



Lead Score Case Study:

The company requires a model wherein a lead score to each of the leads is assigned such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance.

Presented by-

Shivam Gupta and Simran Jain



Problem Statement :-

- The company markets its courses on several websites and search engines like Google.
- Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos.
- When these people fill up a form providing their email address or phone number, they are classified to be a lead. Moreover, the company also gets leads through past referrals.
- Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not.
- The typical lead conversion rate at X education is around 30%.
- Now, although X Education gets a lot of leads, its lead conversion rate is very poor.



Problem Statement :-

- To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'.
- X Education has appointed us to help them select the most promising leads, i.e. the leads that are most likely to convert into paying customers.
- The company requires you to build a model wherein we need to assign a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance.
- The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.

Process done in lead score

Advertising through:

- 1.google , Facebook
- 2.Through referrals

People visit the website and provide their mail and contact no. They are identified as leads.

Company starts contacting them via mails or contact number as a result out of 100 only 30 of them are converted to customers.



Desired Solution :-

- The company aims in identifying the hot leads so that they can focus more on them.
- Hot leads are those which are more capable of converting into customers.
- Since we have a target of 80% conversion rate, we would want to obtain a high accuracy in obtaining hot leads so that company can get the maximum benefit of it.



How to Implement ?

Steps involved :-

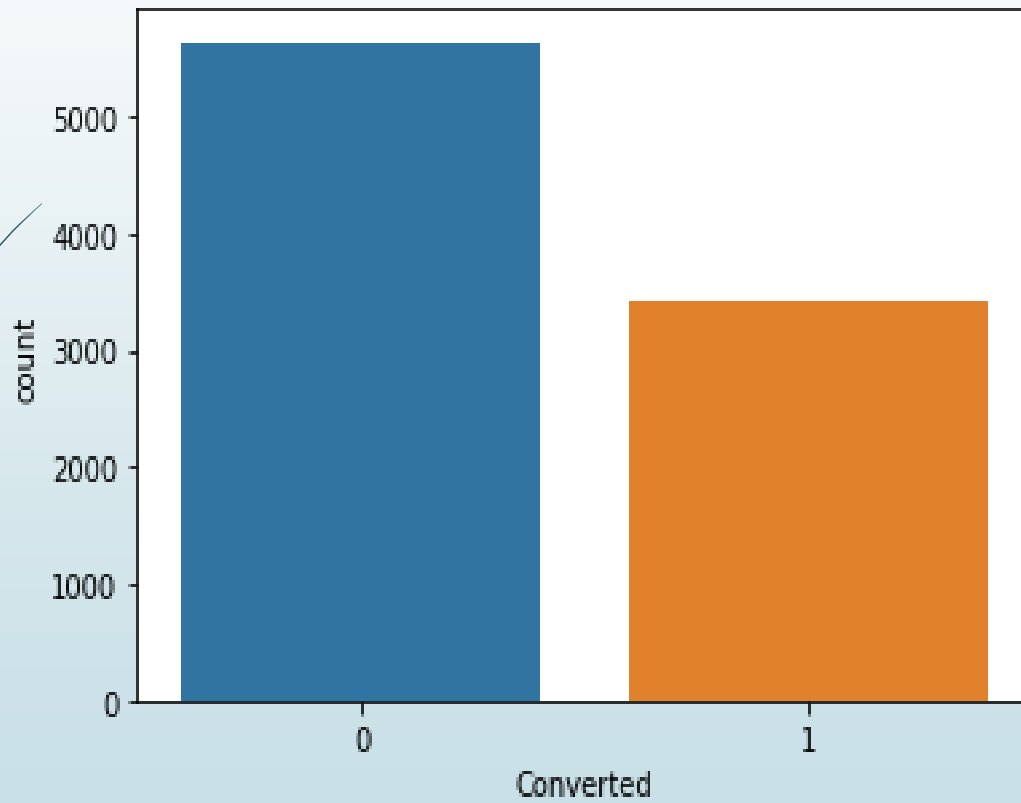
**Data Collection
: Loading the
data**

**Data Cleaning : Cleaning of
data involve removing the
redundancies and missing
values**

**Performing EDA i.e.
univariate and bivariate
analysis and finally building
the model using RFE and
Logistic Regression**

Visualization 1:

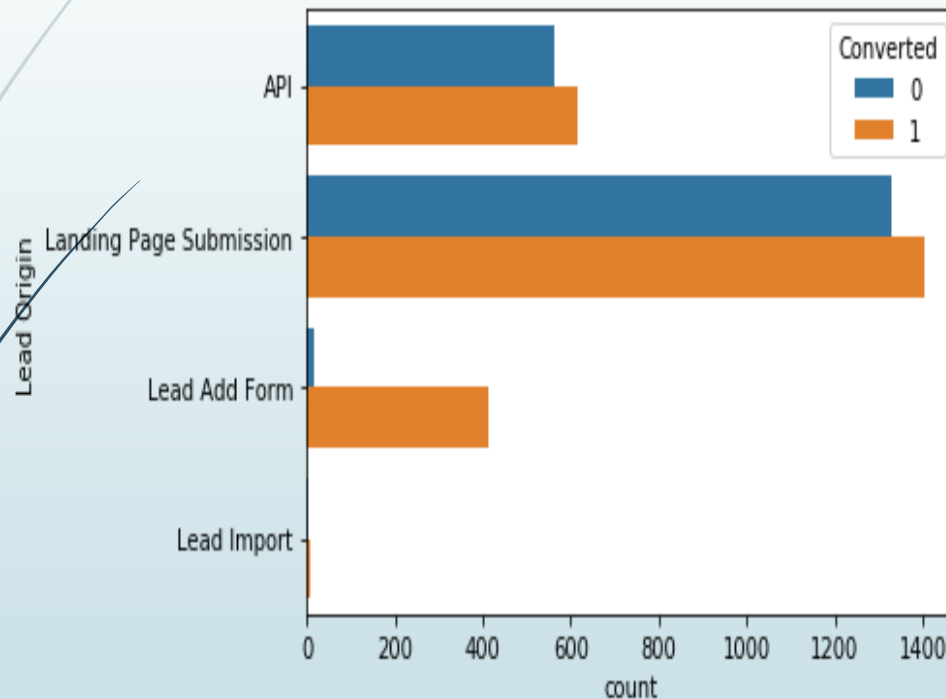
Leads that are converted v/s that are not converted:



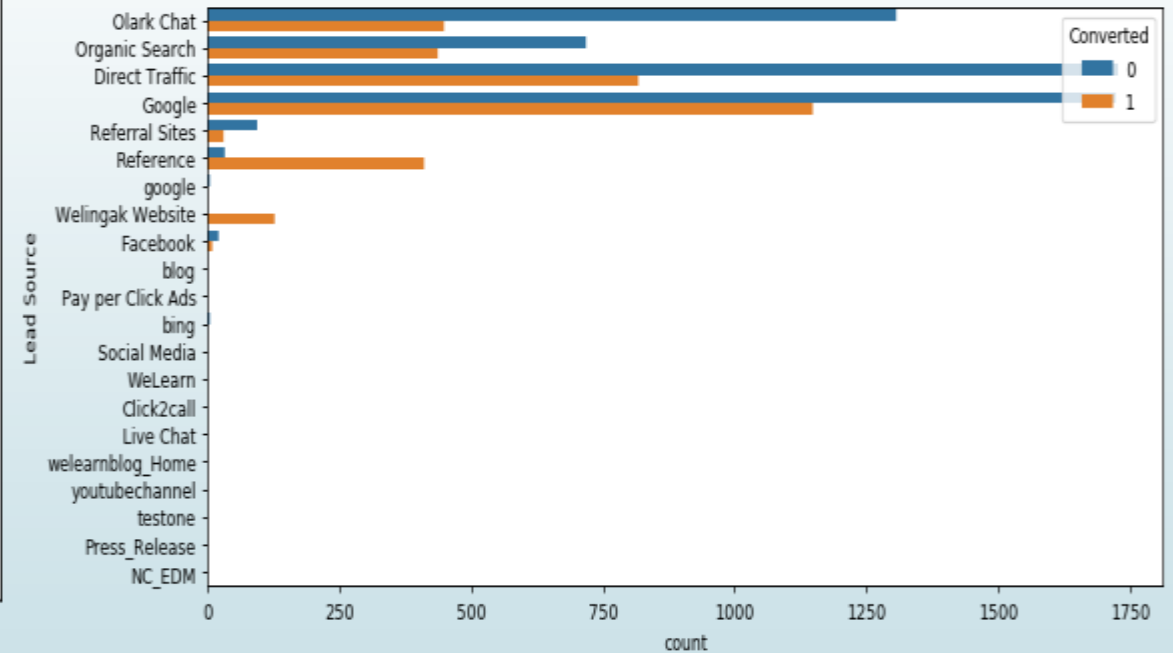
- On analysis we can observe that only 37.86 % of people are converted that means out of 100 only 37 people are converted to customers for the company.

