Insights

Insights about Exploratory data analysis:

1)The number of cases filed in Indian courts has been increasing over the years, leading to a backlog of cases. However, the Indian government has been taking steps to address this issue, such as implementing the use of technology in the court system and increasing the number of judges.

2) In Ladakh, Diu, and daman, DNH at Silvassa has a high number of pending cases and also has a high duration to complete cases.

3) From 2010 to 2018 many types of cases are coming to the court, But in 2010 we have a low number of cases those cases are also criminal/noncriminal they are only political issues family cases are very low at that time when compared to the next years.

4) Average duration of the case from filing to disposal of small states taking more than 3 years of time.

5) According to the population in every state and total cases in every state having similar race, which implies population is proportional to the total cases.

6) Maharashtra has a high number of cases filed every year but pending cases are very low among these total cases.

Classification Problem:

Predicting the state judicial system needs any other help from the government like needing judges or courts in that state. Obviously, if the state has fewer judges then it is the government's responsibility to send judges to that state

1. The state and district of the court cases are important factors in predicting the needs of the court.

2. The filing date and closing date of the case can be used to determine which type of cases take more days to complete.

3. Pending cases which were idle can be used to determine the lack of judges.

Insights about classifier:

I build a classifier that tells about the need for judges in that state.

1) By using previous data it can predict which district people need some awareness or counseling in a particular state.

2) It predicts the state has a higher number of dismissed cases, withdrawal cases, and criminal cases based on the total cases count.

3) It predicts if the target variable has yes then that state needs judges to clear cases.

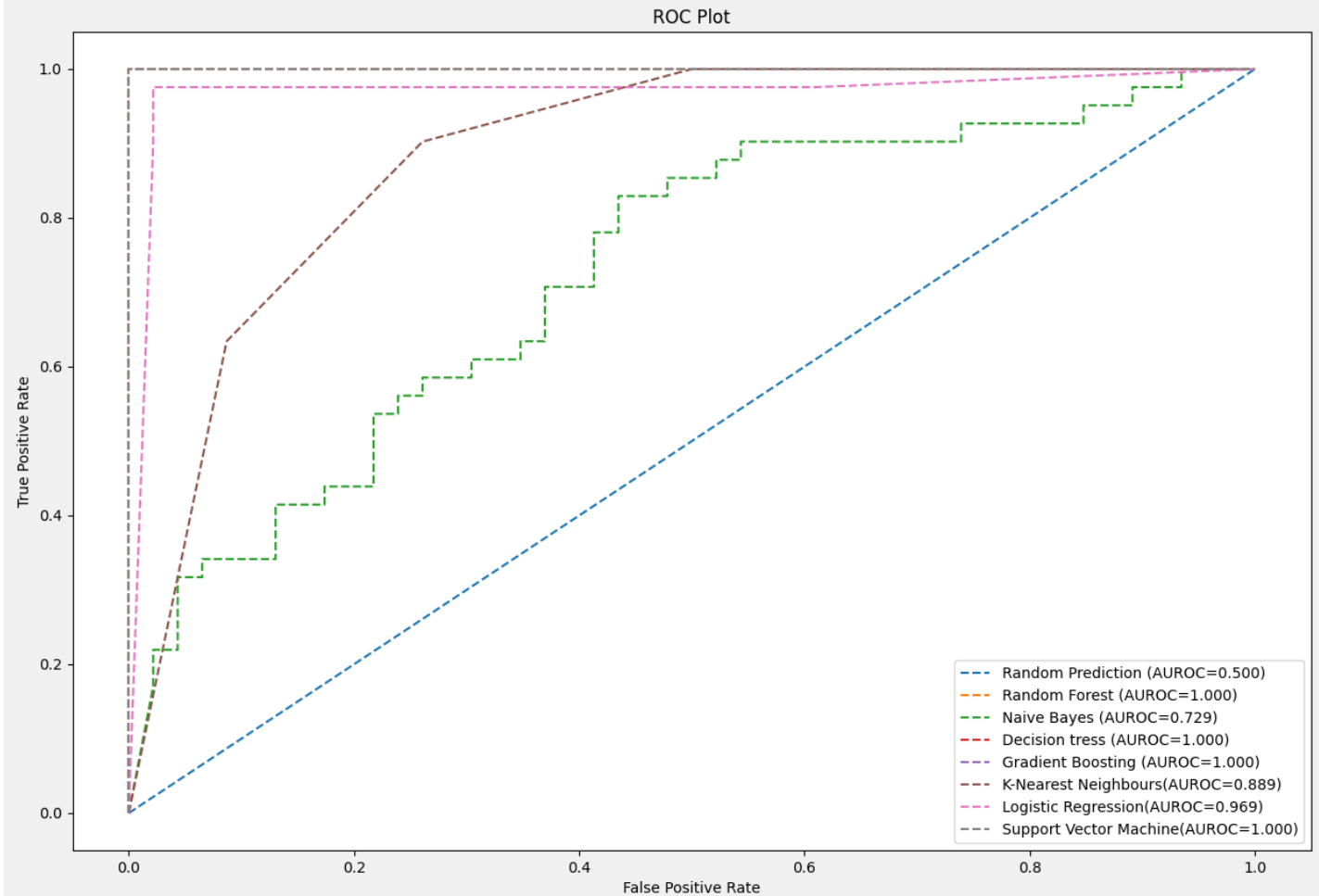## Accuracy of my Classifier:

