Week8:

Team member's details:

Banking Details, Vasu Sharma,

vasu.vs45@gmail.com,

United States of America,

Indiana University Bloomington,

Data Science

Problem description: ABC bank wants to sell its term deposit term. Before launching the product, the bank wants to develop a model so that they can understand which customer will buy this product.

An ML model so that they can use their resources only for the customers who will buy their product which will save them resources and time

Data understanding :

Basically data is about customers who have bought banks products in the past. Various characteristics of these customers such as age, job, marital status, various types of loan such as housing loan, personal loan, education loan etc. have been provided regarding customer's details.

What type of data you have got for analysis:

The data that is provided is of customers with their various characteristics so that based on this data, it can be predicted weather they will be able to buy the new product or not.

What are the problems in the data :

There are a number of unknown values for some of the features in the data as follows:

poutcome: 36959

contact : 13020
education : 1857
job : 288

These needs to be handled accordingly.

What approaches you are trying to apply on your data set to overcome problems like NA value, outlier etc and why?

For jobs column removed those rows with unknown values. We have used techniques to remove imbalance in the dataset such as oversampling to remove these imbalances in poutcome variable which is dependent variable in our dataset.