**IMPLEMENTATION AND ANALYSIS OF PROJECT**

Business Idea: - Predict Customers which can take term plan .

Supervised Binary Classification Project

Domain – Banking Domain

**Implementation**

Data Shape – (45211,17)

Data Cleaning

Column wise data cleaning

age – Remove outliers so as to remove skew

job – keep as it is

marital – keep as it is

education – keep as it is

default – drop as most of the values are same

balance - Remove outliers so as to remove skew

housing – keep as it is

loan – drop as most of the values are same

contact – keep as it is

day – keep as it is

month – keep as it is

duration - Remove outliers so as to remove skew

campaign - Remove outliers so as to remove skew

pdays – drop

previous – drop

poutcome – drop

New Data Shape – (45211,12)

**Data Encoding**

Nominal Variables – One Hot Encoding and drop the originals

Ordinal Variables – Label Encoding

Numerical Variables – Do Nothing

**MODELLING**

Logistic Regression

Decision Tree

Random Forest Classifier

Support Vector Machine Classifier

Naïve Bayes Classifier

**MODEL ANALYSIS**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODEL** | **ACCURACY** | **PRECISION** | **RECALL** | **F1-SCORE** |
| **DECISION TREE**  **CLASSIFIER** | 0.88 | FOR NO-0.92 FOR YES-0.41 | FOR NO-0.92 FOR YES-0.42 | FOR NO-0.92 FOR YES-0.41 |
| **NAÏVE BAYES CLASSIFIER** | 0.83 | FOR NO-0.91 FOR YES-0.32 | FOR NO-0.92 FOR YES-0.41 | FOR NO-0.90 FOR YES-0.34 |
| **RANDOMFOREST CLASSIFIER** | 0.89 | FOR NO-0.91 FOR YES-0.59 | FOR NO-0.97 FOR YES-0.27 | FOR NO-0.94 FOR YES-0.37 |

**INFERENCE**

According to me best model is decision tree or naïve bayes classifier

Improvement for prediction in case of yes scenario will be better .