Lead Score Case Study

Submitted By:

Supriya Pawar

Problem Statement

□ X Education sells online courses to industry professionals. The company markets its courses on several websites and search engines like Google.

□X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say,theyacquire100 leads in a day ,only about 30 of them are converted.

□To make this process more efficient, the company wishes to identify the most potential leads, also known as 'HotLeads'.

□If they successfully identify this set of leads ,the lead conversion rates should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone.

Business Goals

□X education wants to know the most promising leads.

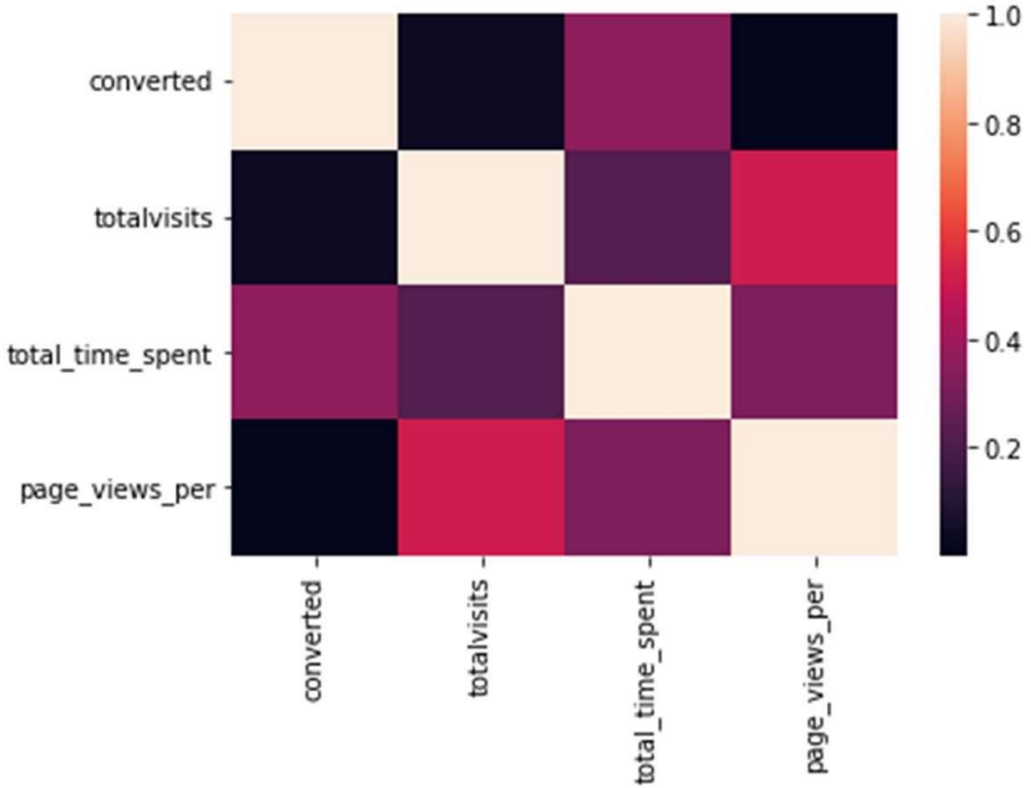
□The company needs a model wherein you a lead score is assigned to each of the leads such that the customers with higher lead score have a higher conversion chance.

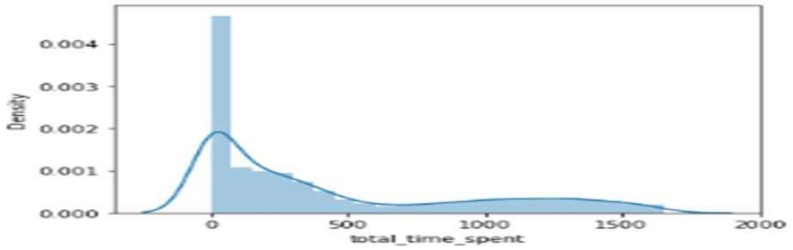
□Deployment of the model for the future use.

