**CREDIT CARD FRAUD DETECTION**

**AIM:** TO DETECT THE CREDIT CARD FRAUDS BY USING MACHINE LEARNING.

**ABSTRACT:**

The credit card fraud detection was done in the google colab which detects the frauds from the credit card.

Today the usage of CREDIT CARD is about high.At Least an average of 1.86 billion transactions are done by credit card. So some fraud transactions are taking place.

Here we detect the transactions of credit cards.

**PROGRAM:**

**->** Firstly import the libraries

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt #it is used for plot the graph in matplot

import seaborn as sns #it is used for bins

**->** convert the csv file into dataframe using PANDAS library

dataset = pd.read\_csv('creditcard.csv')

**->**know the shape of the csv file which contains rows and columns.

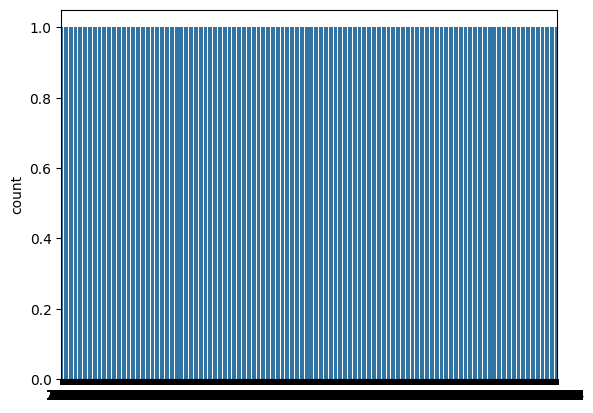
->Check the missing values in the data

dataset.isna().sum(

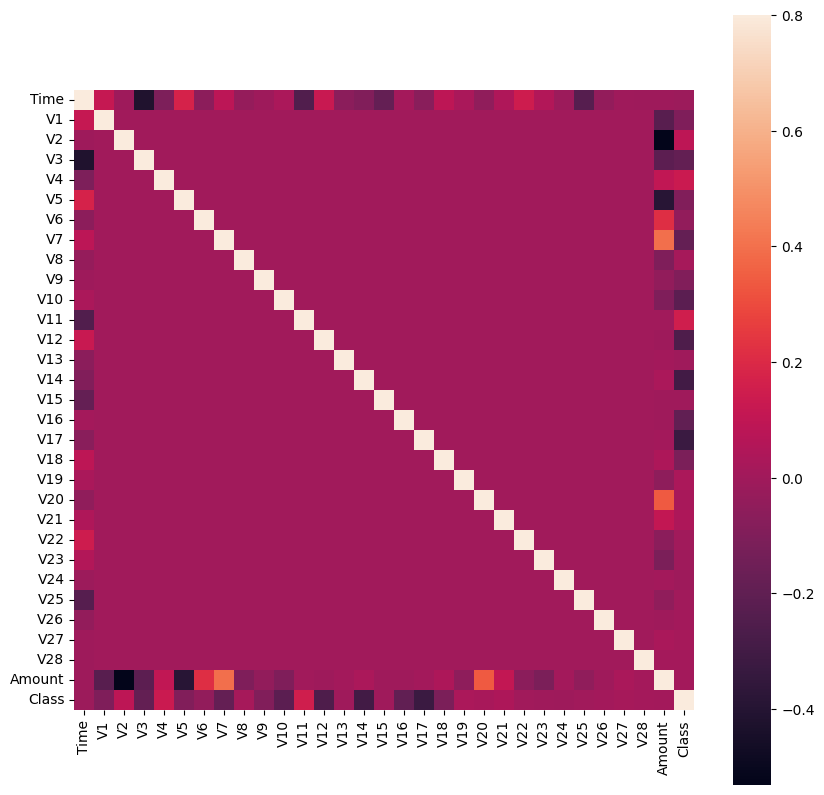
-> Show the first 5 transactions by using head() method

1. Here v1,v2,.,.,...v28 are the properties of the transactions.
2. The time denotes the transaction within the time.

->Now disturbed the normal and fraud transactions.

1. Here 0 is the normal transactions and
2. The 1 denotes the fraud transaction.
3. #distrubtion of real and fraud transcations
4. pd.value\_counts(dataset['Class'])
5. Class
6. 0 284315
7. 1 492
8. Name: count, dtype: int64
9. The data is highly imbalanced.

->Now cleansing the data by using the seaborn method .



->Now separate the data for analytics.

->Compare the values for both transaction.

dataset.groupby('Class').mean()

->Now split the data into train and test for classification

X = dataset.iloc[: , :-1].values

y = dataset.iloc[: , -1].values

from imblearn.over\_sampling import RandomOverSampler

ros = RandomOverSampler(random\_state=0)

x\_res , y\_res = ros.fit\_resample(X,y)

->Now find the shape of X

(284807, 30)

->find the shape of x

(568630, 30)

->from collections import Counter

print(Counter(y))

print(Counter(y\_res))

->from sklearn.model\_selection import train\_test\_split # data preprocessing

x\_train , x\_test , y\_train , y\_test = train\_test\_split(x\_res , y\_res , test\_size=0.3 , random\_state=42)

1. This tells that the data splits into train and test

->from sklearn.ensemble import RandomForestClassifier

classifier = RandomForestClassifier(n\_estimators = 641 , random\_state=0)

classifier.fit(x\_train , y\_train)

1. By using RandomForest we can classify the training and testing data.

->

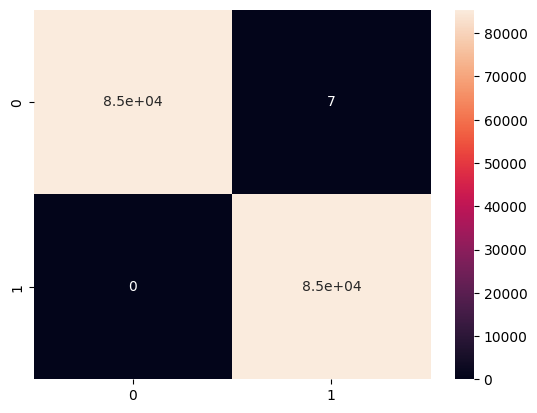
# implemetation report

from sklearn.metrics import confusion\_matrix , accuracy\_score

cm = confusion\_matrix(y\_test , y\_pred)

sns.heatmap(cm , annot=True)

print(accuracy\_score(y\_test , y\_pred))

1. It implements the report by prediction of data.
2. 

->Now find the accuracy point for the classification.

0.9999180778728335

-> precision recall f1-score support

0 1.00 1.00 1.00 85149

1 1.00 1.00 1.00 85440

accuracy 1.00 170589

macro avg 1.00 1.00 1.00 170589

weighted avg 1.00 1.00 1.00 170589

Finally the accuracy data is here the fraud of transactions.

**FINAL REPORT:**

This proves the credit card transactions are some of fraud transaction and we can detect the fraud by using this.