LEAD SCORE CASE STUDY

Submitted by:

Bhumika Jawa

Himanshu Ranjan

Problem Statement

A company named X Education sells it's online course to working professionals and its lead conversion rate is 30%.

The company aims to increase it's lead conversion rate to around 80%

Solution

Our goal is to build a Logistic regression model and give a score from 0 to 100 to each lead and the hotter leads (with higher score) will be targeted by the company's sales team in order to achieve a higher lead conversion rate.

Top features that effect conversion of leads

- **Tags_Lost to EINS**
- **Tags_Closed by Horizzon**
- **Tags_Will revert after reading the email**
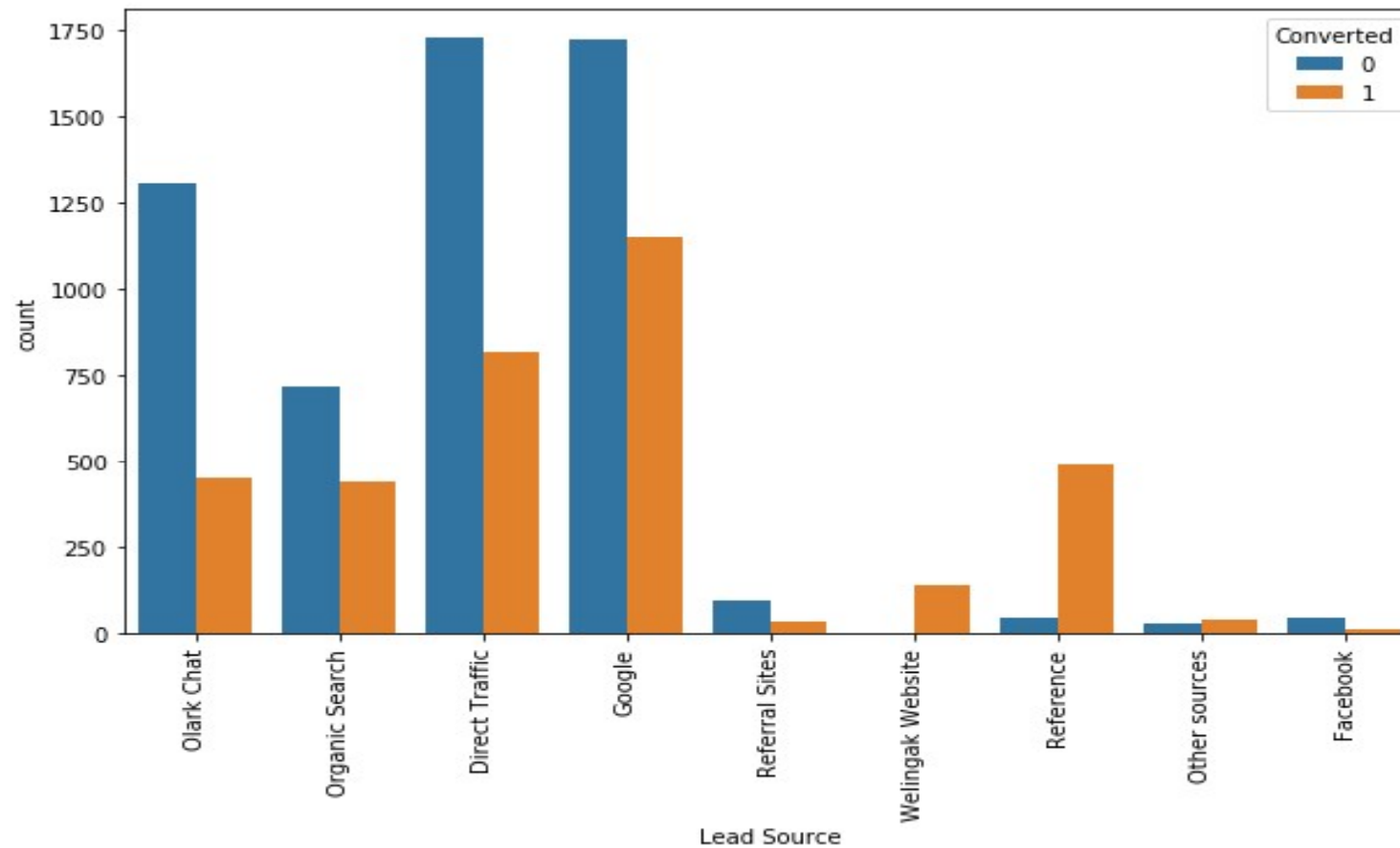- **Tags_Unknown**
- **Total Time Spent on Website**
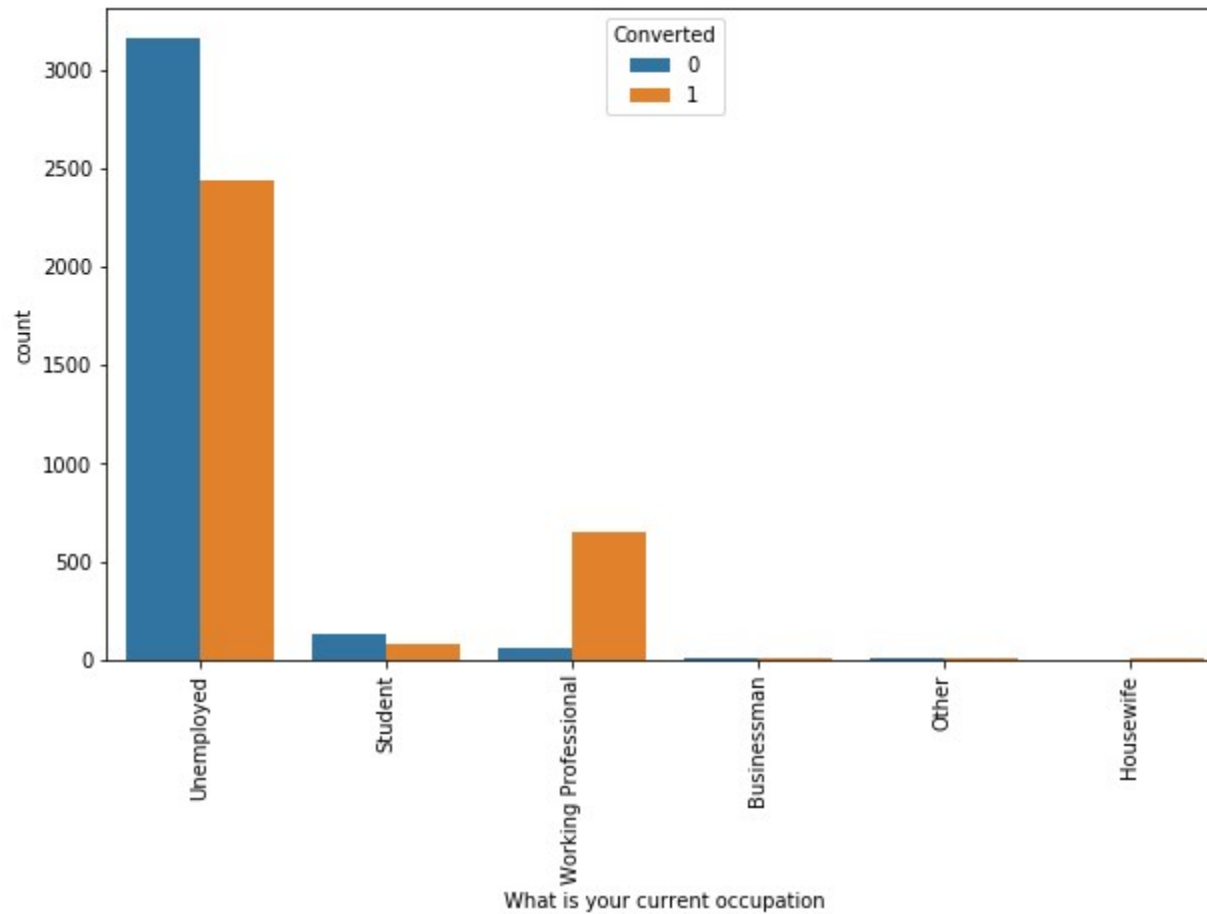
Graph for top features