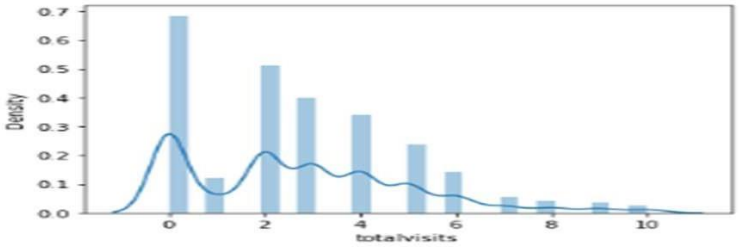
Solution Strategy

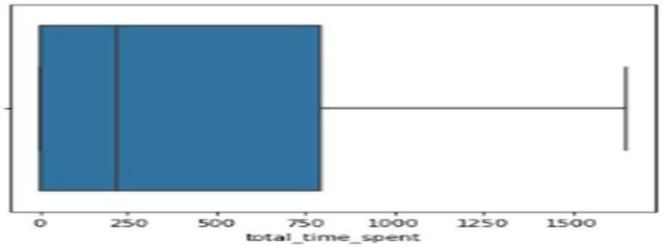
- Data cleaning and data manipulation.
  - Check and handle NA values and missing values.
  - Drop columns, if it has more than 40% nulls as it will not be useful for analysis.
  - Imputation of the values ,if necessary.
  - Check and handle outliers in data.
- EDA (Exploratory Data Analysis)
  - Univariate data analysis: value count, distribution of variable etc.
  - Bivariate data analysis: correlation coefficients and pattern between the variables etc.
- Feature Scaling & Dummy Variables and encoding of the data.
- Classification technique: Logistic regression used for the model making and prediction.
- Validation of the model.
- Model presentation.
- Conclusions and recommendations.

