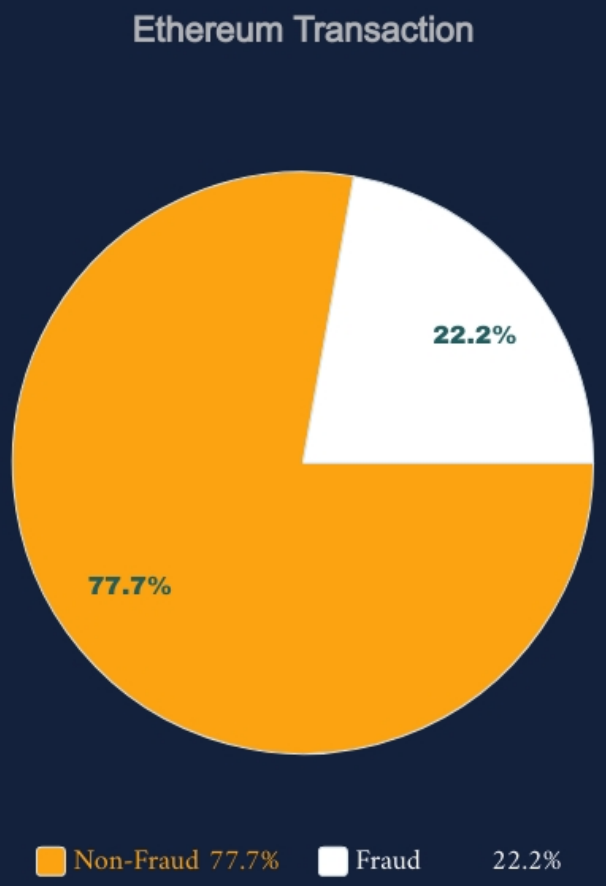
Detecting Fraud Transactions of ETHEREUM

- Akash V

Table of Content

Problem Statement

To find the patterns of transactions performed and help algorithms to learn those patterns in identifying the fradulent transactions and flag them.

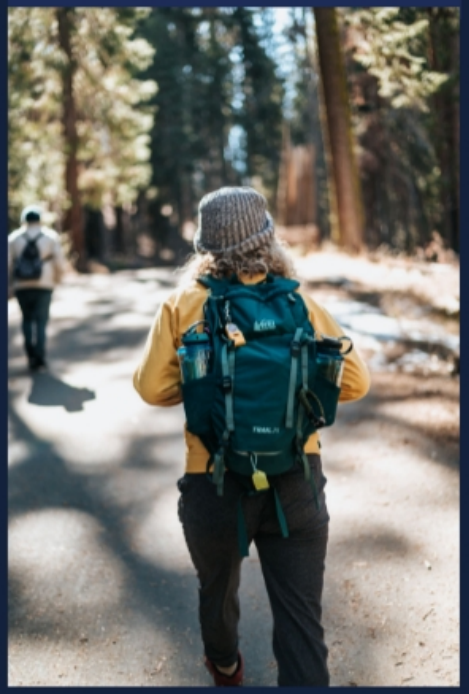


Need of the study

I present my interest in the detection Fraud Transaction of Ethereum, specifically on the Ethereum network, and a structured approach to address the numerous anomalous transactions.

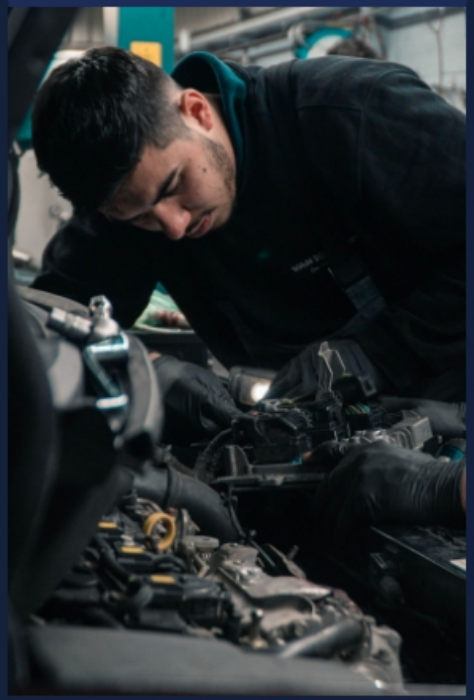
Purpose of this Project

Exploratory analysis of data to extract the pattern of fraudlent activites



Build a machine learning model to classify fraud and non-fraud transactions.

Reduce the error by tuning the model.



