CREDIT CARD DEFAULT PREDICATION

High Level Desing

Ankita Chaudhari

1.INTRODUCTION

There are times when even a seemingly manageable debt, such as credit cards, goes out of control. Loss of job, medical crisis or business failure are some of the reasons that can impact your finances. In fact, credit card debts are usually the first to get out of hand in such situations due to hefty finance charges (compounded on daily balances) and other penalties. A lot of us would be able to relate to this scenario. We may have missed credit card payments once or twice because of forgotten due dates or cash flow issues. But what happens when this continues for months? How to predict if a customer will be defaulter in next months? To reduce the risk of Banks, this model has been developed to predict customer defaulter based on demographic data like gender, age, marital status and behavioral data like last payments, past transactions etc.

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 credit card owner's

characteristics and payment history.

3.DATASET INFORMATION

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 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)

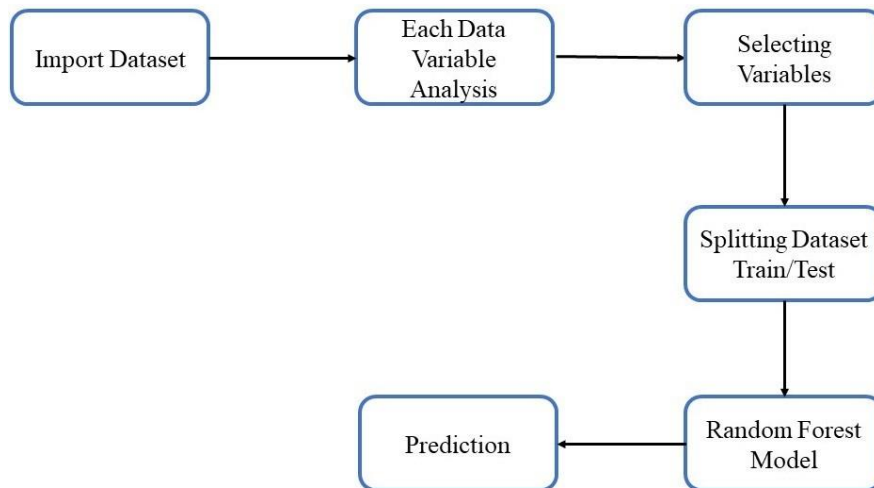**default.payment.next.month:** Default payment (1=yes, 0=no)

4.TOOLS USED

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

5.DESIGN DETAILS

Process flow

```
┌──────────────┐      ┌──────────────┐      ┌──────────────┐
│              │      │  Each Data   │      │              │
│Import Dataset│─────▶│  Variable    │─────▶│  Selecting   │
│              │      │  Analysis    │      │  Variables   │
└──────────────┘      └──────────────┘      └──────────────┘
                                                    │
                                                    ▼
                                            ┌──────────────┐
                                            │Splitting     │
                                            │Dataset       │
                                            │Train/Test    │
                                            └──────────────┘
                                                    │
                                                    ▼
              ┌──────────────┐            ┌──────────────┐
              │              │            │Random Forest │
              │  Prediction  │◀───────────│  Model       │
              │              │            │              │
              └──────────────┘            └──────────────┘
```

CONCLUSION

The above designing process will help banks and loan lenders predict whether customers will default the credit card payment or not, so the bank or respective departments can take necessary action, basedon the model's predictions.