Classifier score with different classification Algorithms
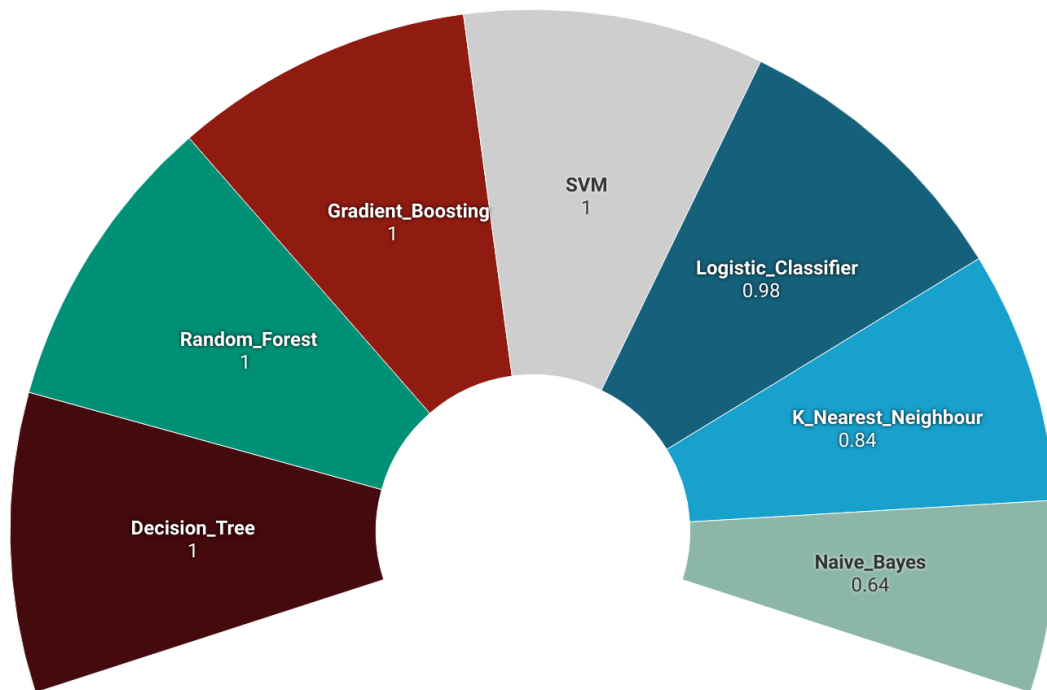
I used different types of classification algorithms to check my classifier for whether the target variable worked perfectly or not. I checked the AUROC score using the below graph

AUROC scores of the classifier:

ROC Plot

Legend:
- Random Prediction (AUROC=0.500)
- Random Forest (AUROC=1.000)
- Naive Bayes (AUROC=0.729)
- Decision tress (AUROC=1.000)
- Gradient Boosting (AUROC=1.000)
- K-Nearest Neighbours(AUROC=0.889)
- Logistic Regression(AUROC=0.969)
- Support Vector Machine(AUROC=1.000)

Observations from the analysis:

1) If percentage of dismissed cases is high, it could indicate that prosecution is not presenting a strong case, this is a lack of evidence

2) If percentage of withdrawn cases is high, it could indicate that the process indicates prosecution or defense is choosing to withdraw cases before they reach trial.

3) If percentage of pending cases is high, it indicates that court system is overwhelmed or understaffed, resulting in delays in processing cases.

4) If the time taken to close case after filing is long, it could indicate that the court system taking too much time in processing cases.

5) If the time taken to reach court after filing is long, it could indicate that there are delays in the initial stage of case such as investigation or indictment

6) If the time taken to close a case after taking a decision is long it could indicate that the court system is slow in executing the decision.