Tool

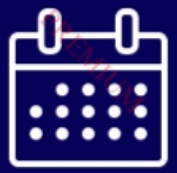
Google
Colab

Library's & Dataset

The dataset as I already download from Kaggle. The dataset is huge and would take a long time to process. Let's see the overview of data

EDA Process

Descriptive Statistics



Start by looking at descriptive statistic parameters for the dataset. We will use describe() for this

Missing value imputation

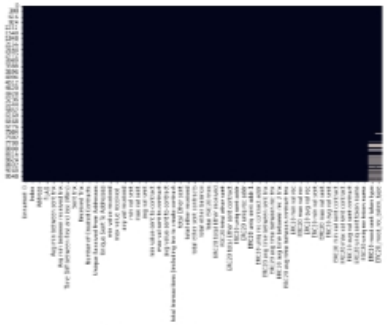
I will check for missing values in our dataset. In case there are any missing entries, we will impute them with appropriate values

Graphical Representation

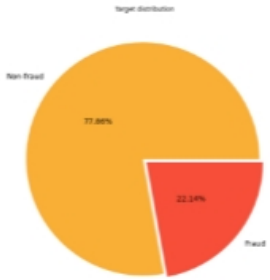
This processing steps is usefull to see the data in graphical view .

Graphical Representation

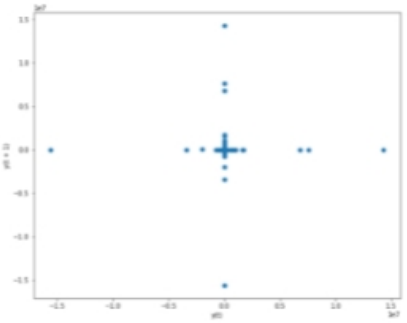
Plots



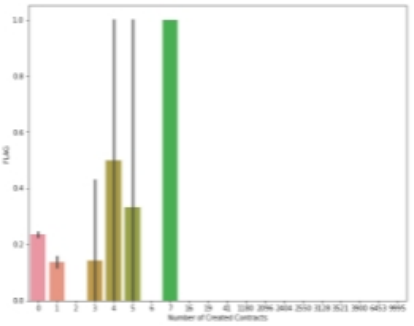
Heat Map



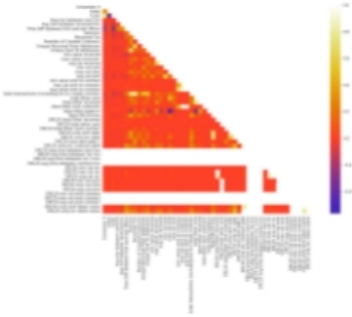
Pie Plot



Lag Plot



Count Plot



Corr Plot

HEAT MAP	PIE PLOT	LAG PLOT	COUNT PLOT	CORR PLOT
As we expected every missing value is in fraud rows. That means that almost 40% of fraud rows have missing values	In this pie plot shows us the Distribution of the Flag(Target Vairable).By looking this plot we have to conclude that the Non-fraud transaction having the most of the distribution in the dataset	The lag plot shows that with many of the points clustered at the corner, showing us that the data has a few non-zero values, but the points are mainly non-zero, and do not follow any trend	By looking this above plot, We can see that the more contacts a user has created, the more likely they are to be of a fraudulent transaction	This corr plot shows that How the features are correlated with other features.By looking this plot we have to conclude that which features are highly correlated with others and lowely correlated with others

