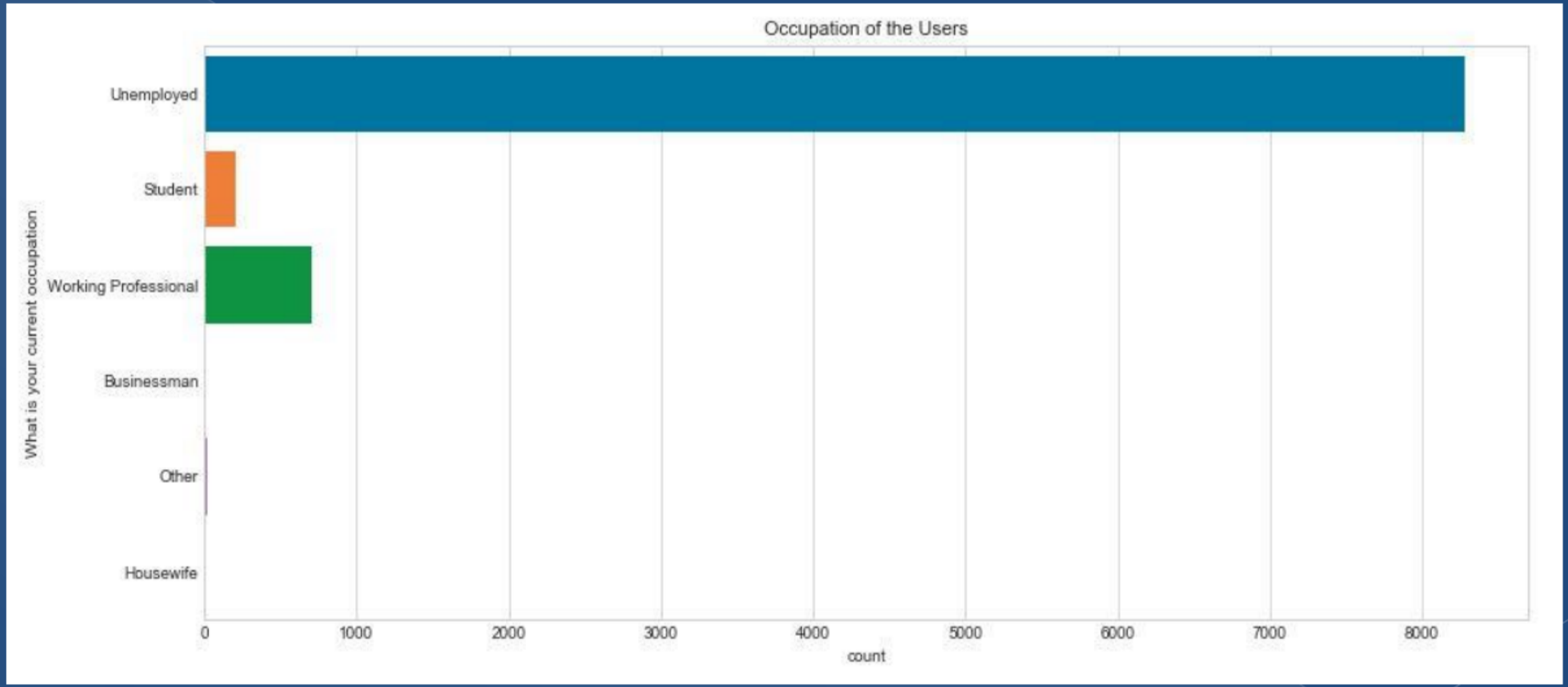
LEAD SCORING CASE STUDY

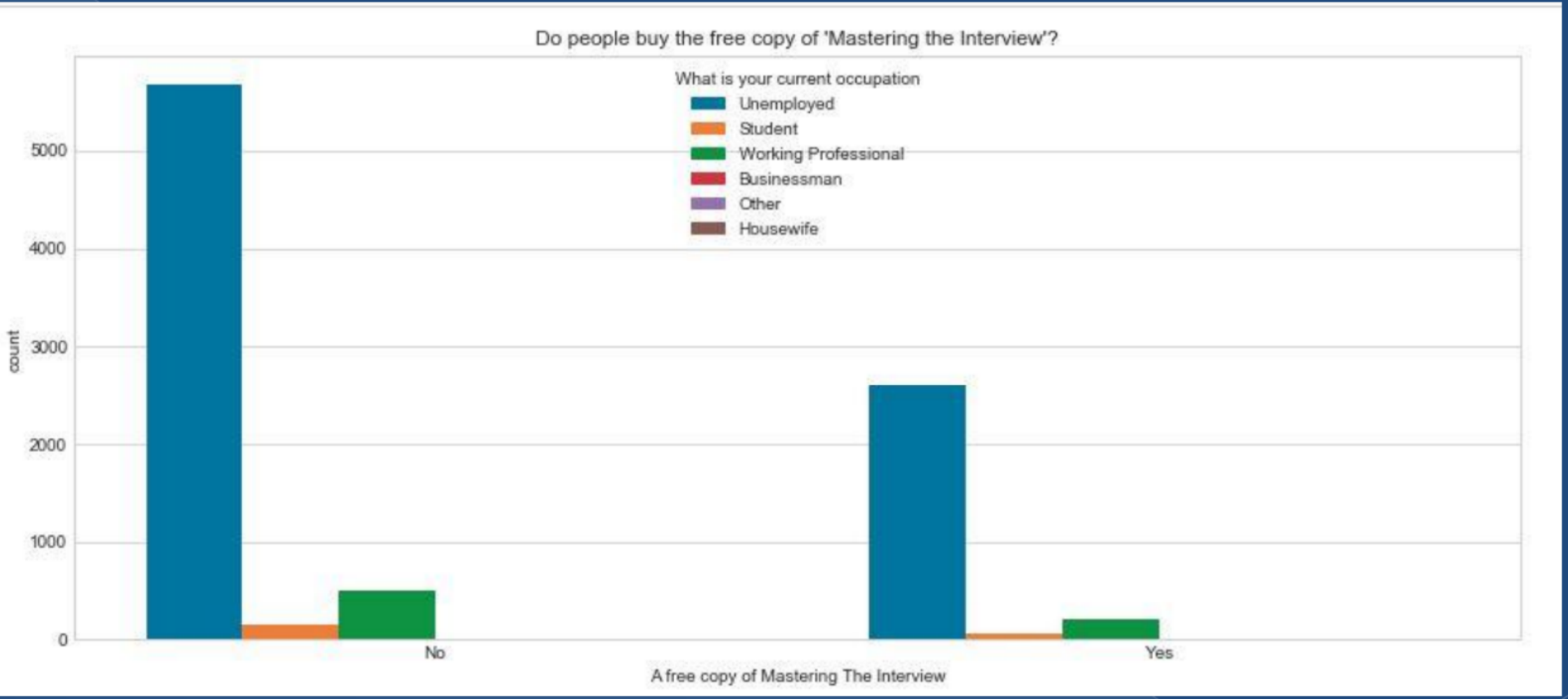
ANALYSIS BY- ANKIT GOREGAONKAR

