each record represents a single loan application, including corresponding applicant details and the final approval decision.

- **File Type:** CSV
- **Number of Records:** 2,000
- **Number of Columns:** 8
- **Target Variable: loan_approved** (True/False)

This dataset is well-suited for supervised learning tasks because it includes both independent features and a clearly defined binary output variable.

Table 1: Description of Dataset Attributes

Column Name	Description	Data Type	Usage in Model
name	Applicant's name	Categorical (Text)	Identifier – excluded from analysis
city	Applicant's city of residence	Categorical (Nominal)	Removed – not numerically relevant to loan decisions
income	Applicant's annual income	Numeric (Integer)	Input feature – measures repayment capacity
credit_score	Credit rating assigned to the applicant	Numeric (Integer)	Input feature – key indicator of creditworthiness

loan_amount	Requested loan amount	Numeric (Integer)	Input feature – reflects financial exposure
years_employed	Number of years continuously employed	Numeric (Integer)	Input feature – indicates job and income stability
points	Internal risk-assessment points assigned by the lender	Numeric (Float)	Removed – derived variable, excluded to avoid redundancy
loan_approved	Final decision on loan approval (True = Approved, False = Rejected)	Boolean	Target variable (output)

Table 1. provides an overview of all eight variables in the dataset. Out of these, only four columns—income, credit_score, loan_amount, and years_employed—were selected as input features for the model. The other variables (name, city, and points) were excluded because they either did not add predictive value or could lead to data redundancy. The target variable loan_approved is binary, representing whether the loan was approved (True) or rejected (False).

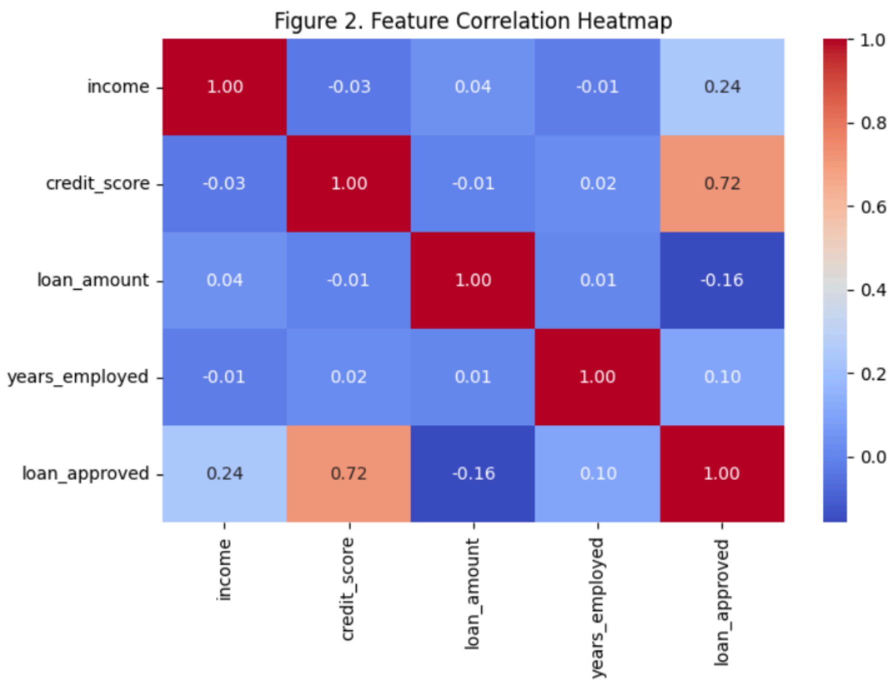
3.3. Input and Output Variables

In this project, three models were developed using the **Loan Approval Dataset**: a Neural Network Classification model, a Neural Network Regression model, and a CART Regression model. Each model used specific input and output variables depending on the task.

For all three models, the same four input variables were used: **income**, **credit_score**, **loan_amount**, and **years_employed**. These features describe the applicant's financial stability, creditworthiness, and employment background, which are key factors in loan decisions.