- **Tags**
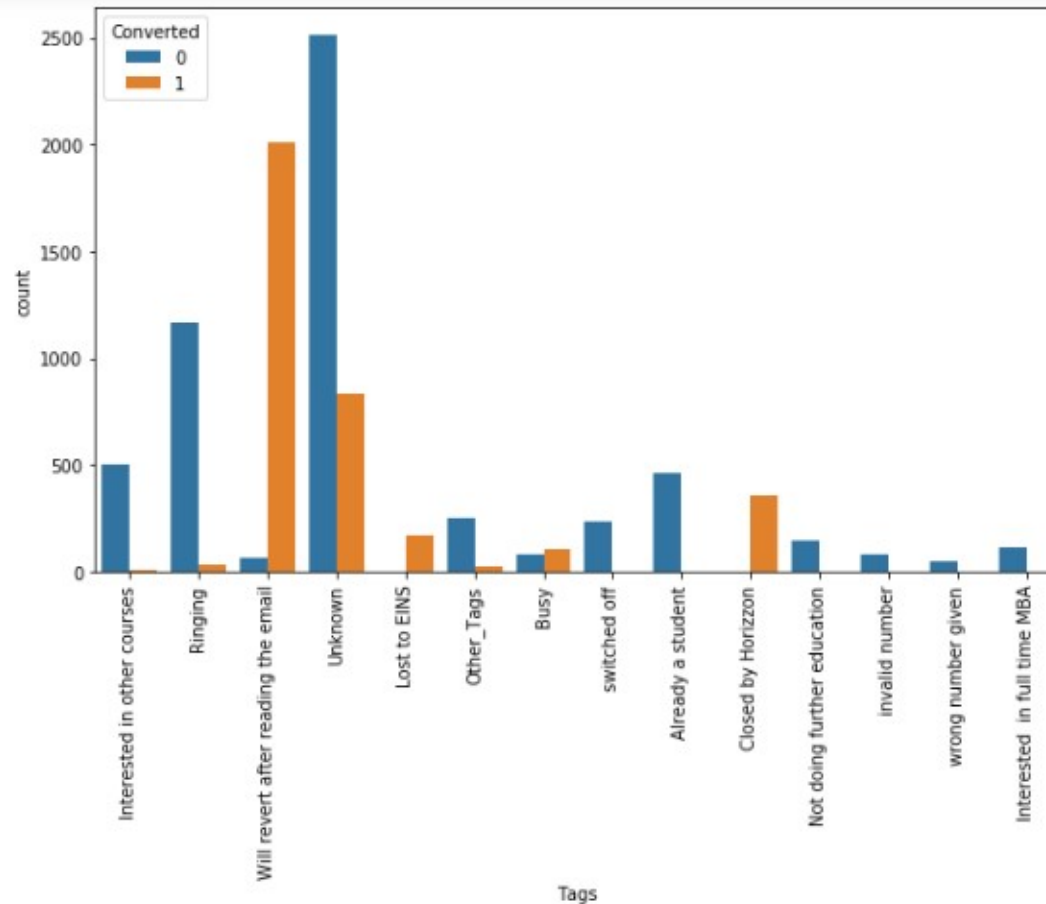
"Will revert after reading the email" have the higher number of conversion and those