**Fraud Detection**

**Project Summary:**

The project is about to detect fraud transaction with or without card by using ML model.

The steps performed in this project are :

1. Understanding problem statement
2. Data visualisation & observing outliers by boxplot
3. Data cleaning (Replacing outliers with mean, median & highest value)
4. Comparing the performance of the model after replacing outliers with mean, median & highest value
5. Logarithmic transformation of data to improve model performance
6. Checking correlations between variables
7. Feature scaling
8. Using various methods in Feature scaling like standard scaler, minmaxscaler, Robustscale, maxabsscaler and comparing its respective model performance
9. Applied Variance inflation factor for removal of some unwanted features
10. Splitted the dataset into train and test data to train the model.
11. Used Logistic Regression and evaluated model performance by confusion matrix
12. Got **94% accuracy** where true positive=**18034**,true negative=**854**,false positive=**167,** false negative=**941**

**Column Description in dataset :**

* The dataset consists of 8 columns namely distance\_from\_home,distance\_from\_last\_transaction,ratio\_to\_median\_purchase\_price,repeat\_retailer,used\_chips,used\_pin\_number,online\_order,

Fraud

* Target variable is **Fraud** which indicates whether the transaction is fraud or not. It is in Boolean form with 0 & 1 values
* Remaining all variables are independent features

1. distance\_from\_home : the distance from home where the transaction happened.
2. distance\_from\_last\_transaction : the distance from last transaction happened.
3. ratio\_to\_median\_purchase\_price : Ratio of purchased price transaction to median purchase price.
4. repeat\_retailer : describes whether the retailer is repeated or not (in Boolean form with 0 & 1 values)
5. used\_chips : describes whether the transaction done by using\_chips or not.The values are in Boolean form (0 & 1 values).
6. used\_pin\_number : indicates whether transaction done by using pin number or not ( in boolean form with 0 & 1 values).
7. online\_order : describes whether online\_ordered or not (values in boolean form)