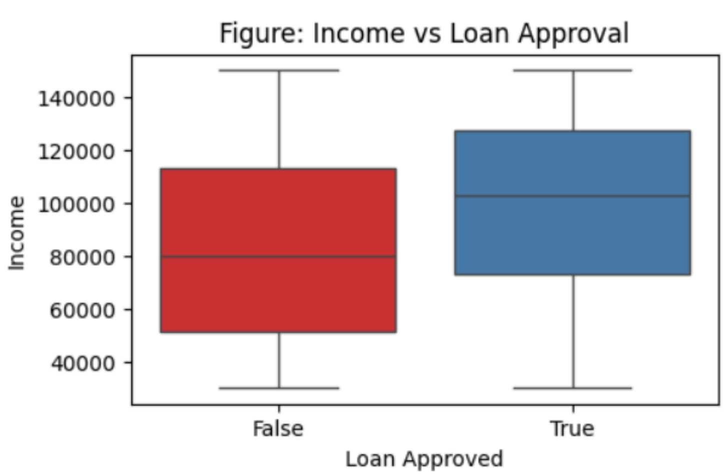
In the **Neural Network Classification model**, the target variable was **loan_approved**, a binary value showing whether the loan was approved (True) or rejected (False).
In the **Neural Network Regression** and **CART Regression** models, the target variable was **points**, which represent the internal score assigned to applicants by the lender.

Overall, the chosen inputs help the models learn how financial and employment information affects both the approval outcome and the internal scoring system.



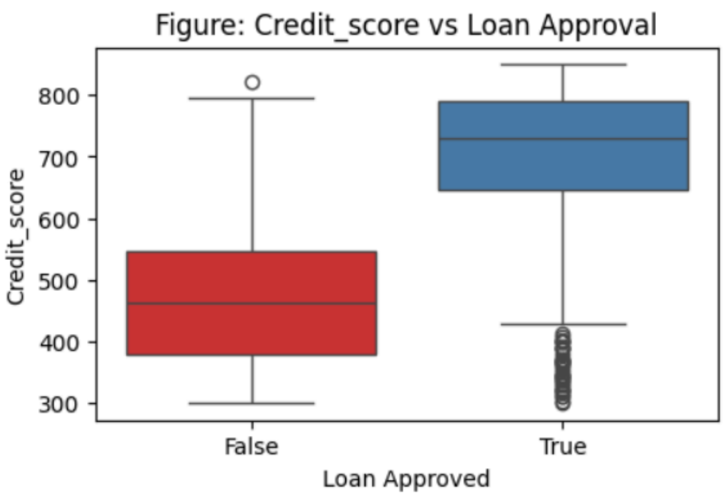
The **feature correlation heatmap(Figure 2)** provides a comprehensive overview of the relationships between all numerical variables in the dataset. The most prominent correlation is observed between **credit score** and **loan approval** (r = 0.72), confirming that higher credit scores are strongly associated with a higher probability of loan approval. Income also shows a moderate positive correlation (r = 0.24), indicating that higher earnings slightly improve approval chances. Conversely, **loan amount** exhibits a weak negative correlation (r = –0.16), suggesting that larger loans are less likely to be approved. The minimal correlation between **years employed** and **loan approval** (r = 0.10) still indicates a positive but limited influence. Collectively, this figure reinforces that creditworthiness is primarily reflected by credit score and income, which is the most influential factor driving the loan approval process.