DOMAIN – SOCIAL MEDIA MARKETING

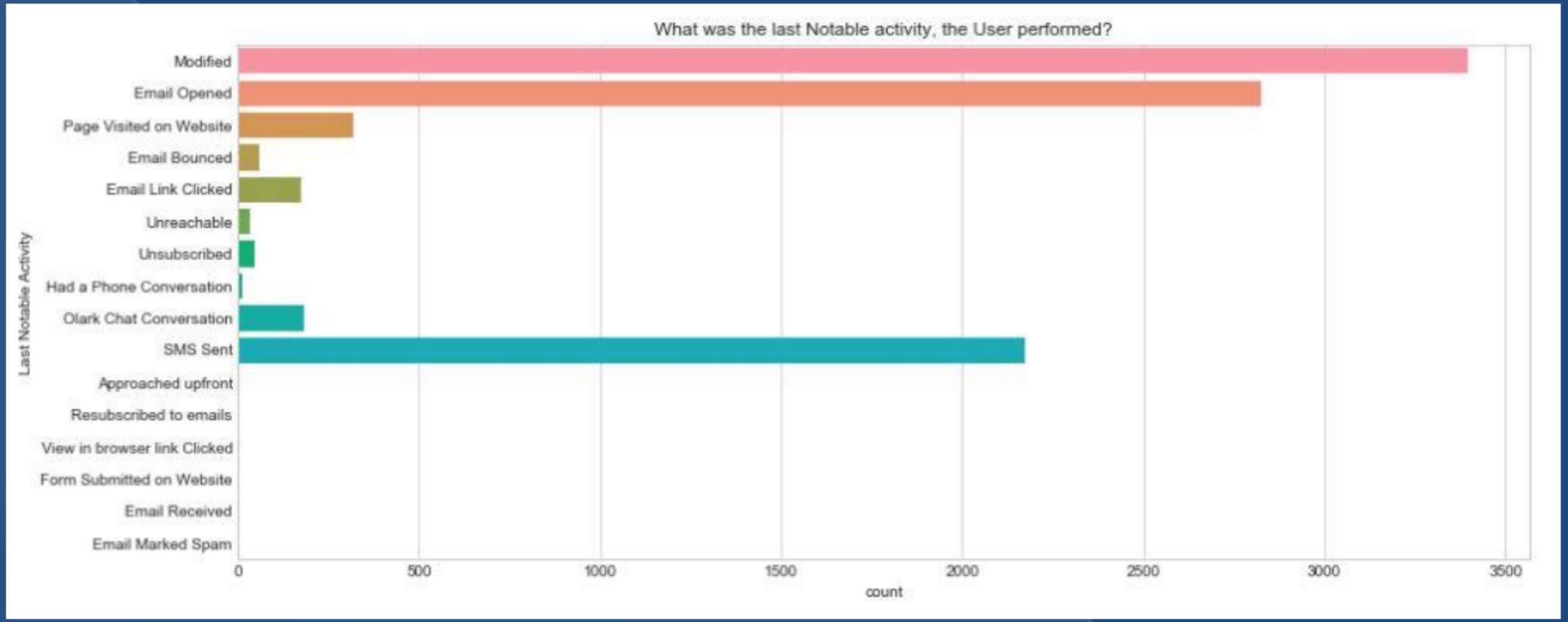
Glimpse from Analysis



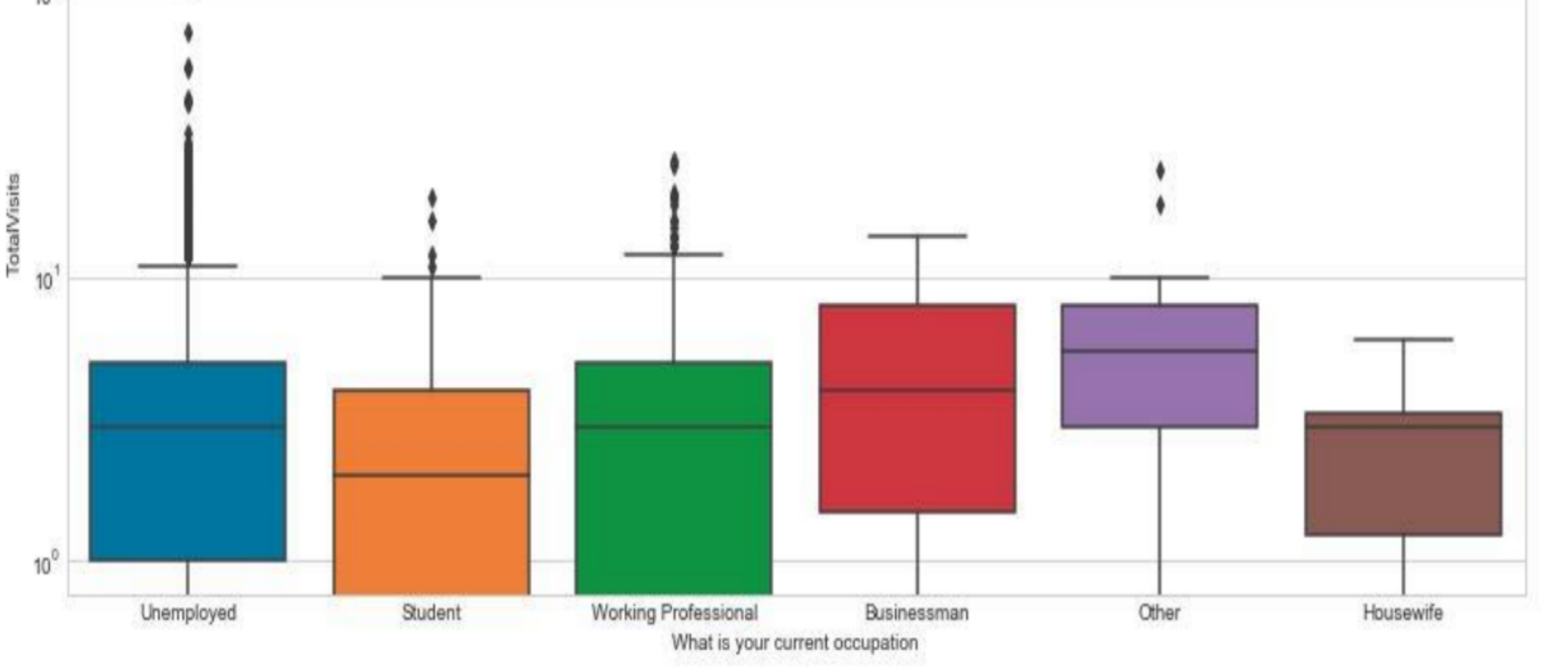
We could infer that a majority of Users are Unemployed whereas Housewife Users are the lowest. Unemployed users are maximum in number followed by Working Professional and Students.



