



CREDIT CARD DEFAULT PREDICTION

Detailed Project Report



Prepared By:
Binaya Kumar Pradhan

Table of Contents

1. Introduction.....2

2. Project Overview.....2

3. Components.....2

 3.1. Data Ingestion.....2

 3.2. Data Transformation.....2

 3.3. Model Trainer.....2

 3.4. Prediction Pipeline.....2

 3.5. Web Application.....3

4. Interaction.....3

5. Architecture.....3

6. Conclusion.....3

Introduction

The project aims to build a machine learning pipeline for predicting credit card defaulters. This report provides an overview of the project, its components, design, and functionality.

Project Overview

The project involves multiple components:

- **Data Ingestion:** Reads raw data and splits it into train and test sets.
- **Data Transformation:** Preprocesses the data, handles missing values, and encodes features.
- **Model Trainer:** Trains various machine learning models and selects the best-performing one.
- **Prediction Pipeline:** Accepts user input and predicts credit card default status.
- **Web Application:** Provides a web interface for user interaction.

Components

1. Data Ingestion

- Responsible for reading raw data from a CSV file.
- Splits the data into training and testing sets.
- Saves the datasets for future use.

2. Data Transformation

- Handles data preprocessing, including missing value imputation, scaling, and encoding of categorical variables.
- Saves a preprocessor object for use in the prediction pipeline.

3. Model Trainer

- Trains multiple machine learning models on the pre-processed data.
- Evaluates model performance using accuracy, classification reports, and confusion matrices.
- Selects the best model based on accuracy and saves it for predictions.

4. Prediction Pipeline

- Accepts user input via a web interface.
- Loads the preprocessor and trained model.
- Applies preprocessing to user input.
- Predicts credit card default status and returns the result to the user.

5. Web Application

- Provides a web-based interface for users to enter input data.
- Sends user data to the Prediction Pipeline for prediction.
- Displays the prediction result to the user.

Interactions

- Data Ingestion prepares the data for Data Transformation and Model Trainer.
- Data Transformation uses the preprocessor object created by Data Ingestion.
- Model Trainer utilizes the pre-processed data to train and evaluate models.
- Prediction Pipeline loads the preprocessor and trained model for predictions.
- Web Application connects to the Prediction Pipeline for user interaction.

Architecture

- The project follows a modular architecture with clear separation of responsibilities.
- Each component has a specific role in the pipeline.
- The Web Application serves as the user interface to access the prediction functionality.

Conclusion

The project successfully creates an end-to-end machine learning pipeline for credit card defaulter prediction. It provides a user-friendly interface for users to input data and obtain predictions. The project is designed to be flexible, allowing for easy addition of new features or models in the future.