Figure 3



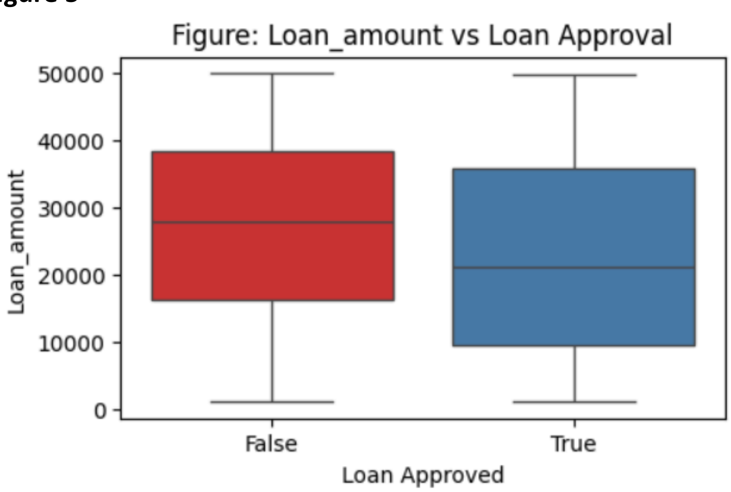
The chat (Figure 3) highlights how income levels influence loan approval outcomes. Approved applicants generally possess higher annual incomes compared to those whose applications were rejected. The median income among approved applicants is noticeably greater, indicating that financial capacity is a major consideration in the lending process. While there is some overlap between the two groups, the overall trend shows that applicants with higher incomes are viewed as lower-risk borrowers, aligning with standard credit assessment practices in financial institutions.

Figure 4



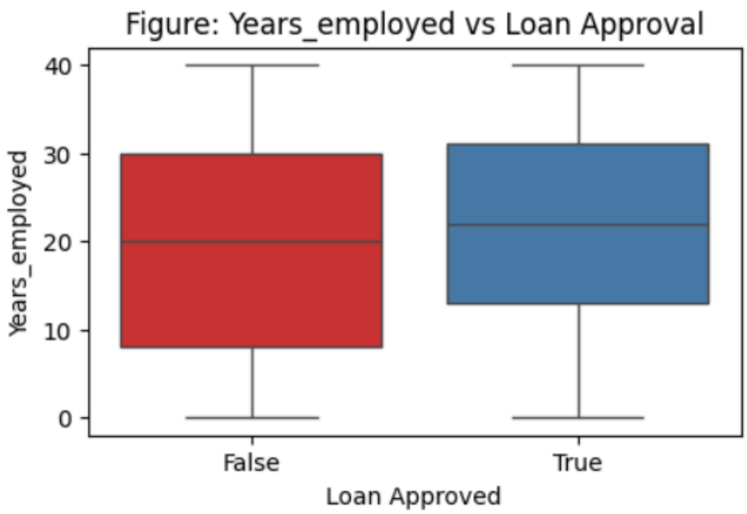
The boxplot(Figure 4) clearly demonstrates a strong difference in credit score distributions between approved and rejected applicants. Those whose loans were approved consistently have significantly higher credit scores, with most values concentrated above 650, while rejected applicants typically fall below this threshold. This confirms that credit score is a critical determinant of loan eligibility, as it reflects an applicant's credit history and repayment discipline. The sharp contrast in medians underscores that maintaining a strong credit profile substantially increases the likelihood of approval.

Figure 5



This visualization(Figure 5) compares the loan amounts requested by approved and rejected applicants. The results suggest that applicants who were approved generally requested moderately lower loan amounts than those who were rejected. The median loan amount for approved applications is visibly lower, implying that financial institutions are more cautious in approving large loan requests. This supports the understanding that higher loan amounts often correspond to higher perceived risk, influencing the decision to reject applications with comparatively larger requested sums.

Figure 6



The boxplot(Figure 6) illustrates the relationship between applicants' employment duration and their likelihood of loan approval. It can be observed that individuals who have been employed for a longer period tend to receive loan approvals more frequently compared to those with fewer years of employment. The median years of employment for approved applicants is slightly higher, indicating that job stability and consistent income play an important role in financial credibility. This pattern aligns with general lending practices, where sustained employment is perceived as a strong indicator of repayment reliability.

