

Untitled13

May 27, 2025

```
[1]: !pip install pandas numpy matplotlib seaborn scikit-learn lightgbm
```

```
Requirement already satisfied: pandas in c:\users\saipa\anaconda3\lib\site-  
packages (2.1.4)  
Requirement already satisfied: numpy in c:\users\saipa\anaconda3\lib\site-  
packages (1.26.4)  
Requirement already satisfied: matplotlib in c:\users\saipa\anaconda3\lib\site-  
packages (3.8.0)  
Requirement already satisfied: seaborn in c:\users\saipa\anaconda3\lib\site-  
packages (0.12.2)  
Requirement already satisfied: scikit-learn in  
c:\users\saipa\anaconda3\lib\site-packages (1.2.2)  
Collecting lightgbm  
  Downloading lightgbm-4.6.0-py3-none-win_amd64.whl.metadata (17 kB)  
Requirement already satisfied: python-dateutil>=2.8.2 in  
c:\users\saipa\anaconda3\lib\site-packages (from pandas) (2.8.2)  
Requirement already satisfied: pytz>=2020.1 in  
c:\users\saipa\anaconda3\lib\site-packages (from pandas) (2023.3.post1)  
Requirement already satisfied: tzdata>=2022.1 in  
c:\users\saipa\anaconda3\lib\site-packages (from pandas) (2023.3)  
Requirement already satisfied: contourpy>=1.0.1 in  
c:\users\saipa\anaconda3\lib\site-packages (from matplotlib) (1.2.0)  
Requirement already satisfied: cycler>=0.10 in  
c:\users\saipa\anaconda3\lib\site-packages (from matplotlib) (0.11.0)  
Requirement already satisfied: fonttools>=4.22.0 in  
c:\users\saipa\anaconda3\lib\site-packages (from matplotlib) (4.25.0)  
Requirement already satisfied: kiwisolver>=1.0.1 in  
c:\users\saipa\anaconda3\lib\site-packages (from matplotlib) (1.4.4)  
Requirement already satisfied: packaging>=20.0 in  
c:\users\saipa\anaconda3\lib\site-packages (from matplotlib) (23.1)  
Requirement already satisfied: pillow>=6.2.0 in  
c:\users\saipa\anaconda3\lib\site-packages (from matplotlib) (10.2.0)  
Requirement already satisfied: pyparsing>=2.3.1 in  
c:\users\saipa\anaconda3\lib\site-packages (from matplotlib) (3.0.9)  
Requirement already satisfied: scipy>=1.3.2 in  
c:\users\saipa\anaconda3\lib\site-packages (from scikit-learn) (1.11.4)  
Requirement already satisfied: joblib>=1.1.1 in  
c:\users\saipa\anaconda3\lib\site-packages (from scikit-learn) (1.2.0)
```

Requirement already satisfied: threadpoolctl>=2.0.0 in
c:\users\saipa\anaconda3\lib\site-packages (from scikit-learn) (2.2.0)
Requirement already satisfied: six>=1.5 in c:\users\saipa\anaconda3\lib\site-
packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Downloading lightgbm-4.6.0-py3-none-win_amd64.whl (1.5 MB)

```

----- 0.0/1.5 MB ? eta -:--:--
----- 0.0/1.5 MB ? eta -:--:--
-  ----- 0.0/1.5 MB 495.5 kB/s eta 0:00:03
-  ----- 0.1/1.5 MB 653.6 kB/s eta 0:00:03
---  ----- 0.1/1.5 MB 1.0 MB/s eta 0:00:02
-----  ----- 0.3/1.5 MB 1.7 MB/s eta 0:00:01
-----  ----- 0.4/1.5 MB 1.8 MB/s eta 0:00:01
-----  ----- 0.6/1.5 MB 2.3 MB/s eta 0:00:01
-----  ----- 0.8/1.5 MB 2.5 MB/s eta 0:00:01
-----  ----- 0.9/1.5 MB 2.6 MB/s eta 0:00:01
-----  ----- 1.1/1.5 MB 2.7 MB/s eta 0:00:01
-----  ----- 1.3/1.5 MB 3.0 MB/s eta 0:00:01
-----  ----- 1.4/1.5 MB 3.1 MB/s eta 0:00:01
-----  ----- 1.5/1.5 MB 3.0 MB/s eta 0:00:00

```

Installing collected packages: lightgbm
Successfully installed lightgbm-4.6.0

```
[2]: import pandas as pd

df = pd.read_csv(r"C:
↪\Users\saipa\OneDrive\Desktop\satsM\codectechnologies\UCI_Credit_Card.csv")
df.head()
```

```
[2]:
```

	ID	LIMIT_BAL	SEX	EDUCATION	MARRIAGE	AGE	PAY_0	PAY_2	PAY_3	PAY_4	\
0	1	20000.0	2	2	1	24	2	2	-1	-1	
1	2	120000.0	2	2	2	26	-1	2	0	0	
2	3	90000.0	2	2	2	34	0	0	0	0	
3	4	50000.0	2	2	1	37	0	0	0	0	
4	5	50000.0	1	2	1	57	-1	0	-1	0	

	...	BILL_AMT4	BILL_AMT5	BILL_AMT6	PAY_AMT1	PAY_AMT2	PAY_AMT3	\
0	...	0.0	0.0	0.0	0.0	689.0	0.0	
1	...	3272.0	3455.0	3261.0	0.0	1000.0	1000.0	
2	...	14331.0	14948.0	15549.0	1518.0	1500.0	1000.0	
3	...	28314.0	28959.0	29547.0	2000.0	2019.0	1200.0	
4	...	20940.0	19146.0	19131.0	2000.0	36681.0	10000.0	

	PAY_AMT4	PAY_AMT5	PAY_AMT6	default.payment.next.month
0	0.0	0.0	0.0	1
1	1000.0	0.0	2000.0	1
2	1000.0	1000.0	5000.0	0
3	1100.0	1069.0	1000.0	0

```
4      9000.0      689.0      679.0      0
```

```
[5 rows x 25 columns]
```

```
[3]: df.rename(columns={'default.payment.next.month': 'default'}, inplace=True)
      print(df['default'].value_counts())
```

```
default
0      23364
1       6636
Name: count, dtype: int64
```

```
[4]: # Total bill and payment amounts
      df['total_bill'] = df[[f'BILL_AMT{i}' for i in range(1, 7)]].sum(axis=1)
      df['total_pay'] = df[[f'PAY_AMT{i}' for i in range(1, 7)]].sum(axis=1)

      # Payment to bill ratio
      df['pay_to_bill_ratio'] = df['total_pay'] / (df['total_bill'] + 1)
```

```
[5]: from sklearn.model_selection import train_test_split

      X = df.drop(['ID', 'default'], axis=1)
      y = df['default']

      X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
      random_state=42)
```

```
[12]: from sklearn.preprocessing import LabelEncoder

      for col in X.columns:
          if X[col].dtype == 'object':
              le = LabelEncoder()
              X[col] = le.fit_transform(X[col])
```

```
[14]: X_train = X_train.apply(pd.to_numeric, errors='coerce')
      X_test = X_test.apply(pd.to_numeric, errors='coerce')
```

```
[ ]:
```