

SUMMARY

1. Data Cleaning:

- a. First step to clean the dataset we chose to remove the redundant variables/features.
- b. After removing the redundant columns, we found that some columns are labeled as 'Select' which means the customer has chosen not to answer this question. The ideal value to replace this label would be null value as the customer has not opted for any option. Hence, we changed those labels from 'Select' to null values.
- c. Removed columns having more than 40% null values

2. Data Transformation:

- a. Changed the multicategory labels into dummy variables and binary variables into '0' and '1'.
- b. Checked the outliers and created bins for them.
- c. Removed all the redundant and repeated columns.

3. Data Preparation:

- a. Split the dataset into train and test dataset and scale the dataset.
- b. After this, we plot a heatmap to check the correlations among the variables.

4. Model Building:

- a. We created our model with rfe count 18, our final model with rfe 18 variables as it has more stability and accuracy.
- b. For our final model we checked the optimal probability cutoff by finding points and checking the accuracy, sensitivity and specificity.
- c. We found one convergent point and we chose that point for cutoff and predicted our final outcomes.
- d. We checked the precision and recall with accuracy, sensitivity and specificity for our final model and the tradeoffs.
- e. Prediction was made now in the test set and predicted value was recorded.
- f. We did model evaluation on the test set like checking the accuracy, recall/sensitivity to find how the model is g, We found the score of accuracy and sensitivity from our final test model is in acceptable range. h. We have given the lead score to the test dataset for indication that high lead scores are hot leads and low lead scores are not hot leads.

5. Conclusion: Learning gathered are below:

- a) Test set has accuracy, recall/sensitivity in an acceptable range.
- b) In business terms, our model is having stability and accuracy with adaptive environment skills. Means it will adjust with the company's requirement changes made in the coming future.