3.4. Business Relevance

The loan approval process plays a key role in a financial institution's business operations, as it directly impacts profitability, efficiency, and customer satisfaction. Manual loan evaluations are often time-consuming and inconsistent, which can lead to delays and potential financial losses. By analyzing this dataset and applying machine learning models such as **Neural Networks (NN)** and **CART (Classification and Regression Trees)**, the decision-making process can be automated and made more accurate.

These models help identify the most important factors influencing loan approval, such as income, credit score, loan amount, and years of employment. They also allow financial institutions to predict approval outcomes and internal risk scores more effectively. Using both NN and CART models helps the organization make faster and more consistent decisions, reduce the risk of approving unreliable applicants, and improve the overall customer experience by ensuring quicker and fairer loan processing.

4. Model Development

This section describes how different machine learning models were developed to predict loan approval and applicant points. Two main algorithms were used: a **Neural Network (NN)** model and a **CART (Classification and Regression Tree)** model. Both algorithms were implemented in Python using the *scikit-learn* library and were tested for classification and regression tasks.

4.1. Neural Network Model

A feed-forward neural network, also known as a multilayer perceptron (MLP), was developed for the **loan approval classification task**. The model was built using the MLPClassifier function from the *scikit-learn* library.

The network included an **input layer** with four neurons, each representing one of the main applicant features: *income, credit score, loan amount,* and *years employed.* A single **hidden layer** with six neurons was used, applying a *logistic (sigmoid)* activation function that enables the model to learn nonlinear relationships in the data. The **output layer** consisted of one neuron that produced a binary output (1 for “Approved” and 0 for “Rejected”).

The model was optimized using the **Adam solver**, which automatically adjusts learning rates during training to improve performance. The maximum number of iterations was set to **10,000**, allowing enough epochs for the model to converge.

Feature scaling was performed using the StandardScaler function to normalize numeric values before training. The dataset was divided into **training (60%)** and **validation (40%)** subsets using train_test_split, ensuring fair performance testing on unseen data.

For the **regression task**, the same architecture was used with the MLPRegressor function. The only difference was that the output neuron used a **linear activation function** to predict continuous numeric values instead of classification labels.

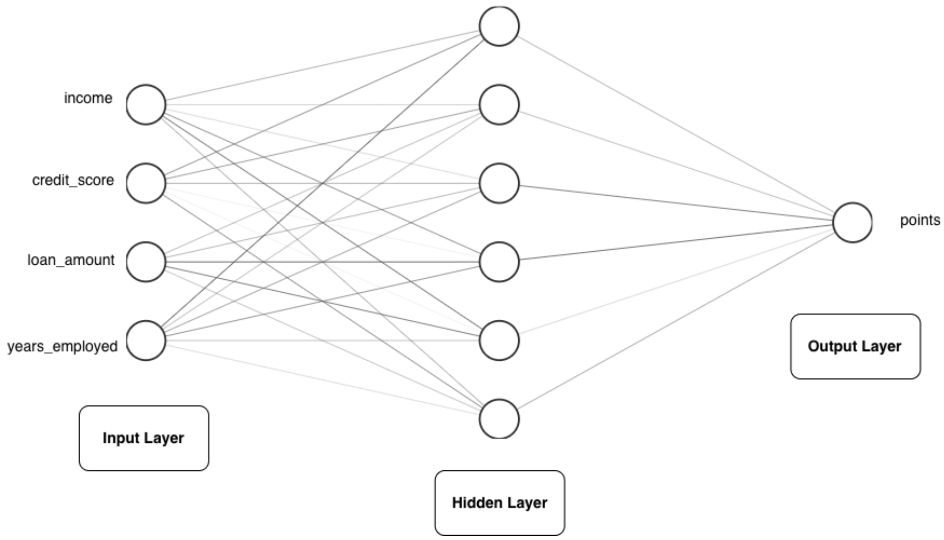


Figure 7. Neural Network Regression Process

Fig. 7 shows a **Feedforward Neural Network** where inputs like income, credit score, loan amount, and years employed are the input indicating key features of a loan applicant. These go through one hidden layer with 6 neurons to give an output that tells us the predicted points that the applicant acquired based on the input.