who won't pick up calls have higher number of people who don't convert.
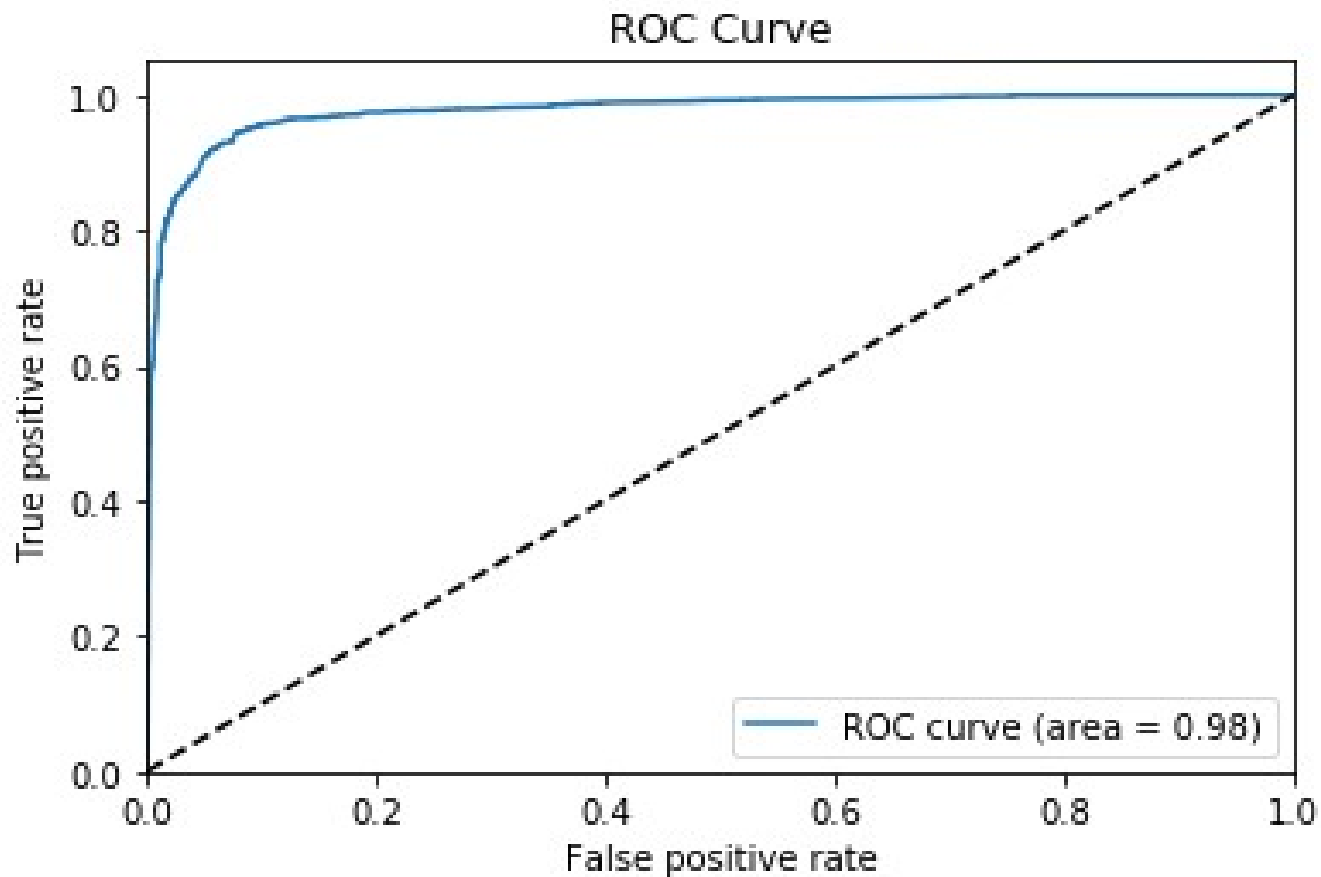
- **Lead Source**