Pre-Processing the Data

Feature Selection

Treat Outliers

Normalization

Balancing the Dataset

Supervised Learning Models

LogisticRegression

Random Forest Classifier

XGB Classifier

K-Nearest Neighbors

LGBMClassifier

SVC

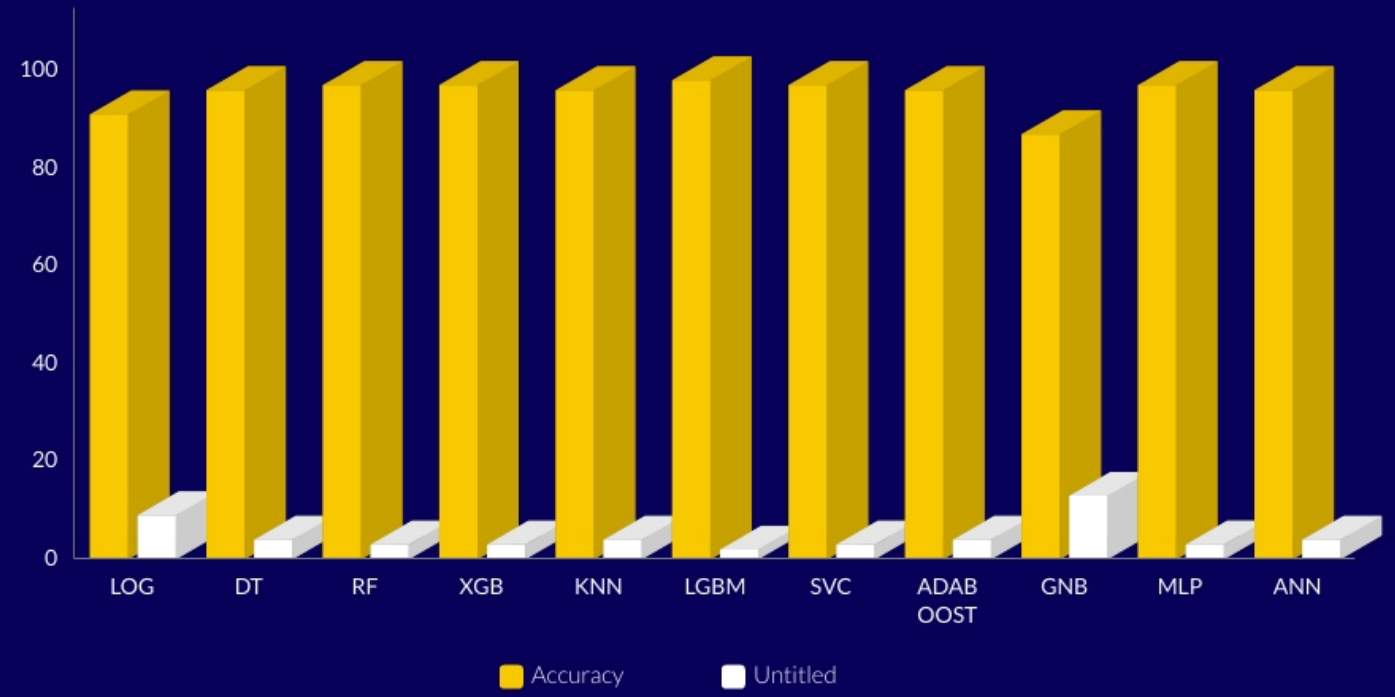
AdaBoostClassifier

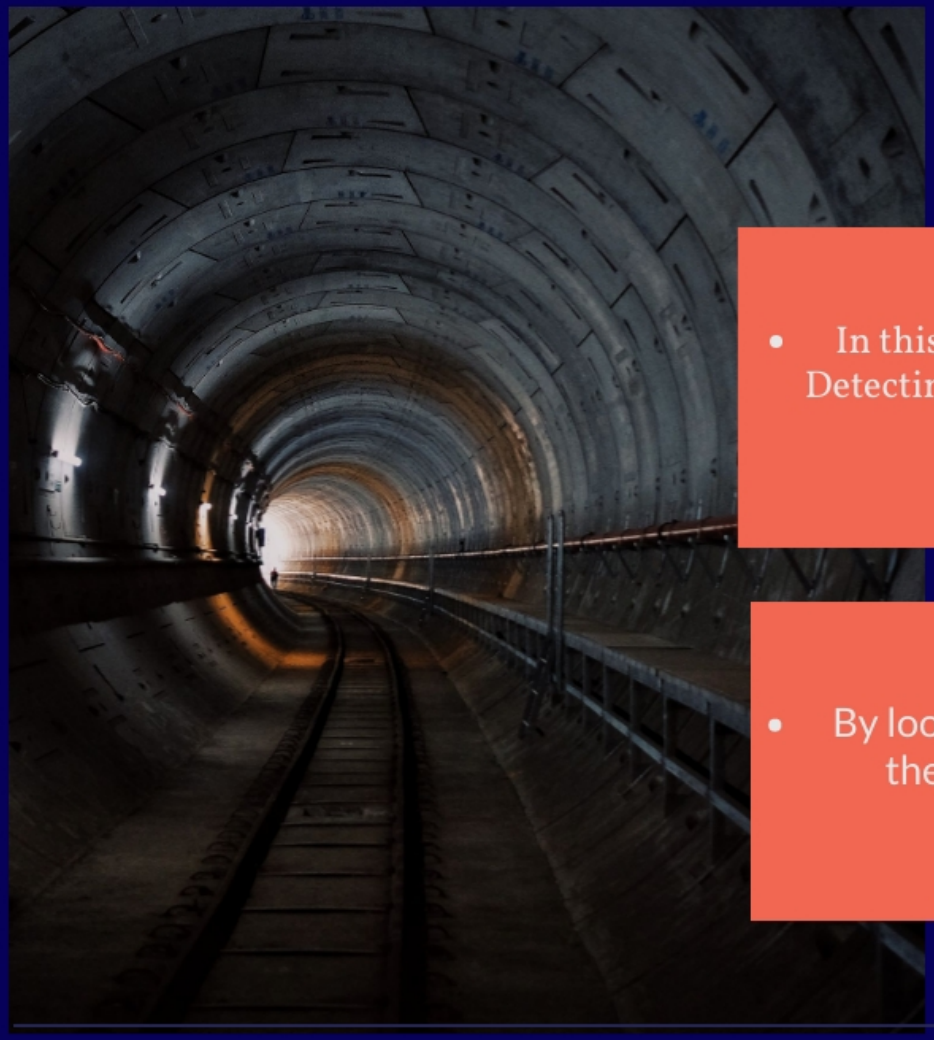
MLPClassifier

Deep Learning Models

Artificial Neural Network

Comparing the Models





Conclustion

- In this Project, we built a Multi models that could approximately Detecting a Fraud transactions of Ethereum given some information about the population.

- By looking this project we have many models . All Models gives us the best result so we say that all models are predicted good.

THE END

Thank you for listening