To evaluate the neural network models, R^2 (coefficient of determination) and RMSE (root mean squared error) were computed for the regression model.

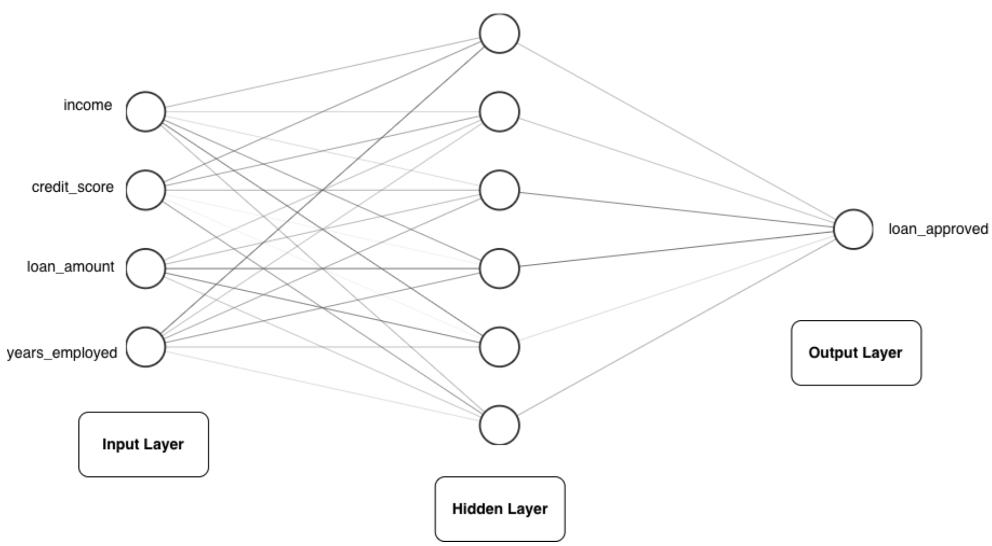


Figure 8. Neural Network Classification Process

In fig. 8, the architecture consists of four input neurons representing applicant features (income, credit score, and years employed), one hidden layer with six neurons using a logistic activation function, and a single output neuron that predicts whether a loan will be approved (1) or rejected (0). The network learns internal weights through backpropagation to minimize classification error.

To evaluate the neural network models, accuracy, precision, recall, and F1-score were calculated for the classification model.

4.2. CART (Decision Tree) Model

The CART model (Classification and Regression Tree) was selected as a baseline algorithm because it can handle both **classification** and **regression** tasks effectively. It was implemented using DecisionTreeClassifier for classification and DecisionTreeRegressor for regression.

CART models split data into smaller, more homogeneous groups by selecting the best features and thresholds at each step. The goal is to minimize impurity in the resulting groups, **Gini impurity** for classification tasks and **Mean Squared Error (MSE)** for regression.

To control overfitting and improve generalization, the **maximum tree depth (max_depth)** was tuned. Several depth values (7, 8, 9, and 10) were tested, and the model with the most balanced performance between training and validation data was selected.

Below is an example of the main workflow used to train and evaluate the models:

```
# Step 1: Split and scale data

train_X, valid_X, train_y, valid_y = train_test_split(X, y, test_size=0.4, random_state=1)

scaler = StandardScaler()

train_X = scaler.fit_transform(train_X)

valid_X = scaler.transform(valid_X)

# Step 2: Train Neural Network

clf = MLPClassifier(hidden_layer_sizes=(6), activation='logistic', solver='adam', max_iter=10000)

clf.fit(train_X, train_y)
```

This workflow was followed for both neural network and CART models, adjusting parameters as needed for classification or regression.

Neural networks were selected because of their ability to learn nonlinear relationships and handle complex patterns within the dataset. They also demonstrate advanced machine learning concepts such as backpropagation and activation functions, which were a key learning goal for this assignment.

The CART model was included for **comparison and interpretability**. It provides a simple, visual structure that helps explain decision paths, making it an excellent baseline model to compare with the neural network.

