Topic :	CREDIT CARD DEFAULT PREDICTION
Document Type:	High-Level-Design (HLD)
Author :	Pooja Rathore
Date: 05/09/2024	Pw Skills: Full Stack Data Science Pro

1. Introduction:

Credit card default occurs when a cardholder fails to make payments on their credit card debt as agreed. This can be caused by various factors such as job loss, medical crises, or cash flow issues. Defaulting on credit card payments can lead to serious consequences, including damage to credit scores, accumulation of debt due to interest and penalties, and potential legal action from creditors.

Credit card default prediction is crucial because it allows financial institutions to proactively identify customers who are at risk of defaulting. By analyzing various data points such as demographics, payment history, and transaction behaviour, predictive models can assess the likelihood of future defaults. This enables banks to take preventive measures such as offering financial counselling, adjusting credit limits, or restructuring payment plans to reduce the risk of default and minimize financial losses.

2. Problem Statement:

Financial threats are displaying a trend about the credit risk of commercial banks as the incredible improvement in the financial industry has arisen. In this way, one of the biggest threats faced by commercial banks is the risk

prediction of credit clients. The goal is to predict the probability of credit default based on the credit card owner's characteristics and payment history.

3. Dataset Information

This dataset contains information on default payments, demographic factors, credit data, history of payment, and bill statements of credit card clients in Taiwan from April 2005 to September 2005.

About the data Variables:
There are 25 variables:

- ID: ID of each client
- LIMIT_BAL: Amount of given credit in NT dollars(includes individual and family/supplementary credit)
- SEX: Gender
(1=male, 2=female)
- EDUCATION:
(1=graduate school, 2=university, 3=high school, 4=others, 5=unknown, 6=unknown)
- MARRIAGE: Marital status
(1=married, 2=single, 3=others)
- AGE: Age in years
- PAY_0: Repayment status in September, 2005
(-1=pay duly, 1=payment delay for one month, 2=payment delay for two months, ... 8=payment delay for eight months, 9=payment delay for nine months and above)
- PAY_2: Repayment status in August, 2005 (scale same as above)
- PAY_3: Repayment status in July, 2005 (scale same as above)
- PAY_4: Repayment status in June, 2005 (scale same as above)
- PAY_5: Repayment status in May, 2005 (scale same as above)
- PAY_6: Repayment status in April, 2005 (scale same as above)
- BILL_AMT1: Amount of bill statement in September, 2005 (NT-New Taiwan dollar)
- BILL_AMT2: Amount of bill statement in August, 2005 (NT dollar)
- BILL_AMT3: Amount of bill statement in July, 2005 (NT dollar)
- BILL_AMT4: Amount of bill statement in June, 2005 (NT dollar)
- BILL_AMT5: Amount of bill statement in May, 2005 (NT dollar)
- BILL_AMT6: Amount of bill statement in April, 2005 (NT dollar)

