PROJECT HLD

Project Title	Credit Card Default Prediction
Technologies	Machine Learning Technology
Domain	Banking
Project Difficulties level	Intermediate
Submitted By	Atyab Hakeem

1. INTRODUCTION

Banks provide loans and credit cards to their customers, allowing them to make purchases and pay later. However, an increasing number of credit card users are defaulting on their payments, which poses problems for banks in terms of profitability and trust from investors and stakeholders. One solution to this problem is to identify potential credit card defaulters ahead of time and implement measures to mitigate the risk of default.

This can be achieved by using machine learning algorithms to identify potential defaulters before they default. By analyzing the financial history and behavior of credit card users, banks can develop predictive models that can identify customers who are at high risk of defaulting on their payments. Once potential defaulters are identified, banks can take steps to mitigate the risk of default, such as by requiring these customers to provide additional collateral or by imposing stricter limits on their credit card usage. By taking these measures, banks can protect their profitability and maintain the trust of their investors and stakeholders.

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 faces 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 COLUMNS DESCRIPTION

Column 1 - ID: ID of each client

Column 2 - LIMIT_BAL: Amount of given credit in NT dollars (includes individual and family/supplementary = credit)

Column 3 - SEX: Gender (1=male, 2=female)

Column 4 - EDUCATION: (1=graduate school, 2=university, 3=high school, 4=others, 5=unknown, 6=unknown)

Column 5 - MARRIAGE: Marital status (1=married, 2=single, 3=others)

Column 6 - AGE: Age in years

Column 7 - 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)

Column 8 - PAY_2: Repayment status in August, 2005 (scale same as above)
Column 9 - PAY_3: Repayment status in July, 2005 (scale same as above)
Column 10 - PAY_4: Repayment status in June, 2005 (scale same as above)
Column 11 - PAY_5: Repayment status in May, 2005 (scale same as above)
Column 12 - PAY_6: Repayment status in April, 2005 (scale same as above)

Column 13 - BILL_AMT1: Amount of bill statement in September, 2005 (NT dollar)
Column 14 - BILL_AMT2: Amount of bill statement in August, 2005 (NT dollar)
Column 15 - BILL_AMT3: Amount of bill statement in July, 2005 (NT dollar)
Column 16 - BILL_AMT4: Amount of bill statement in June, 2005 (NT dollar)
Column 17 - BILL_AMT5: Amount of bill statement in May, 2005 (NT dollar)
Column 18 - BILL_AMT6: Amount of bill statement in April, 2005 (NT dollar)

Column 19 - PAY_AMT1: Amount of previous payment in September, 2005 (NT dollar)
Column 20 - PAY_AMT2: Amount of previous payment in August, 2005 (NT dollar)
Column 21 - PAY_AMT3: Amount of previous payment in July, 2005 (NT dollar)
Column 22 - PAY_AMT4: Amount of previous payment in June, 2005 (NT dollar)
Column 23 - PAY_AMT5: Amount of previous payment in May, 2005 (NT dollar)
Column 24 - PAY_AMT6: Amount of previous payment in April, 2005 (NT dollar)

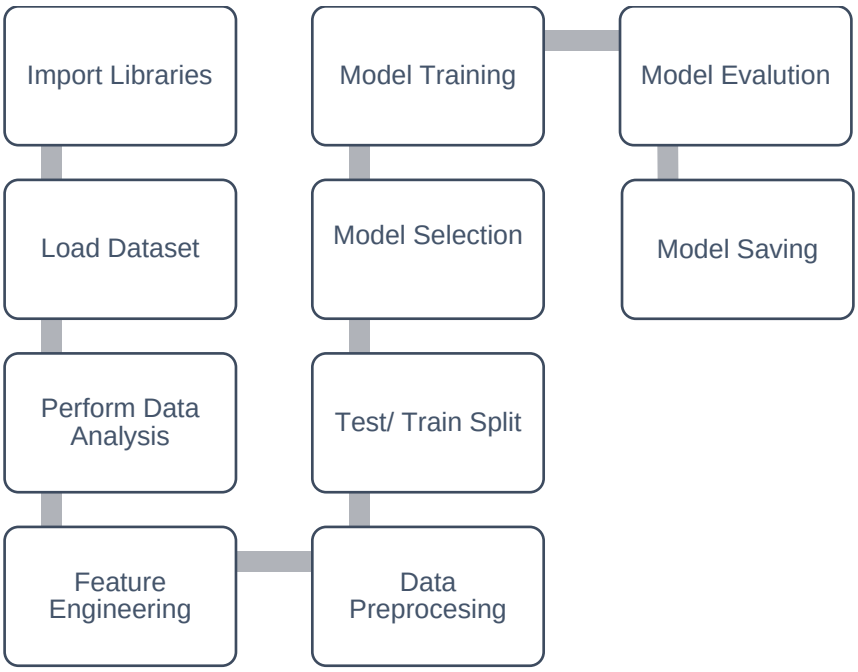
Column 25 - default.payment.next.month: Default payment (1=yes, 0=no)

4. TOOLS USED

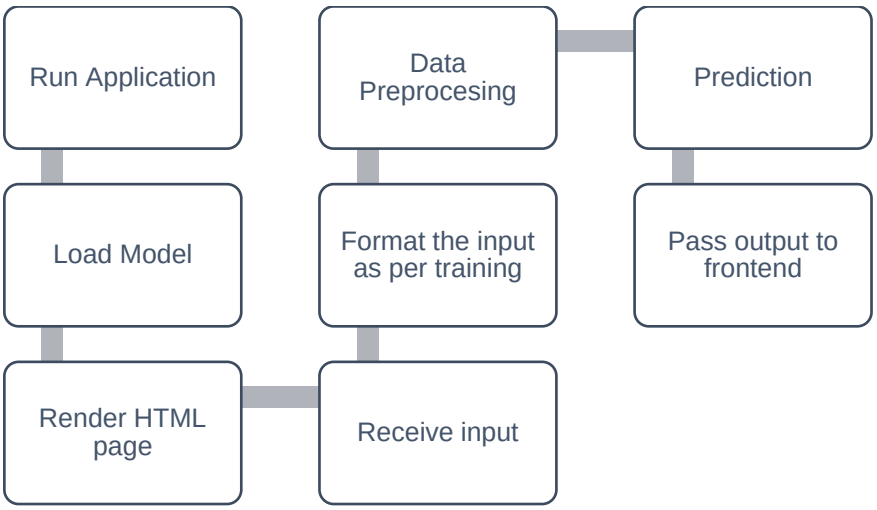
Financial threats are displaying a trend about the credit risk of commercial banks as the incredible improvement in the financial industry has arisen. In this way, one of the biggest threats faces by commercial banks is the risk prediction of credit clients. The goal is to predict the probability of credit default based on credit card owner's characteristics and payment history.

5. DESIGN DETAILS

5.1. Process Flow:



5.2. Deployment Process:



6. CONCLUSION

In conclusion, the web application built using Flask can be deployed on any cloud platform and hence accessible all around the world. With a user-friendly user-interface, the clients can easily predict if their customers are going to default on their next month payment. Armed with this information, the clients can impose restrictions on the suspected customers and/or require them to repay the balance amounts so that the institutions can mitigate as much risk as possible thereby making them less susceptible to incurring high losses.