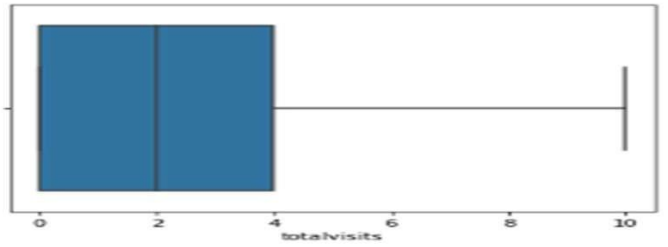
HeatmapofContinuousVaíiables

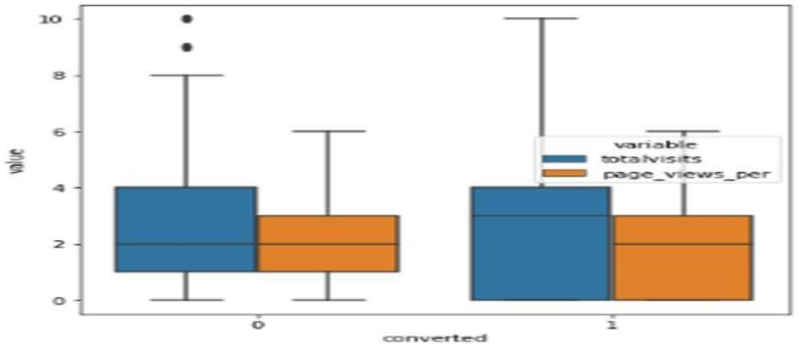


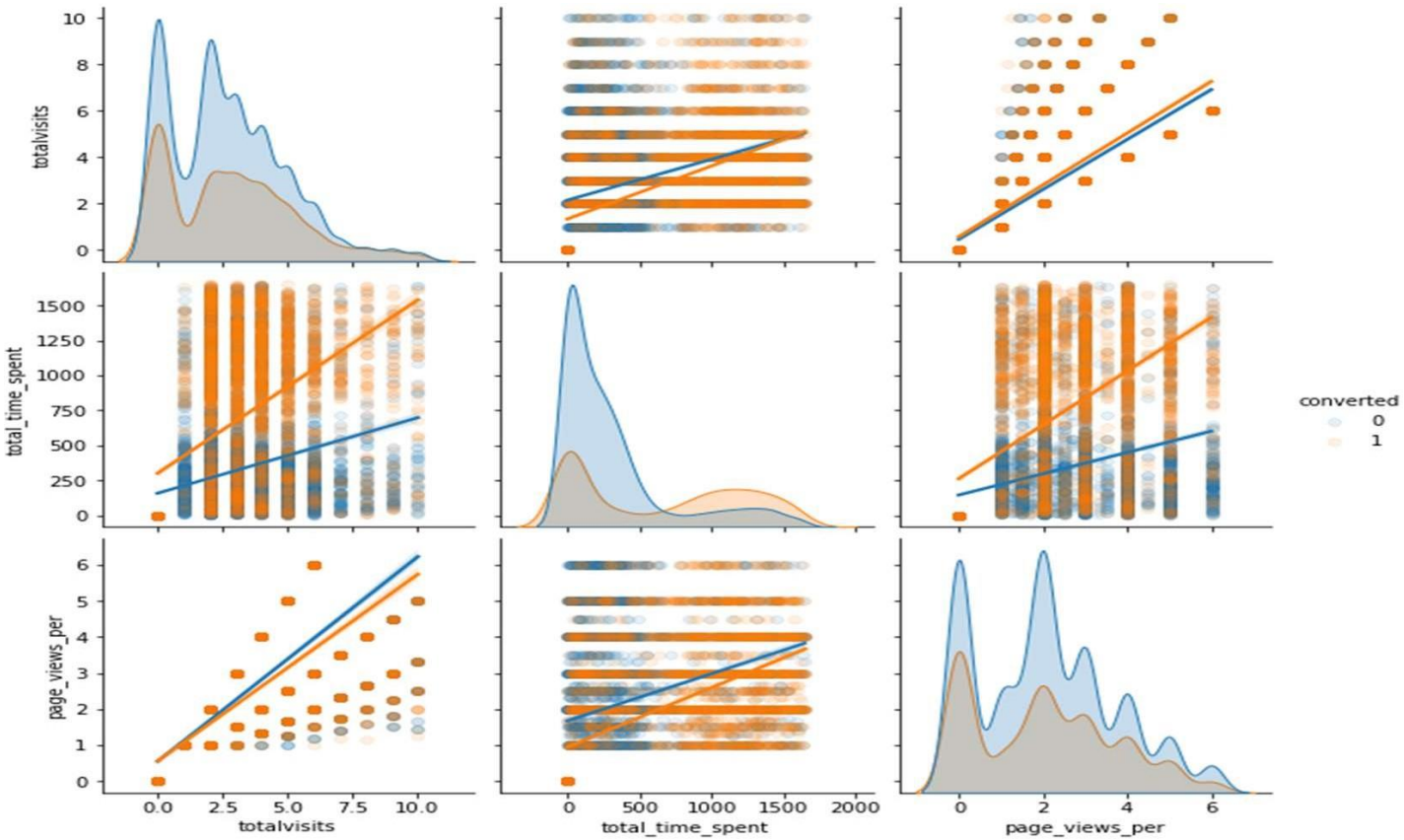

Density
0.004
0.003
0.002
