PHASE-3
CREDIT CARD FRAUD DETECTION

Data preprocessing:

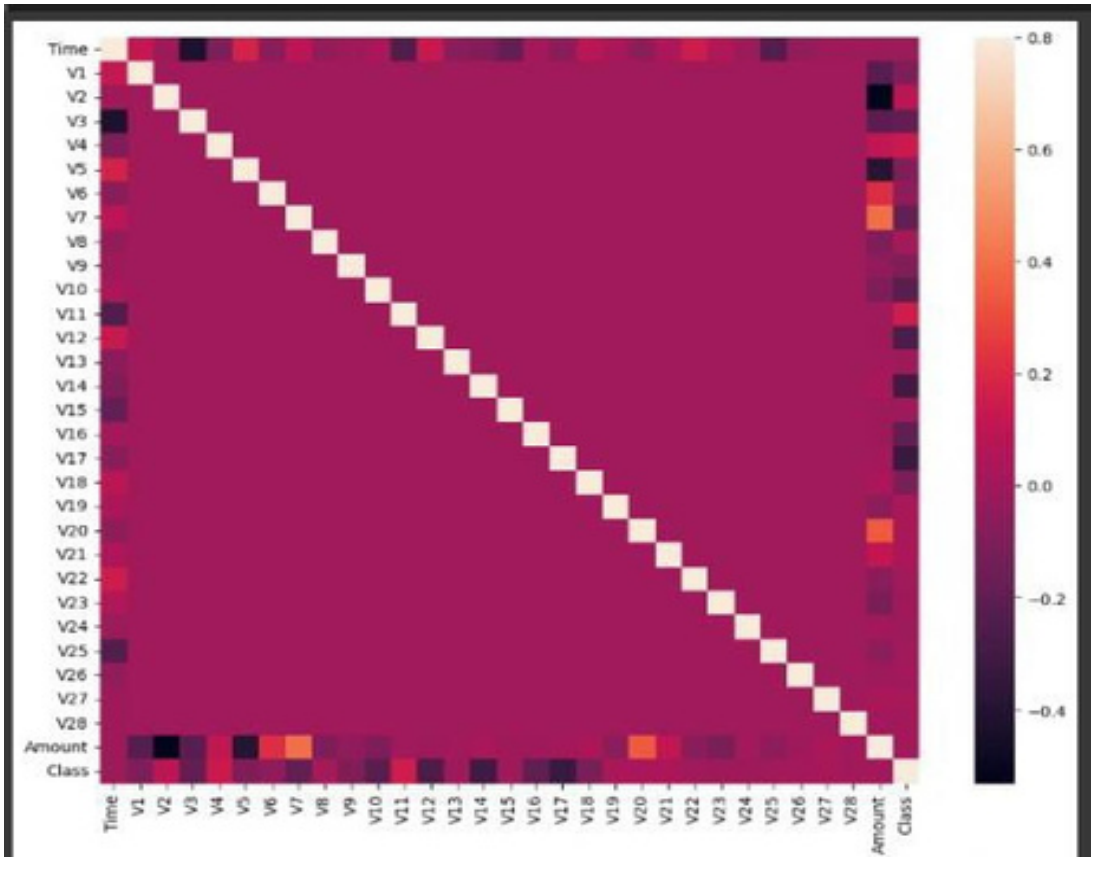
```
from google.colab import drive

drive.mount("/content/drive")


import pandas as pd

data = pd.read_csv('/content/drive/MyDrive/creditcard.csv')

missing_values = data.isnull().sum()

numeric_columns = data.select_dtypes(include=['number']).columns

data[numeric_columns] =
data[numeric_columns].fillna(data[numeric_columns].mean())

missing_values_after = data.isnull().sum()

data.to_csv('preprocessed_dataset.csv', index=False)
```

Program for Data visualization :

```
import numpy as np

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

from sklearn.model_selection import train_test_split
```

```
from imblearn.over_sampling import SMOTE
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import confusion_matrix, classification_report

df=pd.read_csv('/content/drive/MyDrive/creditcard.csv')
df

df.info()

df.describe()

df['Class'].unique()

df['Class'].value_counts()
corr = df.corr()
fig = plt.figure(figsize = (12, 9))
sns.heatmap(corr, vmax = .8, square = True)
plt.show()
```

Visualization:



BY:

R.ABINAYA (abinayaraja1052004@gmail.com)

S.ANBARASI (anbarasideva14@gmail.com)

K.BHUVANA SURUTHI (sagunthalakumar16@gmail.com)

M.GAYATHRI (45gayathri@gmail.com)