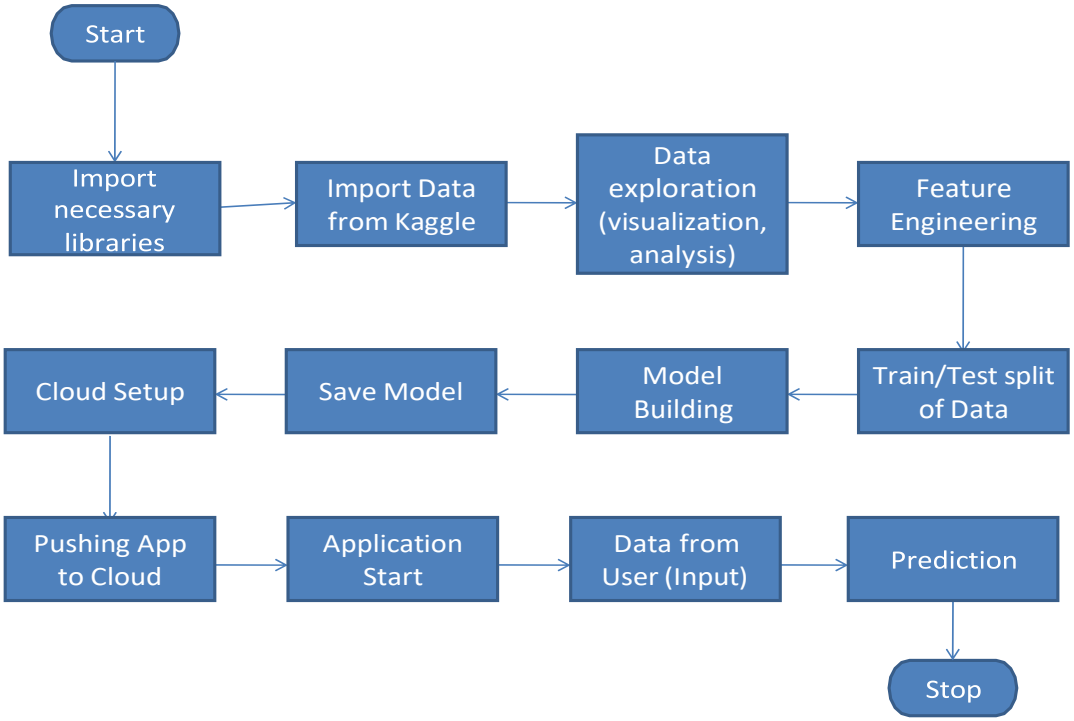
Credit Card Default Prediction

Architecture Document

Report by:
Snehal Kharade
Abhishek Nimje
12 May, 2023

1. Architecture



2. Architecture Description

2.1. Data Description

This dataset is taken from kaggle(url: https://www.kaggle.com/uciml/defaultof-credit-card- clients-dataset). It contains information on default payments, demographic factors, credit data, history of payment, and bill statements of credit card clients in Taiwan from April 2005to September 2005.

2.2. Data Exploration

We divide the data into two types: numerical and categorical. We explore through each typeone by one. Within each type, we explore, visualize and analyze each variable one by one and note down our observations. We also make some minor changes in the data like changecolumn names for convenience in understanding.

2.3. Feature Engineering

Encoded categorical variables.

2.4. Train/Test Split

Split the data into 70% train set and 30% test set.

2.5. Model Building

Built models and trained and tested the data on the models. Compared the performance of each model and selected the best one.

2.6. Save the model

Saved the model by converting into a pickle file.

2.7. Cloud Setup & Pushing the App to the Cloud

Selected Heroku for deployment. Loaded the application files from Github to Heroku.

2.8. Application Start and Input Data by the User

Start the application and enter the inputs.

2.9. Prediction

After the inputs are submitted the application runs the model and makes predictions. The outis displayed as a message indicating whether the customer whose demographic and behavioral data are entered as inputs, is likely to default in the following month or not.