0.001
0.000
0
500
1000
1500
2000
total_time_spent

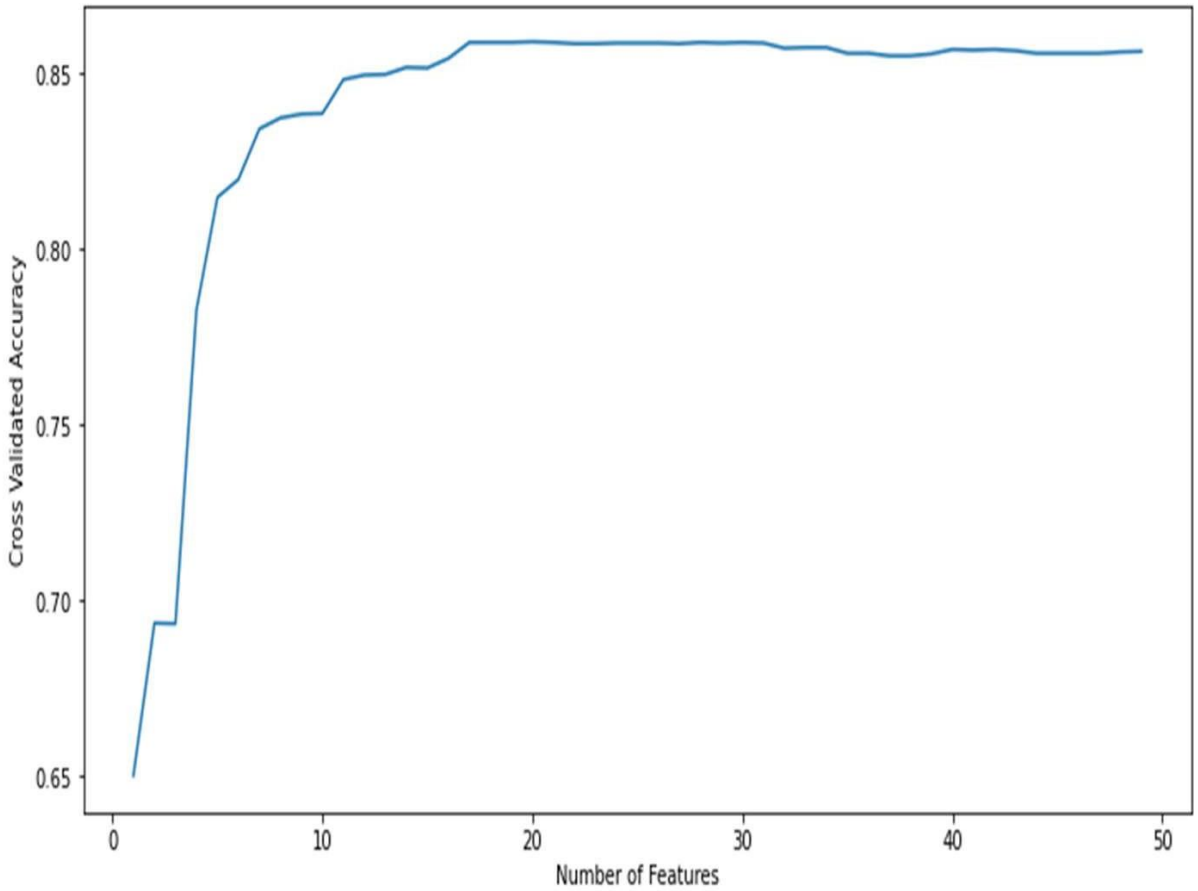

Density
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
2
4
6
8
10
totalvisits