"Google & Direct Traffic" have the highest number of conversion rate.

- **Current Occupation**

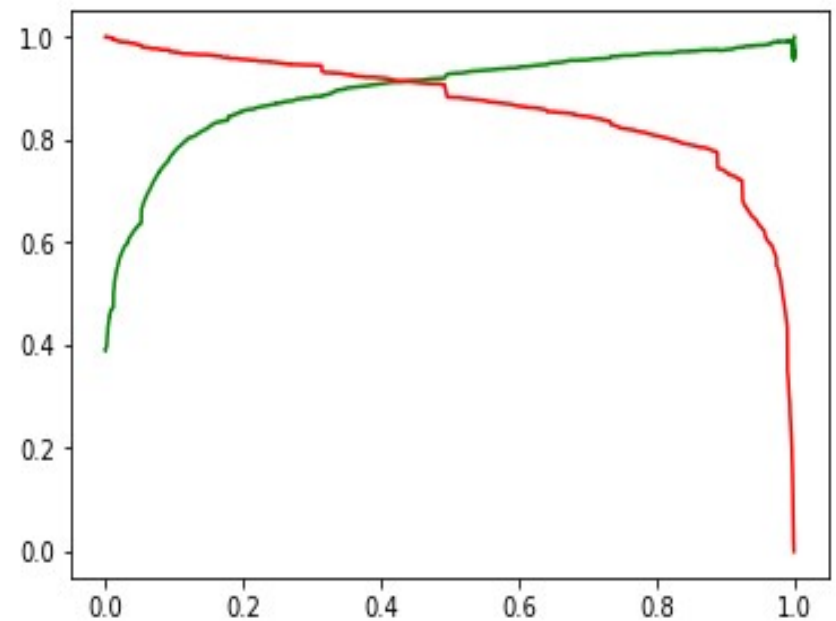"Unemployed" have the highest number of conversion rate.
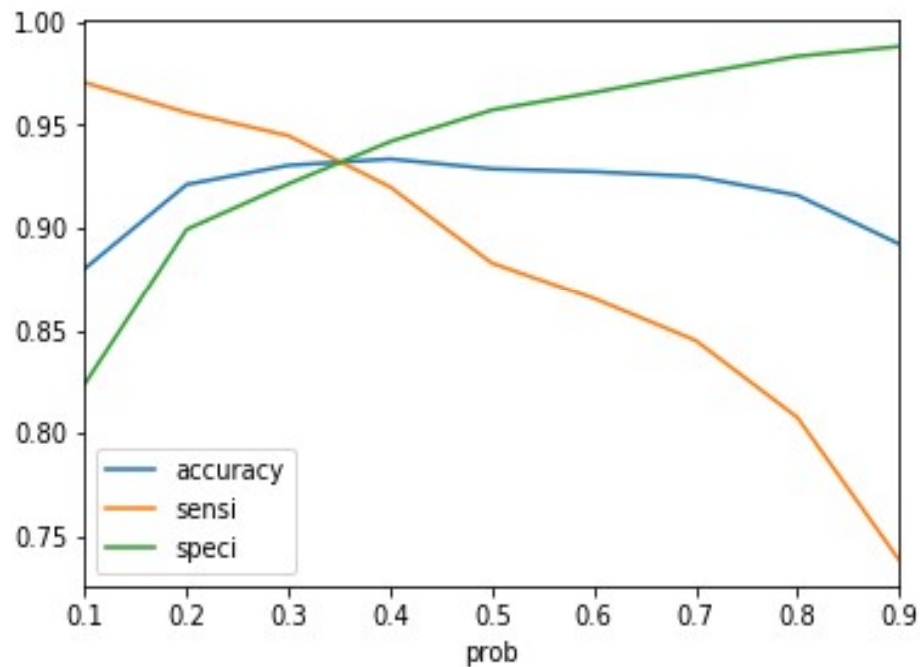
ROC - Curve

auc_score (Area under the curve) = 0.976469

Threshold

1) Using accuracy , sensitivity & specificity
2) Recall – precision tradeoff

Will take threshold value = 0.35

Probability & lead scores

	Prospect ID	Converted	Converted_prob	Lead_Score	final_predicted
4608	4608	1	0.996002	100	1
7935	7935	0	0.050540	5	0
4043	4043	0	0.179519	18	0
7821	7821	0	0.006990	1	0
856	856	0	0.071096	7	0

Final Conclusion

- We have got lead scores (converted_score) of each person between 0-100 now the sales team will contact top hottest leads (LEAD SCORE>80) to attain increase in their conversion rate up to 80%.

Thank You