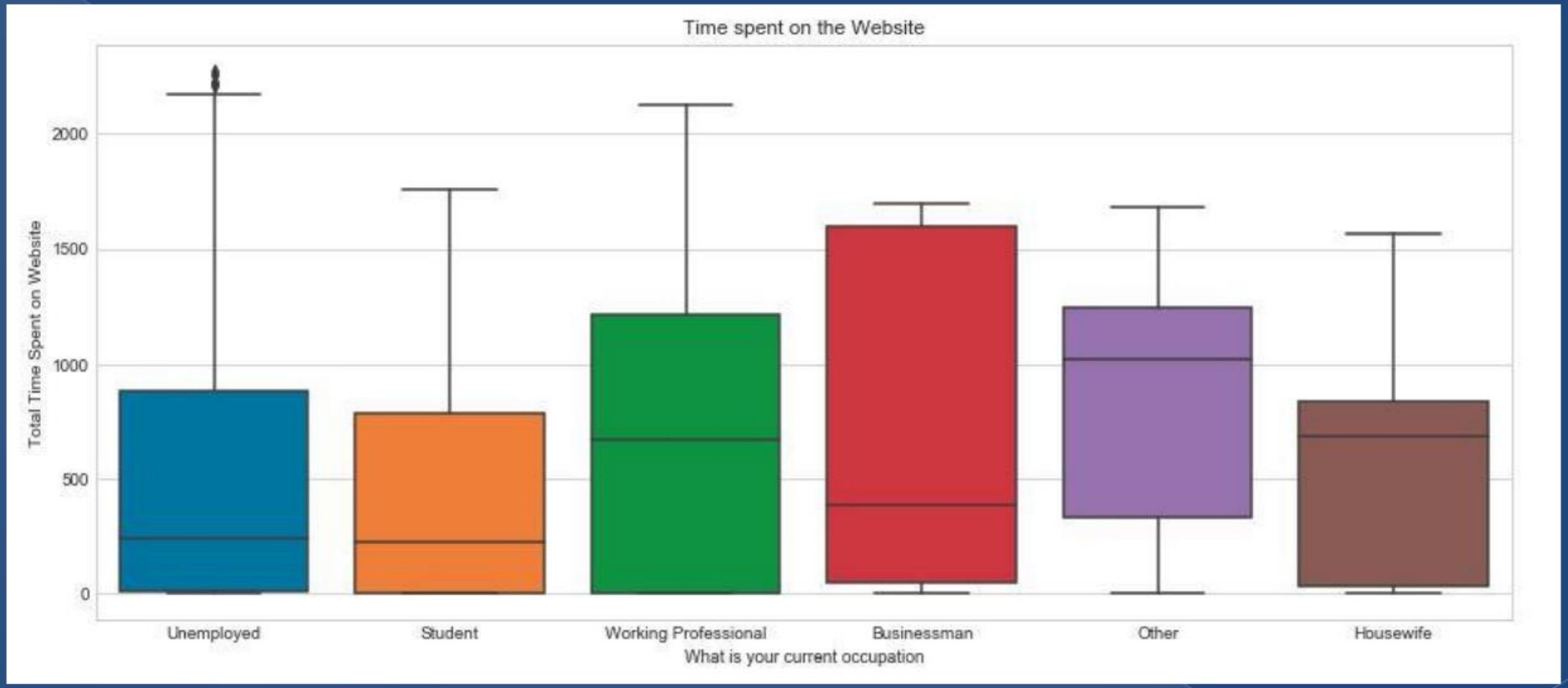
Of all the Users, we could infer that significant amount of Users usually don't take the free copy of "Mastering The Interview".