0
250
500
750
1000
1250
1500
total_time_spent

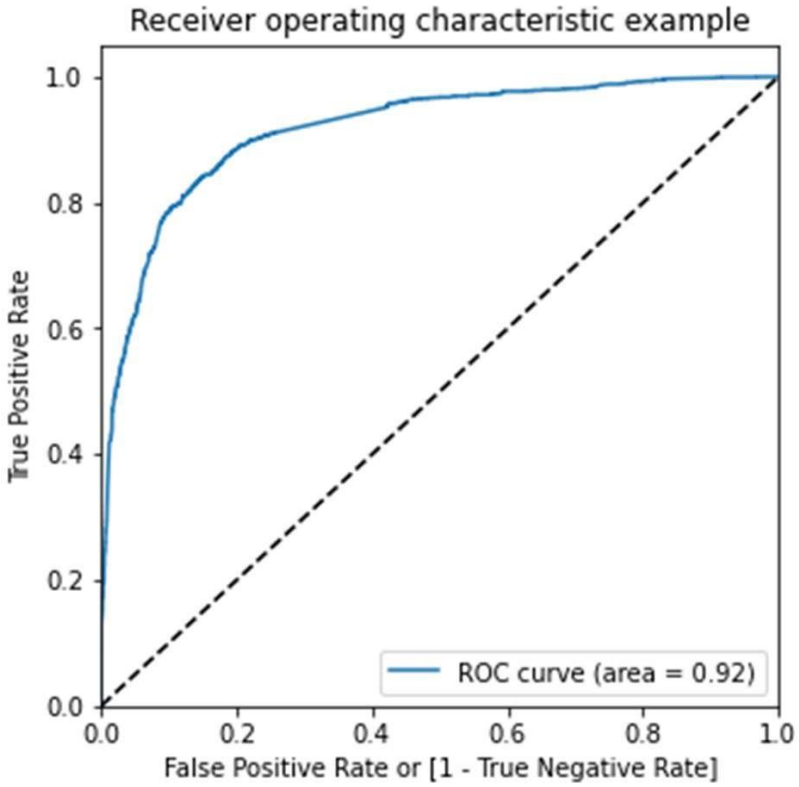

0
2
4
6
8
10
totalvisits

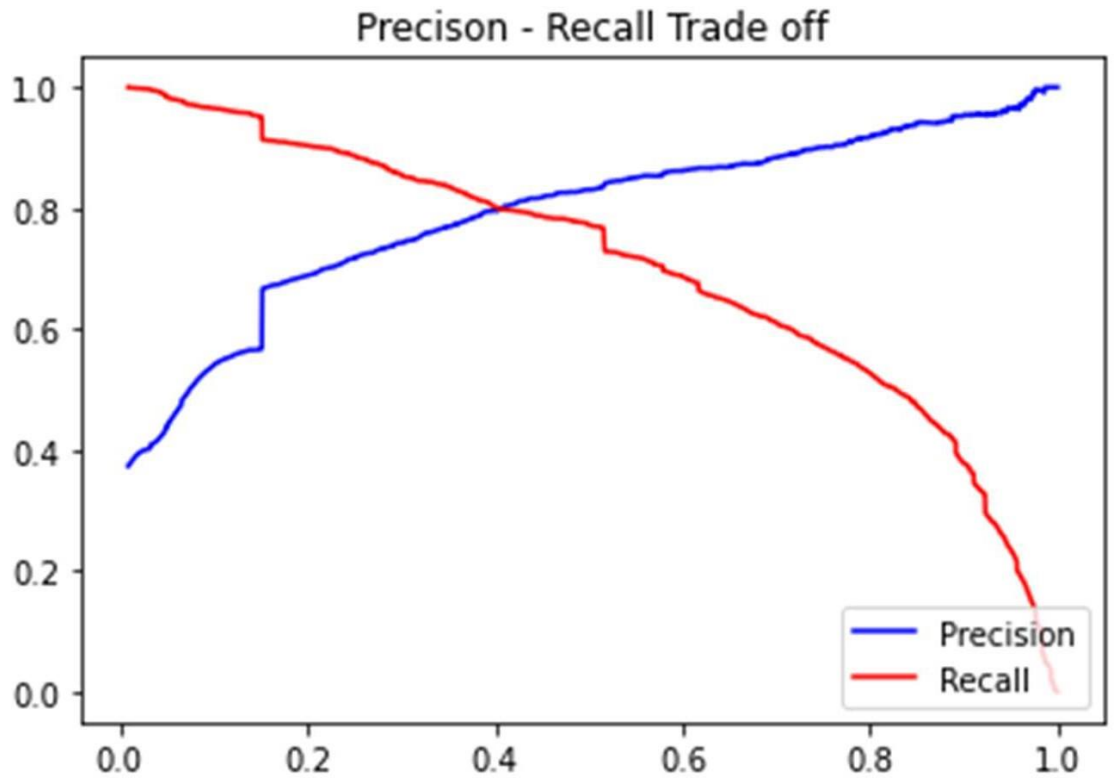

value
10
8
6
4
2
0
variable
totalvisits
page_views_per
0
1
converted

