Project Description

(Fraud Detection)

In this project , we are analyzing Credit Card FRAUD DETECTING from the selected data , as taken from kaggle.com .

In fraud detection, misclassification costs (false positive and false negative error costs) are unequal, uncertain, can differ from example to example, and can change over time. In fraud detection, a false negative error is usually more costly than a false positive error.

Fraud detection systems operates by adding fraudulent claims/applications/ transactions/accounts/sequences (A) dataset to “black lists” to match for likely frauds in the new instances (E) dataset . Some use hard-coded rules which each transaction should meet such as matching addresses and phone numbers, and price and amount limits. In this way different datasets will be created and at the end we will get the final results by combine training data (the class labels are not required here) with evaluation data (A + C + E + F). These should be processed by single or multiple unsupervised algorithms to output suspicion scores.

Data mining approaches will be utilize for training/testing data with labels.