|  |
| --- |
|  |
| CREDIT CARD DEFAULT PREDICTION |
| FINAL PROJECT |

|  |
| --- |
| P.KIRUTHIKA  112721214009 |

df['IsDefaulter'] =df ['default.payment.next.month']

df.drop('default.payment.next.month',axis = 1)

*# df.rename({'default.payment.next.month' : 'IsDefaulter'}, inplace=True)*

from sklearn.preprocessing import StandardScaler

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LogisticRegression

from sklearn.model\_selection import GridSearchCV

from sklearn.metrics import accuracy\_score, recall\_score, precision\_score, f1\_score, roc\_auc\_score, confusion\_matrix, roc\_curve, auc

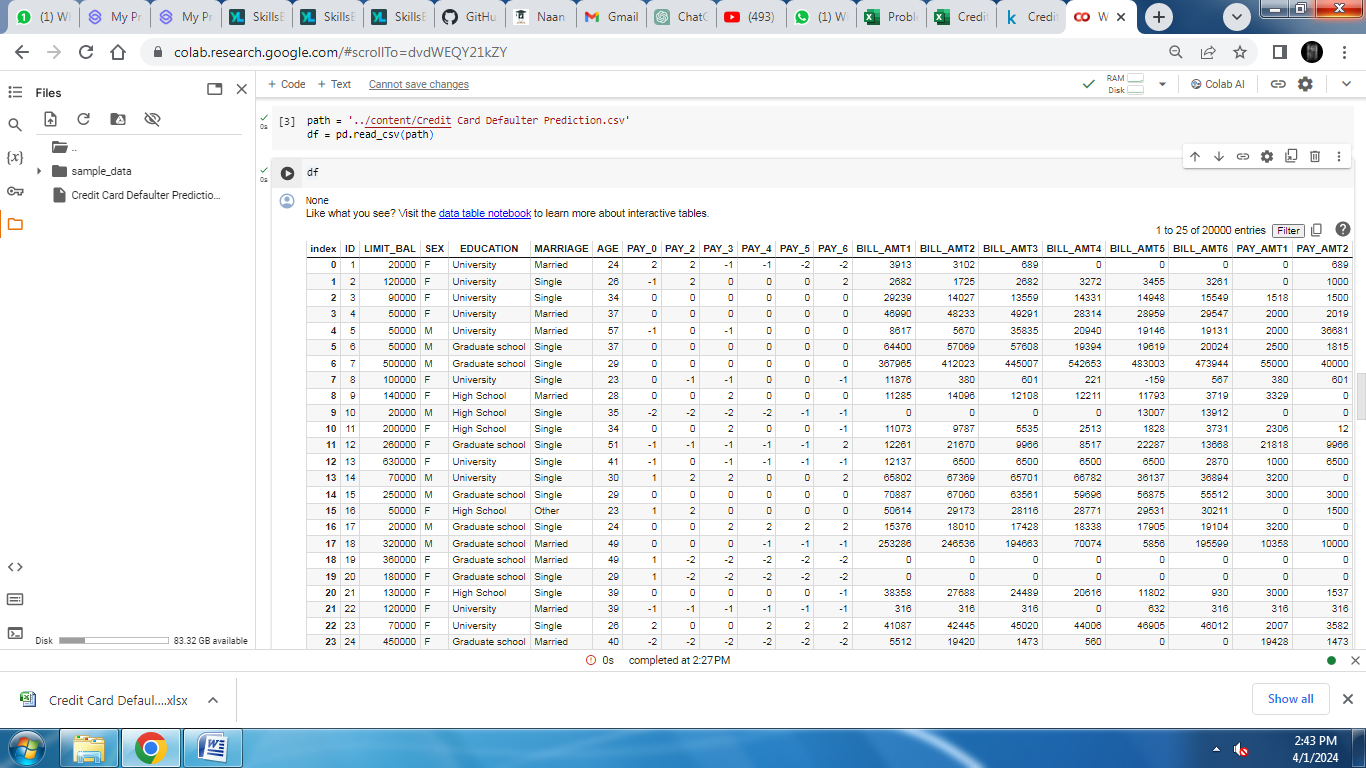
from sklearn.svm import SVC

path = '../content/Credit Card Defaulter Prediction.csv'

df = pd.read\_csv(path)

df

df



from matplotlib import pyplot as plt

df.plot(kind='scatter', x='ID', y='LIMIT\_BAL', s=32, alpha=.8)

plt.gca().spines[['top', 'right',]].set\_visible(False)

df.info()

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 30000 entries, 0 to 29999

Data columns (total 25 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 ID 30000 non-null int64

1 LIMIT\_BAL 30000 non-null int64

2 SEX 30000 non-null object

3 EDUCATION 30000 non-null object

4 MARRIAGE 30000 non-null object

5 AGE 30000 non-null int64

6 PAY\_0 30000 non-null int64

7 PAY\_2 30000 non-null int64

8 PAY\_3 30000 non-null int64

9 PAY\_4 30000 non-null int64

10 PAY\_5 30000 non-null int64

11 PAY\_6 30000 non-null int64

12 BILL\_AMT1 30000 non-null int64

13 BILL\_AMT2 30000 non-null int64

14 BILL\_AMT3 30000 non-null int64

15 BILL\_AMT4 30000 non-null int64

16 BILL\_AMT5 30000 non-null int64

17 BILL\_AMT6 30000 non-null int64

18 PAY\_AMT1 30000 non-null int64

19 PAY\_AMT2 30000 non-null int64

20 PAY\_AMT3 30000 non-null int64

21 PAY\_AMT4 30000 non-null int64

22 PAY\_AMT5 30000 non-null int64

23 PAY\_AMT6 30000 non-null int64

24 default 30000 non-null object

dtypes: int64(21), object(4)

memory usage: 5.7+ MB

df.isnull().sum()

ID 0

LIMIT\_BAL 0

SEX 0

EDUCATION 0

MARRIAGE 0

AGE 0

PAY\_0 0

PAY\_2 0

PAY\_3 0

PAY\_4 0

PAY\_5 0

PAY\_6 0

BILL\_AMT1 0

BILL\_AMT2 0

BILL\_AMT3 0

BILL\_AMT4 0

BILL\_AMT5 0

BILL\_AMT6 0

PAY\_AMT1 0

PAY\_AMT2 0

PAY\_AMT3 0

PAY\_AMT4 0

PAY\_AMT5 0

PAY\_AMT6 0

default 0

dtype: int64

df.describe()

ID 0

LIMIT\_BAL 0

SEX 0

EDUCATION 0

MARRIAGE 0

AGE 0

PAY\_0 0

PAY\_2 0

PAY\_3 0

PAY\_4 0

PAY\_5 0

PAY\_6 0

BILL\_AMT1 0

BILL\_AMT2 0

BILL\_AMT3 0

BILL\_AMT4 0

BILL\_AMT5 0

BILL\_AMT6 0

PAY\_AMT1 0

PAY\_AMT2 0

PAY\_AMT3 0

PAY\_AMT4 0

PAY\_AMT5 0

PAY\_AMT6 0

default 0

dtype: int64