For the given dataset, we could observe a significant amount of Users who have opened the Email and have been made aware about the program through Text message.

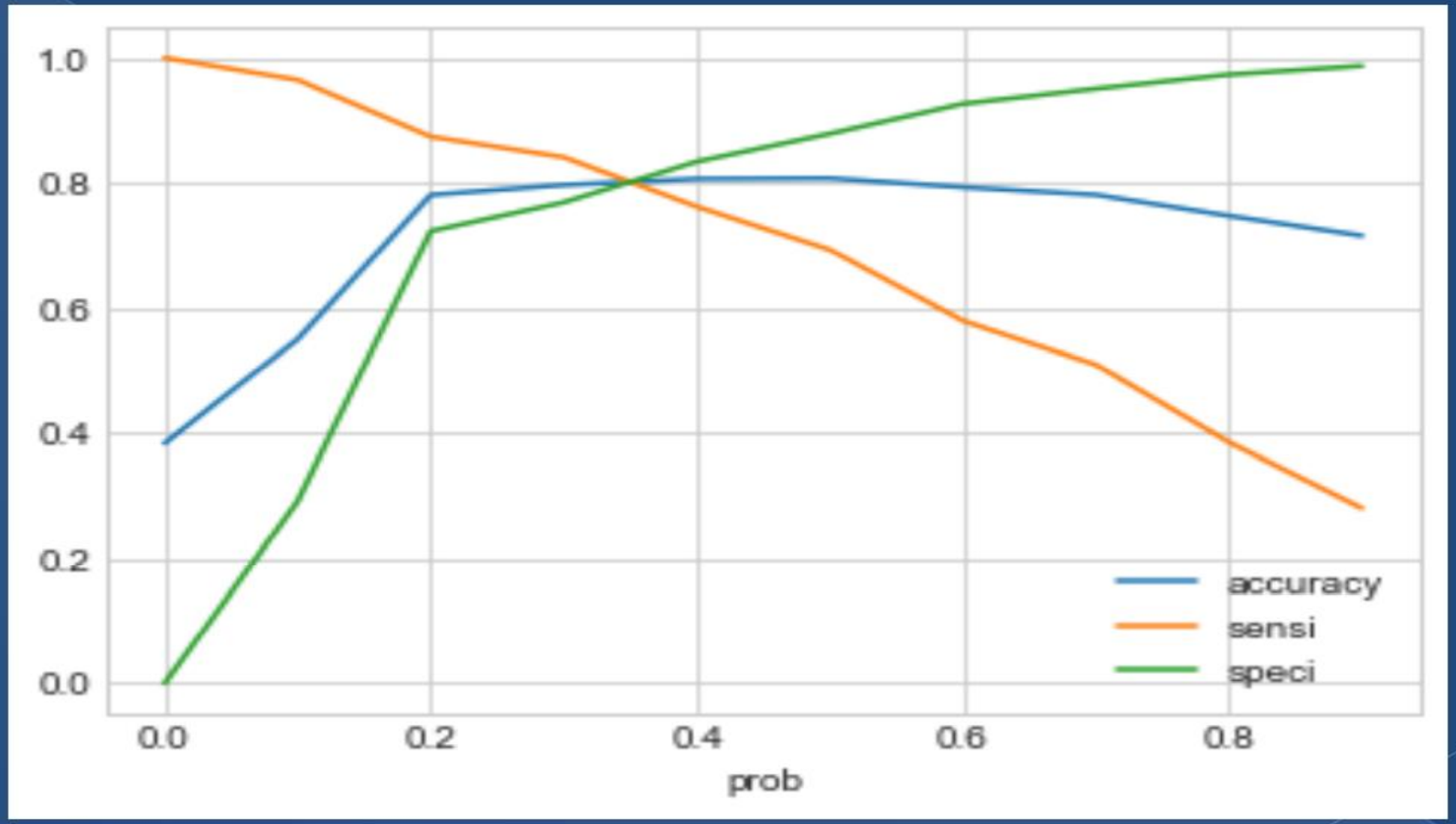


We can see that number of total visits for Businessman and other are greater than all others. However, the median seems lowest for student category. Also, we can see some extended outliers for unemployed category for total visits made to website.



We can see that the median of number of time spent on the website is higher for working professional and other category.
However, the median seems lowest for student category. Also, we can see some outliers for unemployed category for total time spent on the website.