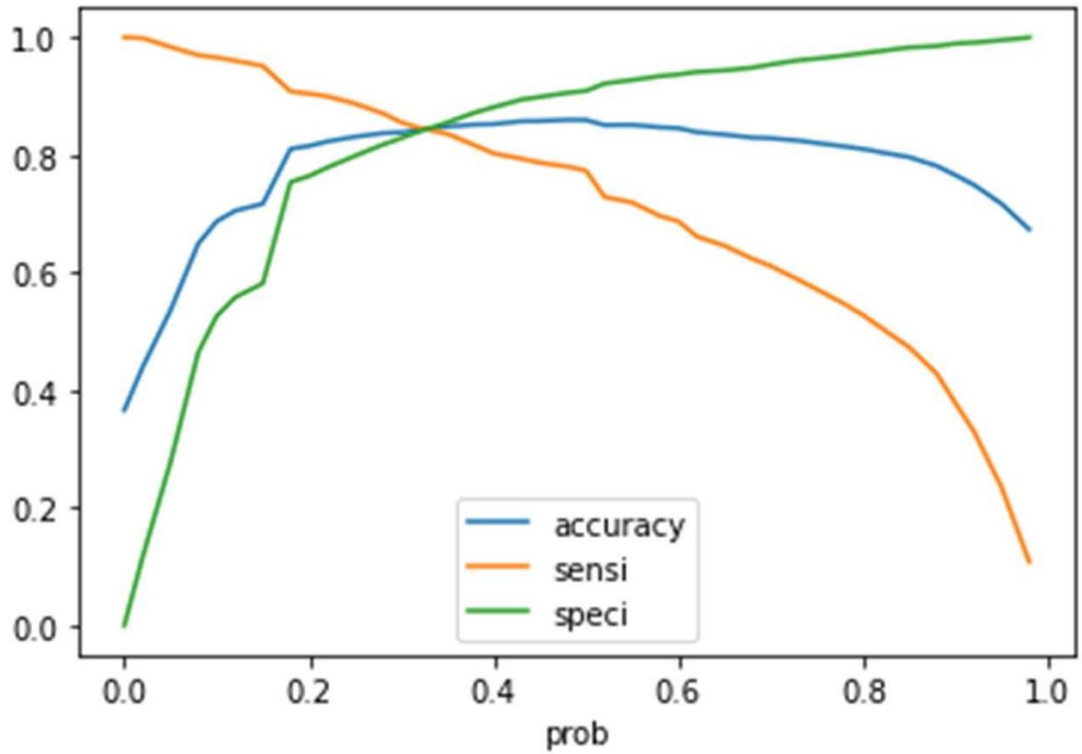

totalvisits
total_time_spent
page_views_per
converted
0
1
0.0
2.5
5.0
7.5
10.0
0
500
1000
1500
2000
0
2
4
6



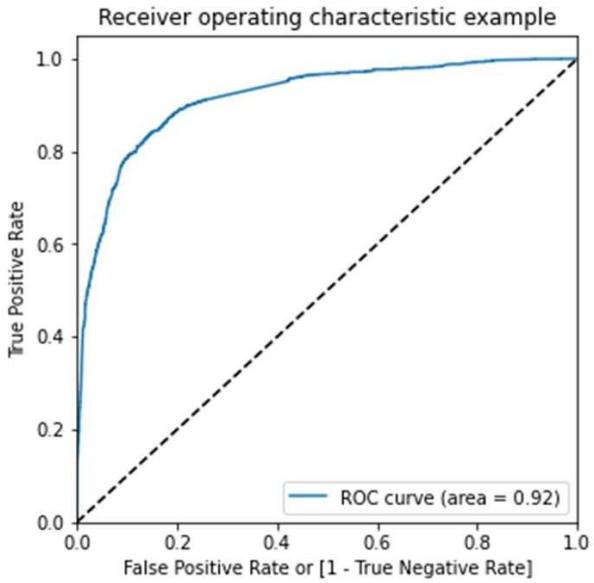
best features are 20.0 which gives 0.8587807097361237 accuracy.