Visualization 2:

G1: Lead_Origin v/s Converted

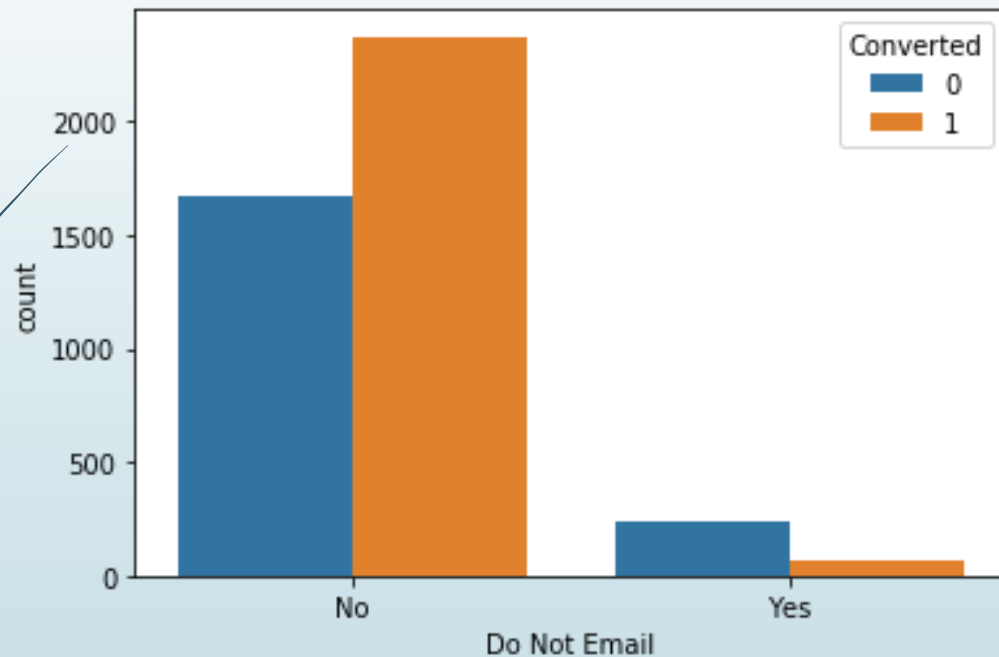


G2 Lead_Source V/s converted

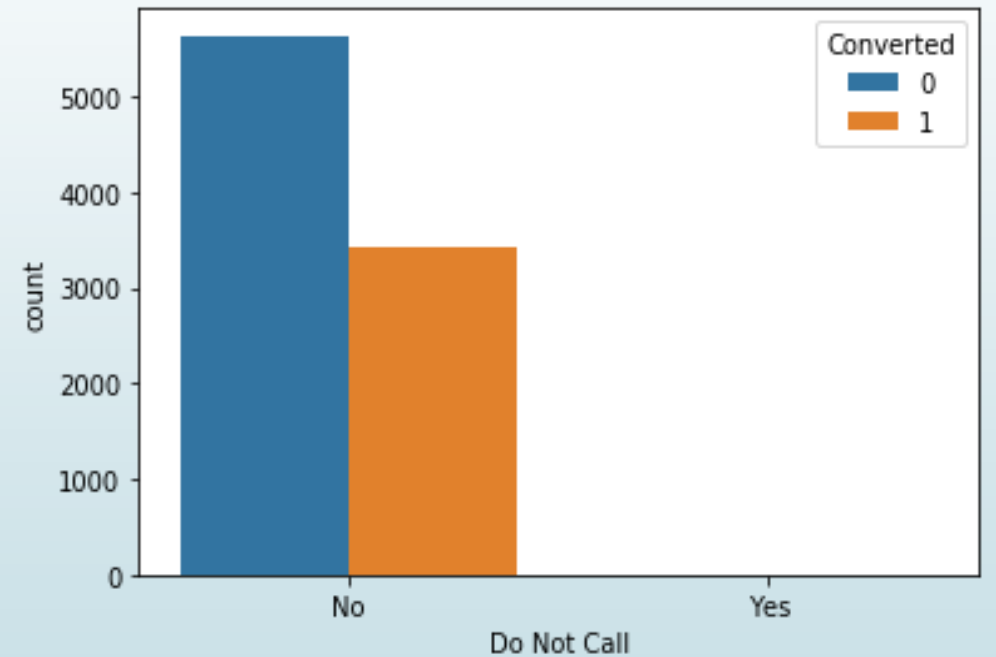


Visualization 3:

G3: Do_not_email v/s converted

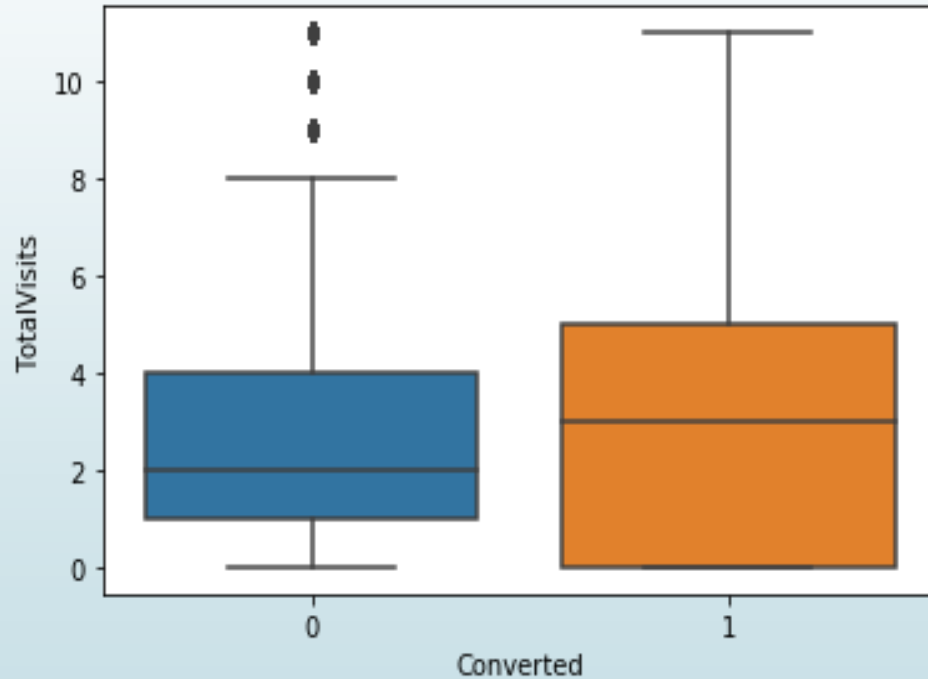


G4: Do_not_call v/s Converted

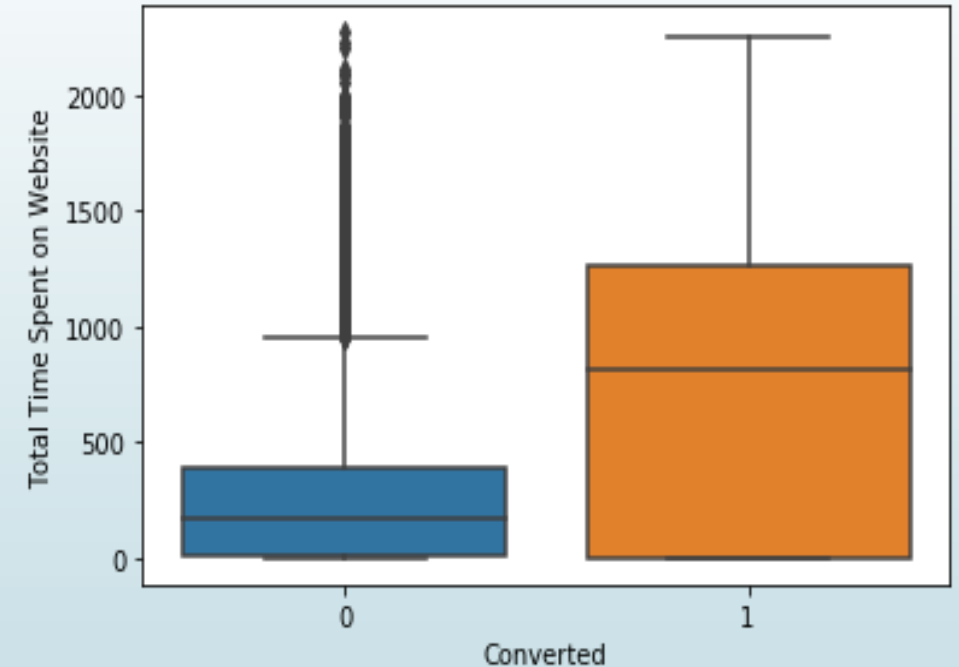


Visualization 4:

G5: Total_visits v/s converted

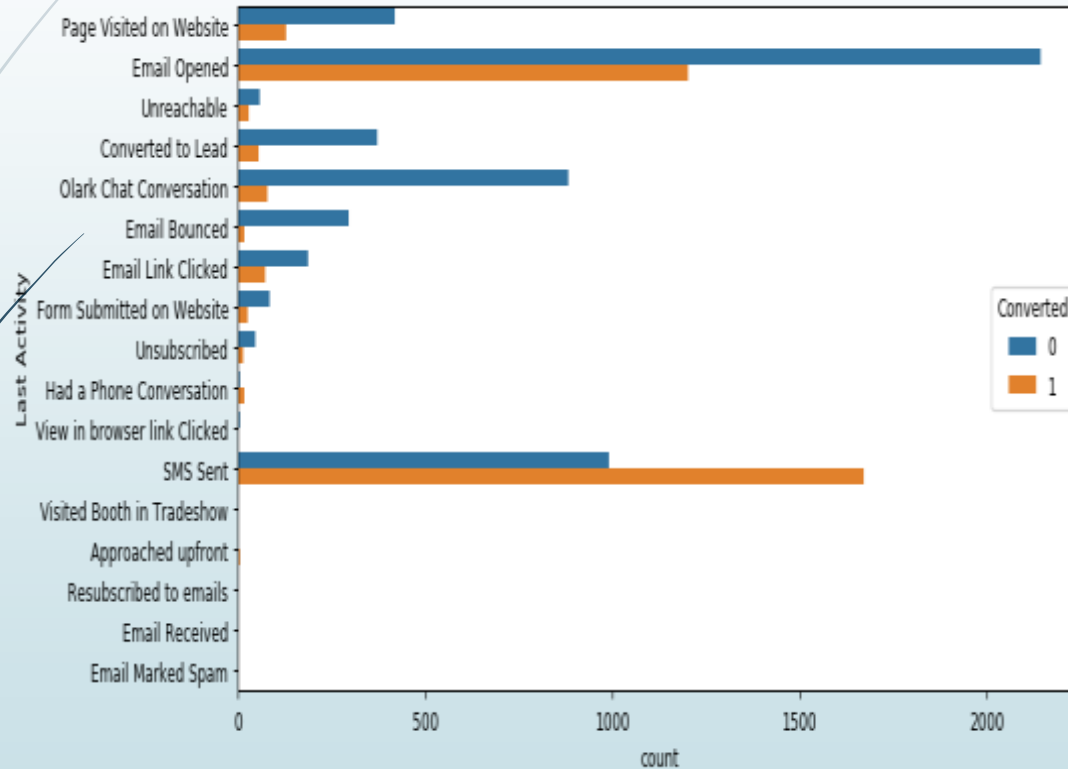


G6: Total time spent on website v/s converted

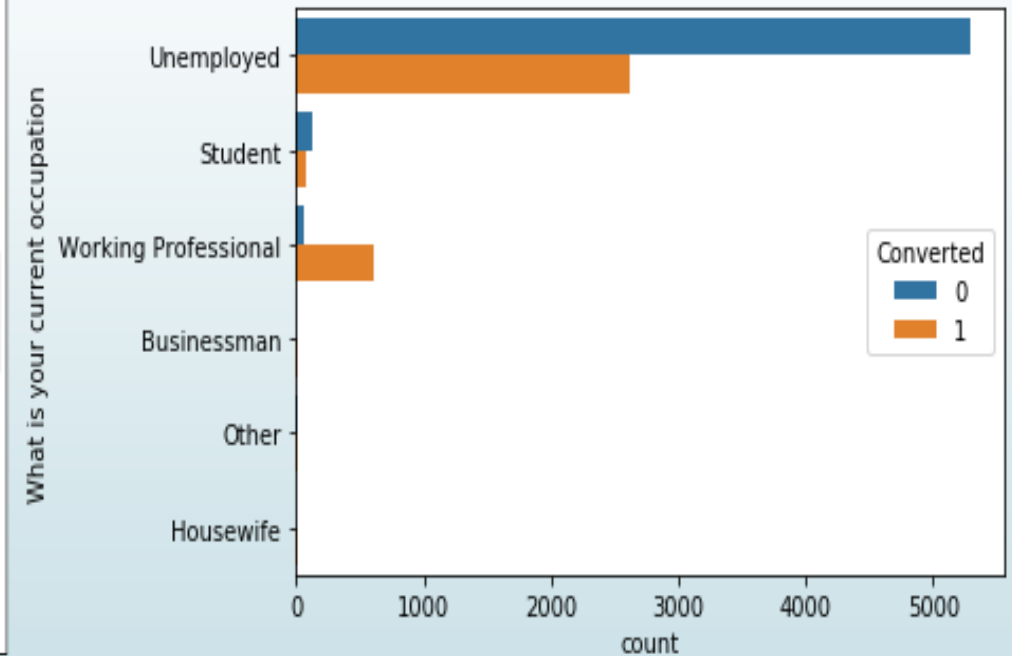


Visualization 5:

