

PROJECT REPORT

ON

LOAN ELIGIBILITY PREDICTION SYSTEM



Submitted to:

UTTAM GRADE

Submitted by:

MEHAK MAJEED

Overview

This internship report presents a study on loan eligibility prediction using machine learning techniques. The aim of this project is to develop a model that can accurately predict whether a loan applicant is eligible for a loan based on various features such as income, credit score, employment status, etc. This report outlines the steps involved in the project, the models used, and the results obtained.

Introduction

With the increasing demand for loans in financial institutions, it is crucial to have an efficient system in place to determine the eligibility of loan applicants. Traditional methods often involve manual evaluation, which can be time-consuming and prone to errors. Machine learning offers a promising solution by automating this process and improving accuracy. In this project, we explore the use of logistic regression and decision tree algorithms for loan eligibility prediction.

Steps Involved

1. Data Collection: Gathered data on loan applicants including features such as income, credit score, employment status, loan amount, etc.
2. Data Pre-processing: Cleaned the data by handling missing values, encoding categorical variables, and scaling numerical features.
3. Model Selection: Chose logistic regression and decision tree algorithms for building predictive models.
4. Model Training: Split the dataset into training and testing sets, and trained the models using the training data.
5. Model Evaluation: Evaluated the performance of the models using metrics such as accuracy, precision, recall, and F1-score.

Model Used

1. Logistic Regression: A binary classification algorithm that estimates the probability that a given input belongs to a particular class. It's well-suited for this task as it provides interpretable results and handles linear relationships well.
2. Decision Tree: A tree-like model where an internal node represents a feature, the branch represents a decision rule, and

each leaf node represents the outcome. Decision trees are capable of handling non-linear relationships and are interpretable.

Results and Conclusion

After training and evaluating both models on the loan eligibility dataset, the following results were obtained:

- Logistic Regression:

- Accuracy: 77%

- Decision Tree:

- Accuracy: 81%

The decision tree model outperforms logistic regression in terms of accuracy for predicting loan eligibility. Decision trees are capable of capturing complex relationships between input features and the target variable, leading to higher predictive accuracy. However, it's essential to consider other factors such as interpretability and computational complexity when choosing between different machine learning models for loan prediction tasks. Overall, the successful application of machine learning techniques demonstrates their potential to enhance

the efficiency and accuracy of loan eligibility assessment in the financial industry.