**PROBLEM STATEMENT OF PROJECT**:

* The main objective of our project is to detect the fraud happening in an insurance company.

>>**IMPORTING LIBRARIES AND DATASET**

1.What are all the necessary libraries to be imported?

2.How to import data set?

>>**DATA UNDERSTANDING/EDA**

3.Perform necessary functions for data understanding?

4.How to convert an object to integer datatype?

5.Rename the necessary columns if necessary?

6. Drop the null values

7. Drop the duplicate values

>>**DATA VISUALIZATION**

8. visualize the necessary columns?

9.Find the relation between Age and Result?

10.How is payment typology related to result?

11. Relation between Admission type and Emergency department?

12. Visualize the pie chart using necessary columns?

13. Visualize the following box plot

a) Relation between age and payment typology?

b) Relation between Gender and Payment typology?

c) Relation between Age and Payment typology?

14. visualize the following using scatter plot

15.Visualize the following bar plot

a) relation between days spent in hospital to total charge?

b) relation between days spent in hospital to total cost?

16. Plot the correlation matrix using heat map?

>> **DATA PREPROCESSING**

17. Drop the unnecessary columns?

18. Perform label encoding?

>> **FEATURE SCALING**

19.To check whether it is standardization or normalization?

>>**MODEL BUILDING**

20. Split the data into X label and Y label?

>>**MODEL TRAINING**

21.LOGISTIC REGRESSION

22.DECISION TREE

23.RANDOM FOREST

* ADA BOOST
* GRADIENT BOOSTING
* LIGHT GRADIENT BOOSTING

24. KNN

25. SUPPORT VECTOR MACHINE

>>**MODEL TESTING/MODEL EVALUATION**

* Accuracy
* Classification report
* Confusion matrix
* auc, roc

>>**MODEL DEPLOYMENT**

* By using pickle

>> **DEPLOYMENT BY USING STEAMLIT**

>> **CONCLUSION**