G7: Last activity v/s converted

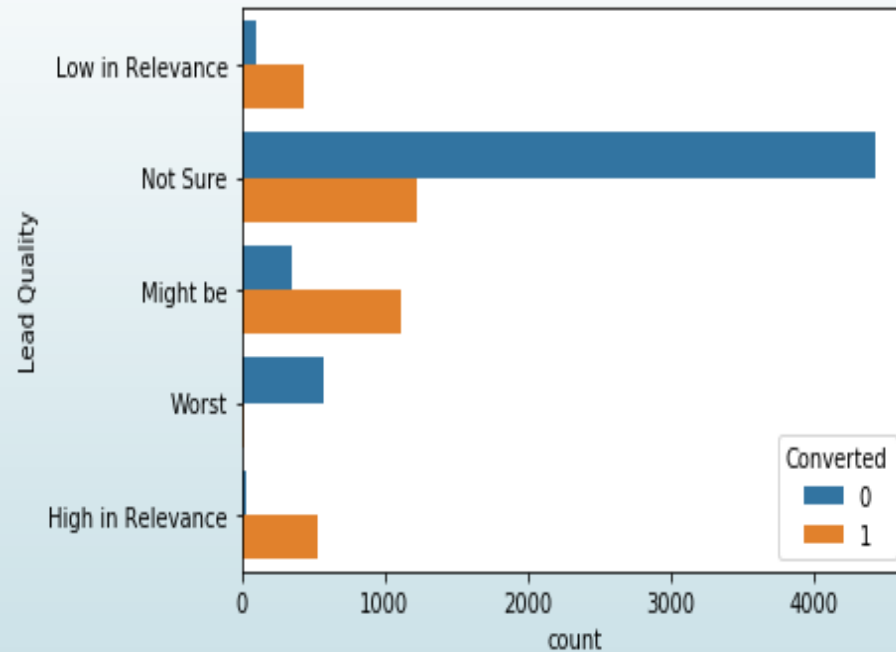


G8 :What is your current occupation v/s converted

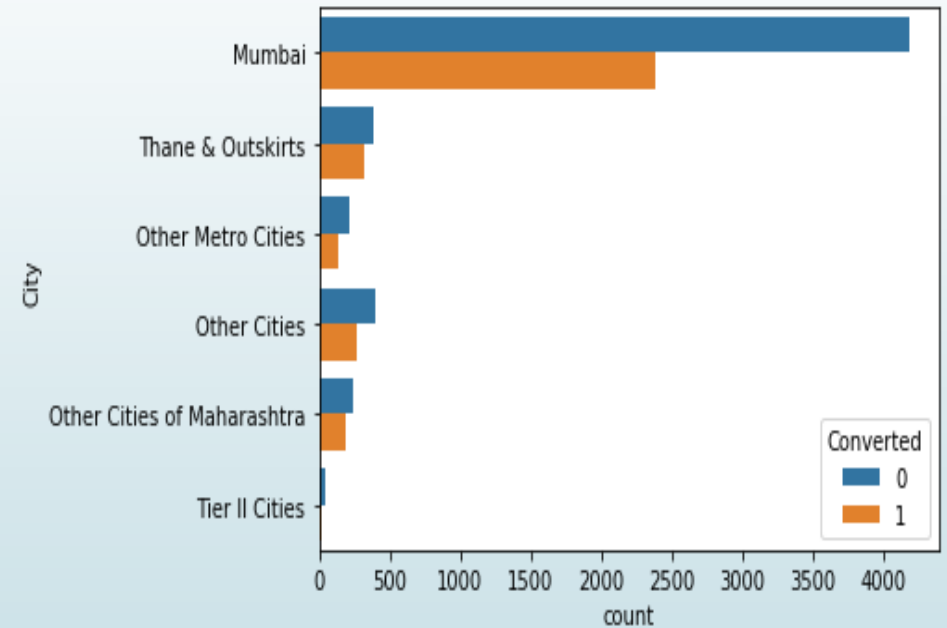


Visualization 6:

