**Blue Print for Credit Card Fraud Detection**

Data Extraction

Importing the dataset

Importing required packages into our python the environment.

Data Processing

* Processing the data to our needs and Exploratory Data Analysis. (Data Structuring)
* Identify mistakes and errors in data.
* Check for missing values and duplicates.

EDA

* View shape of the dataset.
* Check datatypes and other basic information of all the variables of the dataset.
* Checking correlation b/w variables of dataset.

Data Visualization

* Plot sea-map, bar graph, etc.

Feature Engineering

* Scaling - Normalization and Standardization is required (Standardizing time and amount column)

Data Splitting

* Split into train and test dataset using Random Under Sampling since our dataset is highly unbalanced.

Model Creation

* It is an iterative phase where we continuously train and test machine learning models to decide the best one of them.
* Random Forest, Logistic Regression, SVM, and some other algorithms are used to predict the credit card fraud detection.

Model Evaluation

* To estimate the accuracy of the model by calculating any three error metrices : RMSE, MSE, MAE depending on our model.

**Based on the accuracy of different models, we’ll select the best one i.e., one with largest accuracy score.**