- PAY_AMT1: Amount of previous payment in September, 2005 (NT dollar)
- PAY_AMT2: Amount of previous payment in August, 2005 (NT dollar)
- PAY_AMT3: Amount of previous payment in July, 2005 (NT dollar)
- PAY_AMT4: Amount of previous payment in June, 2005 (NT dollar)
- PAY_AMT5: Amount of previous payment in May, 2005 (NT dollar)
- PAY_AMT6: Amount of previous payment in April, 2005 (NT dollar)
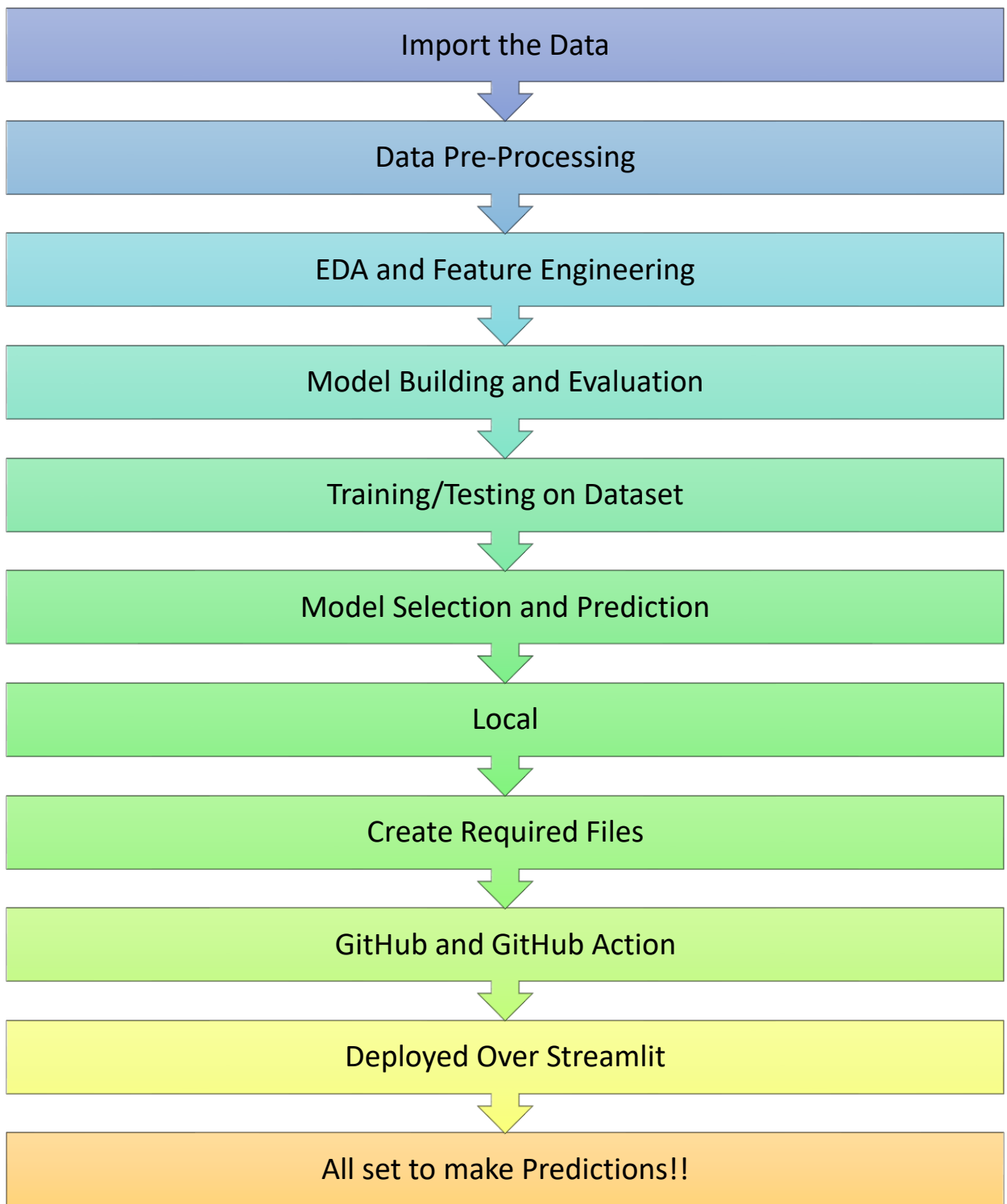- default.payment.next.month: Default payment (1=yes, 0=no)

4. Tools Used:

Python programming language and frameworks such as NumPy, Pandas, Scikit-learn,
Matplotlib, Seaborn and Flask are used to build the whole model.

5. Design Details:

Methodology and Deployment:

Import the Data

↓

Data Pre-Processing

↓

EDA and Feature Engineering

↓

Model Building and Evaluation

↓

Training/Testing on Dataset

↓

Model Selection and Prediction

↓

Local

↓

Create Required Files

↓

GitHub and GitHub Action

↓

Deployed Over Streamlit

↓

All set to make Predictions!!

6. Conclusion:

The project is designed in flask; hence it is accessible to everyone. The above designing process will help banks and loan lenders predict whether customers will default on credit card payments or not, so the bank or respective departments can take necessary action, based on the model's predictions. The UI is made to be user-friendly so that the user will not need much knowledge of any tools but will just need the information for results.