Lead quality v/s converted



City v/s converted





Conclusion of the visualization:-

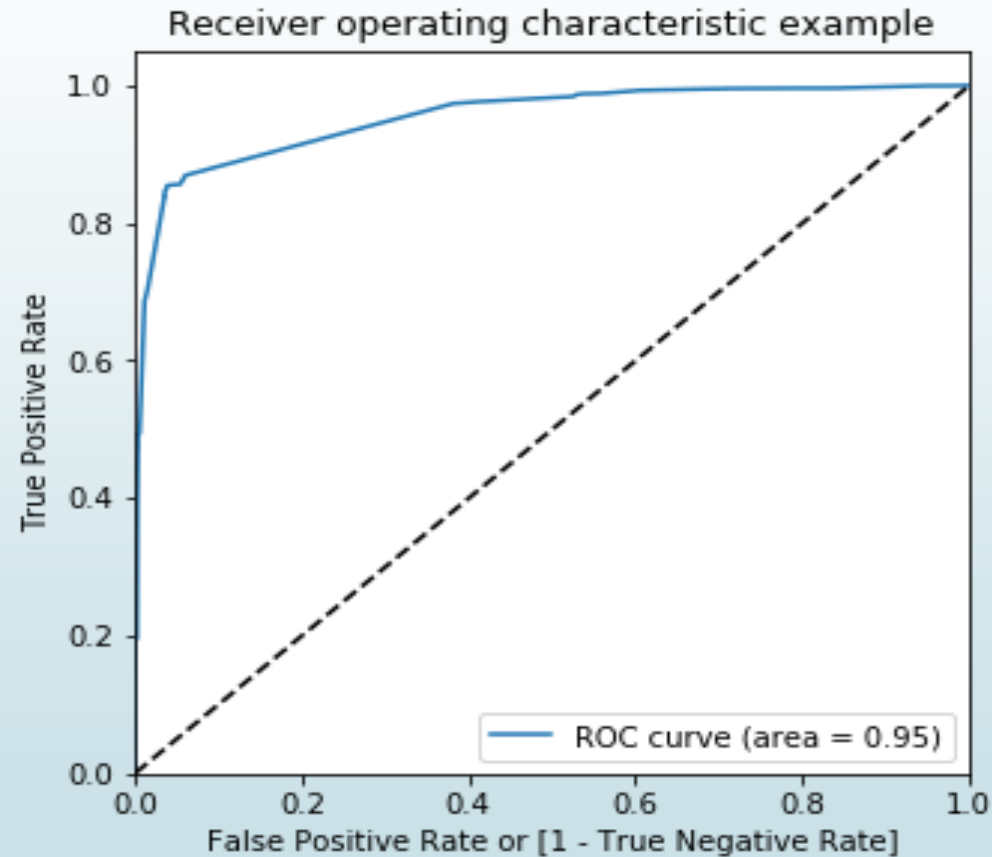
- From above we can conclude that:
- Graph 1: If the lead origin is Lead Add Form then probability that they will become lead is very high.
- Graph 2: Direct Traffic and Google is generating high amount of lead but the non leads are more so focus should be on direct lead and google.
- If the lead source is reference then the prob of becoming lead is very high.
- Graph 3: for do_not_email being yes or no we can observe that convergence is 0.
- Graph 4: for do_not_call being yes or no we can observe that convergence is 0.
- Graph 5: If total visits is high then convergence is more and if it is less then convergence is low.



