knowledgehut
upGrad

CREDIT CARD FRAUD PREDICTION PROJECT DOCUMENT

TABLE OF CONTENTS

Requirement

Key Components

Solution Design

Source code

Deployment plan

knowledgehut upGrad

Requirement

Please refer to the below links for Project requirements:

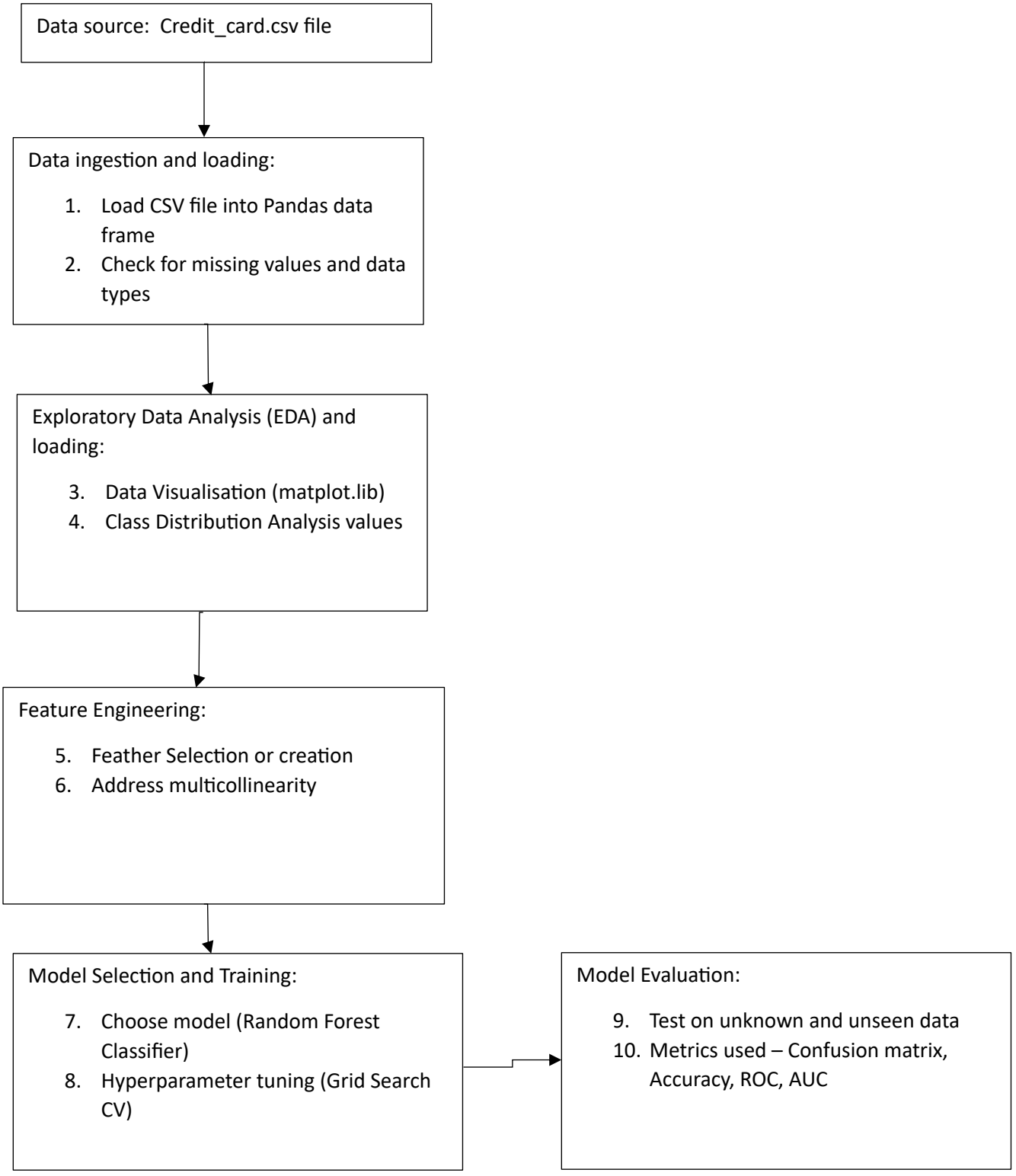
https://github.com/haridoss-annavi/HA-Capstone-project/blob/Credit-Card-Fraud-Prediction/Requirement

Key Components

1. DATA SOURCE: The dataset (creditcard.csv) containing transaction data, including fraudulent and non-fraudulent transactions.
2. DATA INGESTION & LOADING: The CSV file is read into a Pandas DataFrame. Data quality checks, including missing values and data types, are performed.
3. EXPLORATORY DATA ANALYSIS (EDA): Visual and statistical analysis to understand patterns, correlations, and data distribution. Key outputs include class distribution, outlier detection, and correlation matrices.
4. DATA PREPROCESSING: o FEATURE SCALING: Standardize features using StandardScaler. o IMBALANCED DATA HANDLING: Use SMOTE to generate synthetic samples for the minority class. o TRAIN/TEST SPLIT: Split the data into training and testing sets.
5. FEATURE ENGINEERING: Select or create new features and address issues like multicollinearity to improve model performance.
6. MODEL SELECTION & TRAINING: o Select a machine learning model (Random Forest). o Perform hyperparameter tuning using GridSearchCV. o Train the model using the balanced dataset.
7. MODEL EVALUATION: Evaluate the model using metrics like accuracy, confusion matrix, and ROC AUC to ensure the model's robustness.
8. MODEL DEPLOYMENT PLAN: Outline steps to deploy the model, including containerization, serving via an API, and integration into the production environment.

knowledgehut upGrad

Solution Design



knowledgehut upGrad

Source code

Please refer to the GitHub link for Project source code

https://github.com/haridoss-annavi/HA-Capstone-project/blob/Credit-Card-Fraud-Prediction/ha-credit-card-fraud-prediction-rf-smote.ipynb

Installation and Deployment plan

Please refer to the GitHub link for README.md file

https://github.com/haridoss-annavi/HA-Capstone-project/blob/Credit-Card-Fraud-Prediction/README.md