This case study aims at detecting credit card fraud transactions using machine learning models

The process followed is as follows:

* The data is pca of the original data since because of privacy issues it’s not possible to get original data
* Since the variables are generated by pca the variables are normally distributed but skewed
* Skewness of variables can be removed by any of power transformation techniques
* The data is highly imbalanced since number of frauds are very less compared to normal transactions
* Class imbalance problem can be handled using different techniques like Smote, oversampling, under sampling . we wont be using under sampling since data is precisious
* We have to divide the data in to test and train and we can start trying different models like logistic regression, random forest, etc
* We have to do hyper parameter tuning
* Matrices we will be using are AUC roc to calculate efficiency of model since its higly unbalanced data