

Executive Report: Automating Credit Card Approval with Machine Learning

Objective: The goal of this project was to develop an automated credit card approval predictor using machine learning techniques, specifically logistic regression, to assist banks in processing the high volume of applications they receive. Manual review of applications is error-prone and time-consuming, making automation a cost-effective solution to increase efficiency and reduce operational costs.

Data Overview: The dataset used for this project is a subset of the Credit Card Approval dataset from the UCI Machine Learning Repository. It contains various applicant features such as age, income, and credit history, and a target variable indicating whether an application was approved or rejected.

Data Preprocessing:

- Missing Values:** The dataset contained missing values represented by '?'. These were replaced with NaN, and missing values were imputed using the most frequent value for categorical columns and the mean for numerical columns.
- Encoding:** Categorical variables were transformed into numerical values through one-hot encoding.
- Scaling:** Features were standardized using `StandardScaler` to ensure uniformity and improve model performance.

Model Development:

- Logistic Regression** was chosen for its simplicity and effectiveness in binary classification tasks. The dataset was split into 67% training and 33% testing, and the model was trained on the preprocessed training data.
- Hyperparameter Tuning** was conducted using `GridSearchCV` to optimize the `tol` (tolerance) and `max_iter` (maximum iterations) parameters. The best configuration was found to be `max_iter=100` and `tol=0.01`.

Results:

The model was evaluated using a confusion matrix, which showed a high true positive rate and minimal false positives and negatives:

```
[[203  1][ 1 257]]
```

- Accuracy:** The logistic regression model achieved an accuracy of 79.39% on the test set, indicating a strong ability to correctly classify credit card applications.
- Best Hyperparameters:** The optimal values for the model were `max_iter=100` and `tol=0.01`, yielding a training accuracy of 81.81%.

Conclusion: This machine learning model demonstrates the potential to automate the credit card approval process with a high degree of accuracy (79.39%). By reducing the reliance on manual reviews, banks can process applications faster, reduce errors, and enhance operational efficiency. This project lays the groundwork for further exploration into more advanced models and techniques, but it already represents a significant step forward in leveraging AI for financial decision-making.

