Predicting Employee Retention

Problem statement – Identify the factors to improve the employee retention in the organization and increase the turn around.

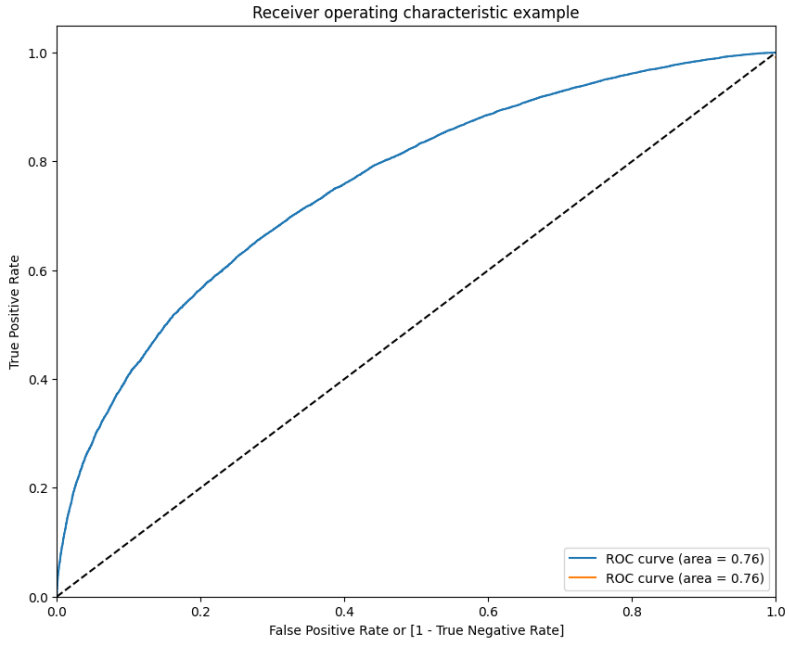
Objectives – Help HR team to identify the factors that they work on to improve the employee retention.

Approach –

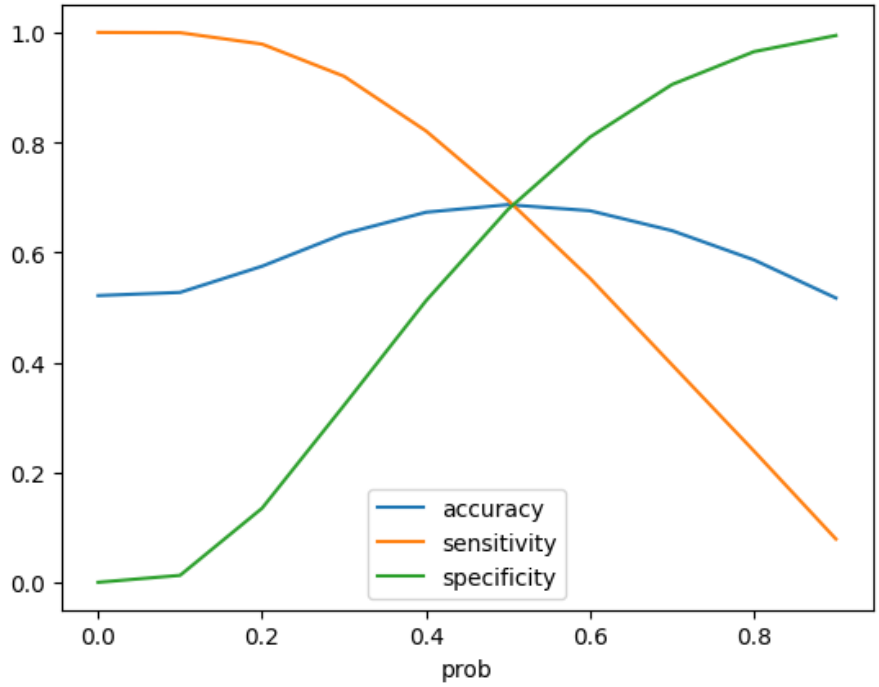
- Perform EDA to identify the gaps in the data set
- Rectify the gaps by removing the null data, removing redundant columns, removing duplicate information.
- Find out the outlier and its impact to the analysis
- Convert categorial data to the numeric values
- Scale all data for better analysis
- Remove high corelated features as they might give incorrect result
- Use Recursive Feature Elimination techniques for finding the suitable features for model preparation
- Check the different measuring technique such as Confusion matrix, Accuracy, Specificity, Recall, Precision for model performance
- Use ROC curve to identify the nearest possible correct probability value for prediction