Conclusion 2:

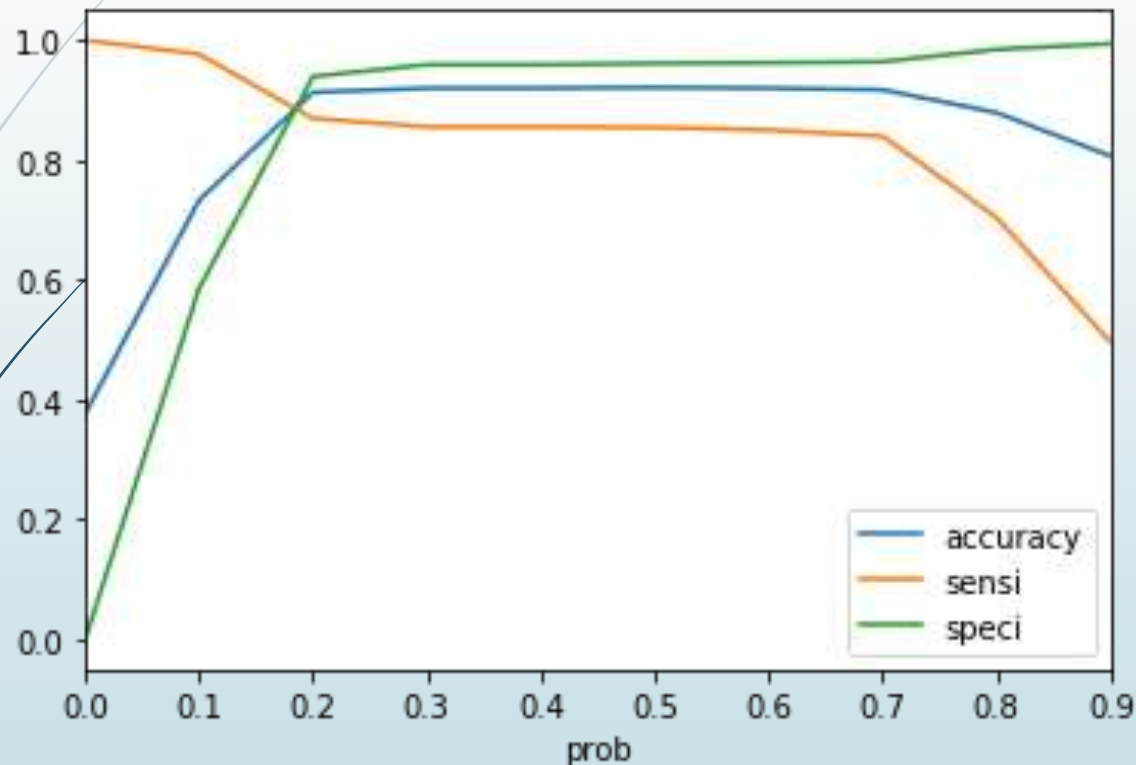
- Graph 6: From the graph of total time spend on website v/s converted we can conclude that if more time is spend on website then chance of being converted is more.
- Graph 7: If the last activity is Email Opened then the lead is higher . If the last activity is SMS sent then the probability that it is going to be lead is high
- Graph 8 : Unemployed leads are the most in numbers. If the occupation is Working Professional then the probability that it is going to be lead is high
- Graph 9: for lead quality being not sure the rate of being converted is more likely to be 0.
- Graph 10: for city in Mumbai the rate of being converted is high for both 0 and 1 .

ROC Curve:



➡ From curve we can conclude that area under the ROC curve is 0.95. This indicated that our graph is good.

Plot accuracy sensitivity and specificity for various probabilities:-



From the curve we can conclude that the cutoff for the graph with sensitivity specificity and accuracy is 0.18.



CONCLUSION:



Model Analysis :-

1. Overall accuracy on Test set: 0.9130
2. Sensitivity of our logistic regression model: 0.8694
3. Specificity of our logistic regression model: 0.9393
4. False positive rate: 0.60629
5. Positive predictive value : 0.89649
6. Negative predictive value:0.92256



Conclusion from model building

Top 3 variables in model, that contribute towards lead conversion are:

- ↳ Total Time Spent on Website
- ↳ Last Notable Activity
- ↳ Total Visits



Statistical conclusion:-

Our Logistic Regression Model is decent and accurate enough with 91.3 % Accuracy on Test Set, 86.94 % Sensitivity and 93.93 % Specificity.

We can vary these parameters by varying the cut-off value and thus predict Hot leads based on scenarios like availability of extra resources and vice-versa.



Business conclusion:-

X Education Company can improve the overall conversion rate by following:

1. Increase user engagement on their website since this helps in higher conversion i.e. try to improve total visits.
2. Increase on sending SMS notifications since this helps in higher conversion by sending them messages.
3. Get Total visits increased by advertising etc. since this helps in higher conversion thus will help user to interact.



Thankyou