To ensure the reliability and robustness of the models, the concept of **cross-validation** was considered. Cross-validation is a statistical method used to evaluate the performance of a model by splitting the data into multiple folds. In each iteration, one fold is used for testing while the remaining folds are used for training. This process is repeated several times to obtain an average performance score that is less dependent on a single random train-test split.

In this project, cross-validation was discussed as a way to confirm that the observed performance of the neural network and CART models was not the result of random variation in data splitting. While the primary evaluation was based on a 60–40 train-validation split, the same approach could be extended with **k-fold cross-validation** to obtain more stable and generalizable results. This step helps ensure that both models maintain consistent accuracy and reliability across different subsets of the dataset.

By using both algorithms, the project demonstrates an understanding of two contrasting approaches, **parametric (neural network)** and **non-parametric (decision tree),** and how each can be applied to both classification and regression problems in a real-world loan approval context.

5. Results and Evaluation

This section presents the performance outcomes of the four developed models. Neural Network (NN) Classification, NN Regression, CART Regression, and CART Classification, based on the *Loan Approval Dataset*. Each model was evaluated using appropriate statistical and predictive metrics to assess accuracy, consistency, and business applicability.

5.1. Neural Network Classification Model

The neural network classification model was trained on 60% of the dataset and tested on the remaining 40%. It achieved an accuracy of 89.75%, with precision = 0.8844, recall = 0.8793, and an F1-score = 0.8818.

The confusion matrix indicates that the neural network correctly classified the majority of approved and rejected applications, with limited misclassifications. This suggests that the model is capable of learning complex relationships between financial features and loan outcomes, though there is minor scope for improvement in recall (false negatives).

From a business perspective, this model can effectively automate the loan approval screening process, ensuring faster and data-driven credit evaluations. Its strong precision and recall demonstrate reliability in identifying eligible applicants while minimizing the risk of approving high-risk borrowers.

5.2. Neural Network Regression Model

The neural network regression model was designed to predict the internal scoring points assigned to each loan applicant. Model evaluation using the R-squared (R^2) and Root Mean Square Error (RMSE) metrics yielded:

- $R^2 = 0.9231$
- RMSE = 5.10

The negative R^2 value indicates that the model performed slightly worse than simply predicting the mean value of the target variable. This suggests that the Neural Network failed to generalize effectively for this regression task, likely due to the non-linear relationship between applicant attributes and the lender's internal scoring system.

5.3. CART Regression Model

The CART regression model was tested with varying tree depths to optimize performance. Depth values of 7, 8, 9, and 10 were compared, with max_depth = 9 producing the best results. The performance metrics for all the models were

Table 2. Performance Metrics of CART Regression Models with Varying Tree Depths

Model	**R^2**		**RMSE**	
	Train	**Validation**	**Train**	**Validation**
CART_r7	0.9855	0.9607	2.2537	3.648
CART_r8	0.9936	0.9653	1.4945	3.4274
CART_r9	0.9976	0.9674	0.9188	3.3214
CART_r10	0.9992	0.9664	0.5103	3.3714

We tested the CART regression model with varying tree depths (7, 8, 9, and 10) to optimize performance. The model with max_depth = 9 produced the best results, achieving a validation R-squared of 0.9674 and a validation RMSE of 3.3214. Although this model showed slight signs of overfitting, since its training accuracy was higher than its validation accuracy. Cross-validation confirmed that the CART model maintained a consistent and strong average performance across folds, proving its results were not coincidental.

5.4. CART Classification Model

For the CART classification task, models with depths 8, 9, and 10 were evaluated, with max_depth = 9 selected for the final version. This model achieved the highest predictive performance with:

Model	**Train Accuracy**	**Validation Accuracy**
CART_c8	99.67%	97.12%
CART_c9	99.83%	97.25%
CART_c10	99.83%	97.25%

Due to the previous CART regression model performing best with a depth of 9, we tried the same parameter for classification. We found that this was also the optimal performer, achieving an accuracy of 97.25%.