Charts –

- **ROC Curve**



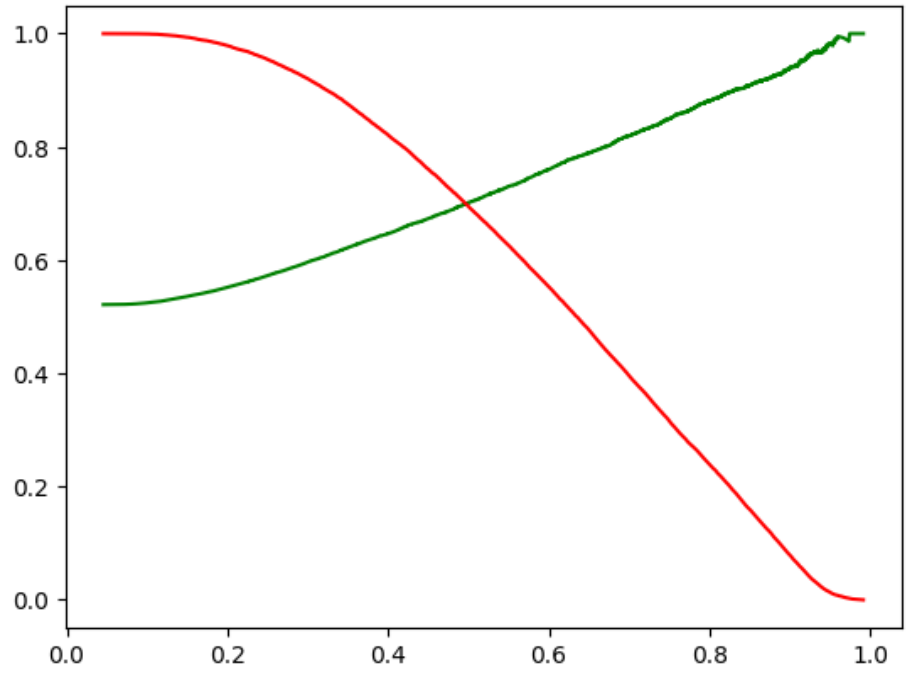
- **accuracy, sensitivity, and specificity at different values of probability cutoffs**



Model Matrix –

- Accuracy - 0.6867567348919991
- True Positive - 17911
- True Negative - 16045
- False Positive - 7618
- False Negative - 7870
- Sensitivity - 0.694736433807843
- Specificity - 0.6780627984617336
- Precision - 0.7015942653452936
- Recall - 0.694736433807843

Precision-Recall trade off –



Based on the above information, we decided to take 0.5 as cutoff.

Evaluation Matrix –

- Accuracy - 0.6842999386532018
- True Positive - 7652
- True Negative - 6849
- False Positive - 3313
- False Negative - 3377
- Sensitivity - 0.6938072354701242
- Specificity - 0.6739814997047825
- Precision - 0.6978568171454629
- Recall - 0.6938072354701242

Conclusion –

From this exercise we conclude below parameters are important for employee retention -
- Years at Company
- Number of Promotions
- Overtime
- Distance from Home
- Number of Dependents
- Work-Life Balance_Fair
- Work-Life Balance_Poor
- Job Satisfaction_Low
- Job Satisfaction_Very High

- Performance Rating_Below Average
- Performance Rating_Low
- Education Level_PhD
- Job Level_Mid
- Job Level_Senior

Few factors need to consider as -
- Give more incentives or benefits to the employees worked more years, or they are doing overtime
- Also need to investigate on why employees are doing overtime and work life balance is poor and job satisfaction is low
- Add some travel facilities for employees commute from long distance
- Also need to think on improving employee's performance rating. Check training needs and other support to do the best in work
- Employees like to grow on laddering so refresh hierarchical pyramid periodically
- Involve PhD employees in to the research and development programming to take advantage of their deep specialization