Choosing an optimal cut off

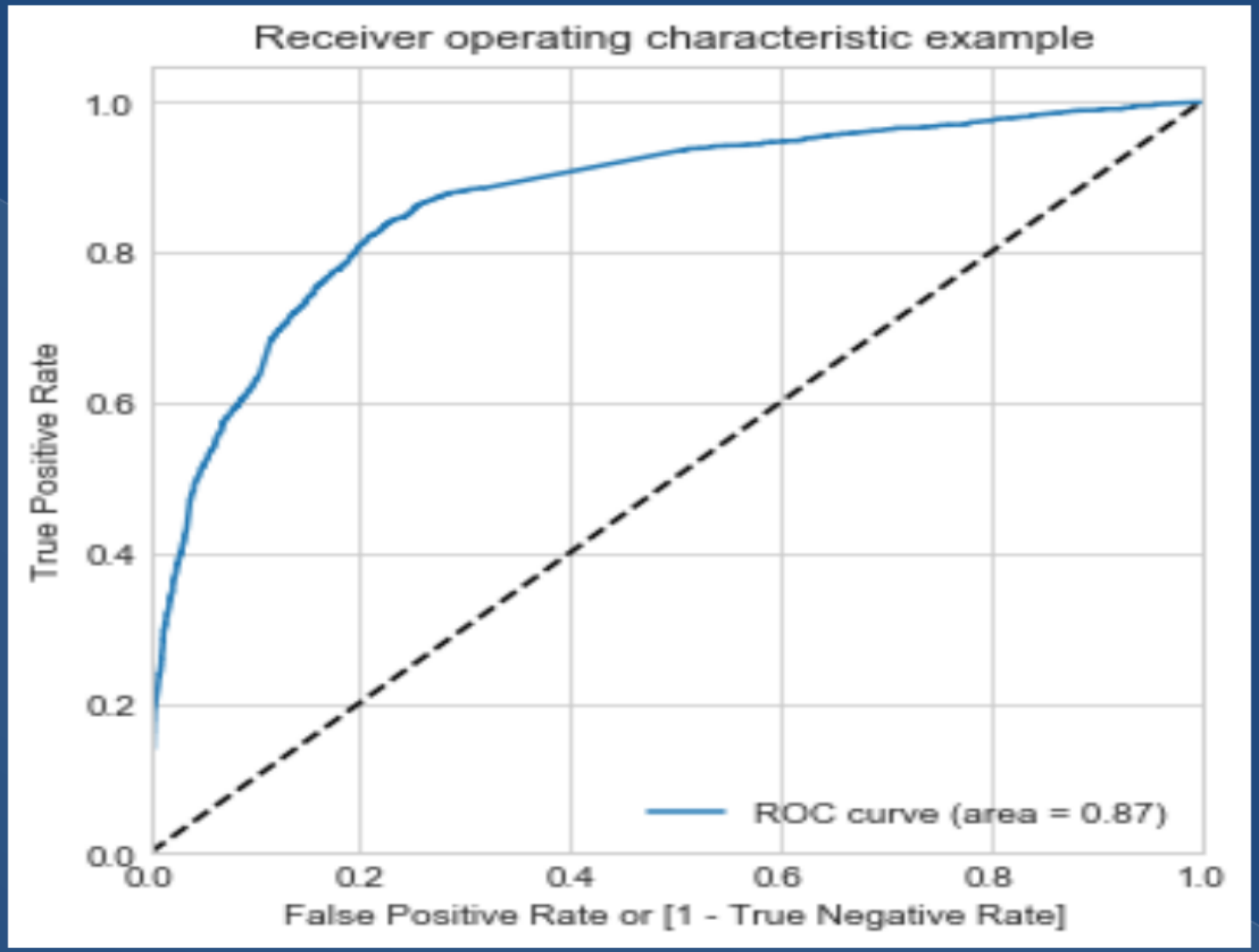


Accuracy, Sensitivity, Specificity plot for choosing optimal cut off for the model.
0.34 has been chosen as optimal cut off as per the plot

Model Evaluation

	Train Data	Test Data
Accuracy	80.24%	78.48%
Sensitivity/ Recall	81.29%	79.07%
Specificity	79.58%	78.11%
Precision	71.26%	69.77

Since, the accuracy and Sensitivity was above 80%, we settled for the model and it equally performed well.
None of the metrics dropped more than 5%



ROC curve for the tradeoff between Specificity and Sensitivity. For our model, area under ROC curve cam out to be 0.87.

THANK YOU