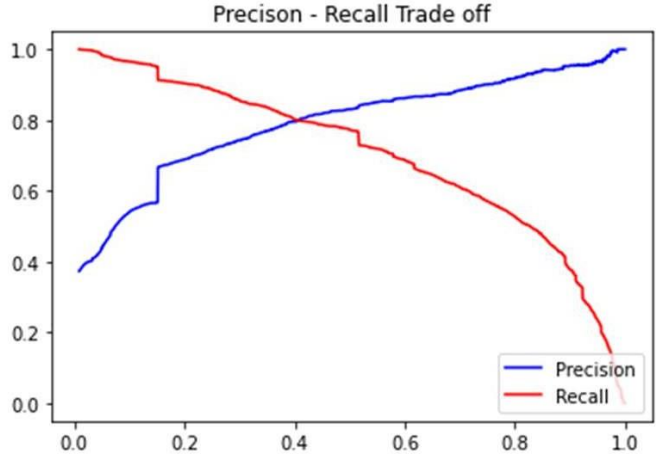

Receiver operating characteristic example
True Positive Rate
False Positive Rate or [1 - True Negative Rate]
ROC curve (area = 0.92)
0.0
0.2
0.4
0.6
0.8
1.0

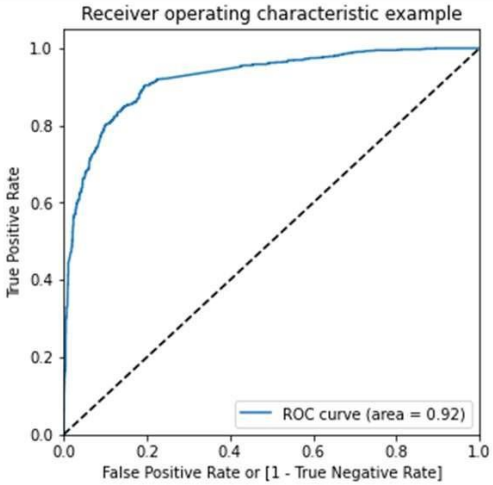

Precison - Recall Trade off
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Precision
Recall

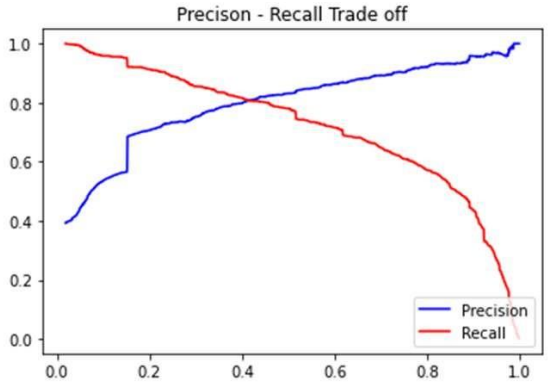



```
Optimum cut-off value is: 0.32
```


Receiver operating characteristic example
True Positive Rate
False Positive Rate or [1 - True Negative Rate]
ROC curve (area = 0.92)
0.0
0.2
0.4
0.6
0.8
1.0


Precison - Recall Trade off
Precision
Recall
0.0
0.2
0.4
0.6
0.8
1.0





Now, we have an overall accuracy of about 0.85 on our Logistic Regression model. This means that there is 85% chance that our predicted leads will be converted. This also mean that it meets the CEO's target of atleast 80% lead conversion.

Recommendations:

X Education needs to improve their engagement on Welingak website, they also must target the working professionals as they are the users that are more attracted by the services X Education provides. X Education can have a huge boost by starting a referral program with incentives to both the person giving reference and the person joining X Education because of the referral. They can also improve the lead conversion score by increasing the time spent on their website and by improving the Olark chat service.

Conclusion:

We can see that all the objectives stated above are satisfied and have achieved a accuracy of 85%, which indicates that we have met the CEO's expectation. We have also found the key factors that indicate if a particular user can be converted, this would help the sales team of X Education use their limited human and financial effectively and efficient