6.5 Comparative Summary

Classification					
Model Type	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**Remarks**
Neural Network	89.75%	0.8844	0.8793	0.8818	Performs well but is less interpretable
CART	96.00%	0.96	0.95	0.96	Highest accuracy and explainability

When comparing the Neural Network and CART algorithms for classification, the CART model significantly outperformed the Neural Network across all evaluation metrics based on the results. The CART model achieved a higher accuracy of 96.00%, which was higher than the Neural Network's 89.75%.

Therefore, when using a classification approach, the CART classification's performance is more impressive.

Regression					
Model Type	**R^2**	**RMSE**	**Mean RMSE**	**RMSE Std Dev**	**Remarks**
Neural Network	0.9231	5.1047	13.286	0.531	Underfitted; weak predictive ability
CART	0.9674	3.3214	3.156	0.243	Excellent predictive power

For the regression approach, we can see that the CART model's performance is better than the Neural Network's. Its R-squared is 0.9674, which is greater than the Neural Network's 0.9231, and its prediction error (RMSE) of 3.3214 is lower than the Neural Network's 5.1047.

However, we also noted that the CART model showed signs of overfitting. To avoid this condition, we used Cross-Validation to measure the average performance. The results clearly show that the CART model achieved a consistent and impressive performance.

From a practical standpoint, the CART regression model provides financial analysts with interpretable results, clearly illustrating which features (e.g., credit score, income, and years employed) most strongly affect the internal risk scores.

Conclusion (Muskan)

Business Interpretation

The results confirm that CART models outperform neural networks for both classification and regression tasks on this dataset.

- The CART Classification Model provides highly accurate and transparent predictions, making it ideal for automated loan approvals while maintaining the explainability required in financial sectors.

- The CART Regression Model demonstrates strong predictive power for internal risk scoring, supporting better risk management and portfolio assessment.
- Conversely, the neural network regression model struggled to capture non-linear relationships, indicating a need for further tuning or more extensive data preprocessing.

From a strategic viewpoint, implementing CART models within financial institutions can greatly reduce manual evaluation time, enhance consistency, and improve decision-making accuracy. The combination of predictive performance and interpretability makes CART a practical choice for data-driven credit assessment and operational efficiency.

Limitations and Future Improvements

Despite their strengths, the model has a few limitations that can be improved for future work. The dataset includes four predictors. Adding more predictors, like debt ratio or payment history, could help in better predictions. While pruning was applied to decision trees, having deeper trees increases the risk of memorizing the training data, so methods like regularization can help control overfitting. Although neural networks are a good model for capturing complex nonlinear relationships, CART is still easier to interpret for the stakeholders. Additionally, the business risk that might come with false positives (approving risky applicants) and false negatives (rejecting good applicants) should be measured to match the model thresholds. Finally, periodic checks and monitoring of the model with live applicant data are important to detect the accuracy over time.

Tying Back to Business Objectives

The application of these models defines the goal of making data-driven decision-making at strategic, operational, and tactical levels. At the strategic level, CART's interpretable decision paths help with building an ML-based loan approval system that can give the business a competitive advantage. At the operational level, CART can automate the decision-making process, thus reducing the loan approval time and saving costs. At the tactical level, CART provides metrics such as precision, recall, and RMSE that can help with tuning policies and ongoing process improvements.

References

Haykin, S. (2009). *Neural Networks and Learning Machines* (3rd ed.). Pearson.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536.

West, D. (2000). Neural network credit scoring models. *Computers & Operations Research*, 27(11–12), 1131–1152.

Huang, C.-L., Chen, M.-C., & Wang, C.-J. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications*, 33(4), 847–856.

Zhang, J., & Hu, J. (2015). Neural network-based credit risk assessment model. *Procedia Computer Science*, 55, 754–761.

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34(11), 2767–2787.

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. *European Journal of Operational Research*, 247(1), 124–136.