

REPORT

REPORT

Team:

Hrushikesh Reddy T

GopiKrishna N

Title: PREDICTION OF RECEIVING A LOAN

This project aims to leverage machine learning techniques to predict loan approval outcomes. Models including KNearestNeighbors (KNN), Bernoulli Naive Bayes (BNB), Multinomial Naive Bayes (MNB), and Logistic Regression (LR) have been developed using a dataset of customer financial and demographic information. Cross-validation techniques have also been implemented to evaluate model performance. This report summarizes the current progress on the project.

Models Developed The following models have been developed so far:

KNeighborsClassifier Model

- The KNN algorithm classifies data points based on the labels of their nearest neighbors in the feature space. A KNeighborsClassifier model was implemented using scikit-learn.
- K-fold cross validation was utilized to assess performance across multiple train-test splits.

BernoulliNB Model

- The Bernoulli Naive Bayes model assumes features follow a binomial distribution. A BernoulliNB model was built using scikit-learn.
- Cross-validation was again leveraged to gauge performance.

MultinomialNB Model

- The Multinomial Naive Bayes algorithm is suitable for discrete features. A MultinomialNB classifier was implemented.
- Performance was evaluated via cross-validation on unseen test folds.

LogisticRegression Model

- A logistic regression model was developed to predict loan approval probability based on customer attributes in the dataset.
- Here too cross-validation was used to score and analyze the model.

NeighborsClassifier

- Nearest neighbor classifier storing instance data. Cross-validation accuracy exceeds 80%.

Next Steps

- Further tune model hyperparameters to optimize performance.
- Explore ensemble modeling to combine algorithms.
- Critically evaluate predictions to surface insights.
- Expand test data to assess real-world generalization ability.

Results: Reported accuracy of the models

LogisticRegression: 0.955 . ----High Accuracy

MultinomialNB: 0.895

KNeighborsClassifier: 0.96