

SUMMARY

Data Cleaning:

- There were about 13 values which had null values.
- Dropped 10 values for making the data frame easier.
- 22 columns cleared off the null values.
- Also dropped the column with the unique values.

EDA

- With EDA it was possible to remove unwanted categorical values
- No outliers were found.

Train-Test split:

- A condition of 70% and 30% for train and test was split.

Model Building:

- RFE was done to attain 15 variables
- VIF values were <5 and p-value were <0.05
- ROC was at 0.89

Prediction:

- Cut off was at 0.36
- With 0.36 cut-off accuracy attained at 80%

Precision and Recall:

- Rechecked with the cut off of 0.42 with precision of 74% and Recall of 77%.