



# Lead Scoring Case Study

Group Activity

Ramakrishna Appalla

Karan Wadhawan



# AGENDA

- ☐ Problem Statement
- ☐ Overall Approach
- ☐ Analysis
- ☐ Conclusion



# Problem Statement

An education company named X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses. 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%. There are a lot of leads generated in the initial stage (top) but only a few of them come out as paying customers from the bottom. In the middle stage, you need to nurture the potential leads well (i.e. educating the leads about the product, constantly communicating etc. ) in order to get a higher lead conversion. X Education has appointed you 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 you 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%.



# Overall Approach

## Our Goals of the Case Study:

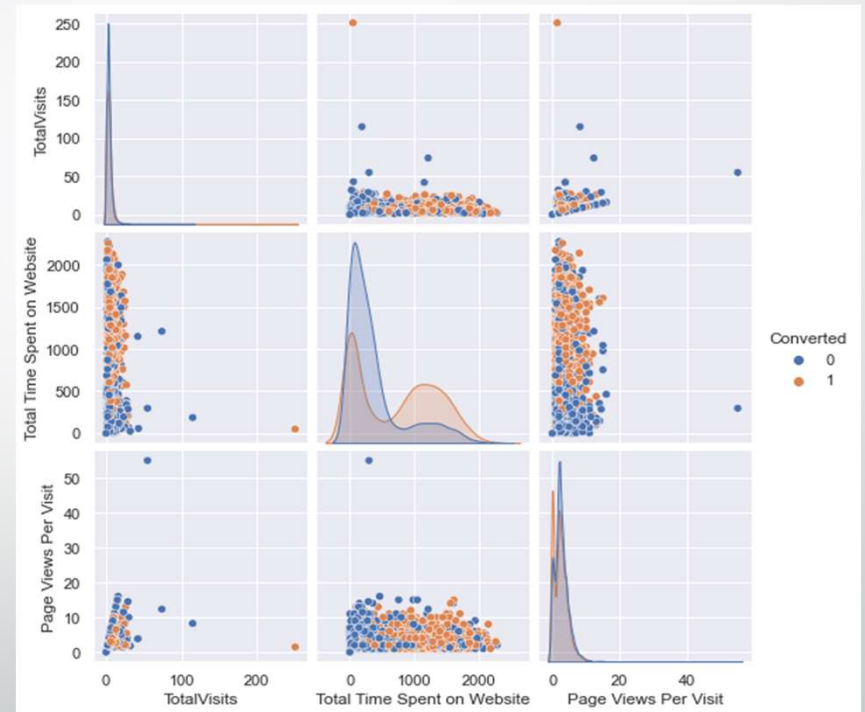
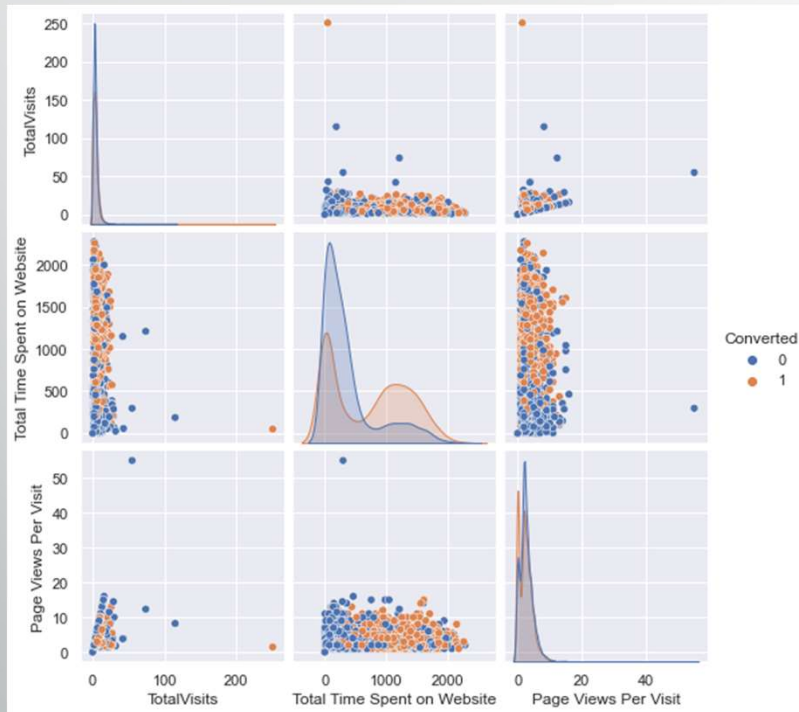
- To **build a logistic regression model to assign a lead score** between 0 and 100 to each of the leads which can be used by the company to target potential leads.
- To **adjust to if the company's requirement changes** in the future so you will need to handle these as well.



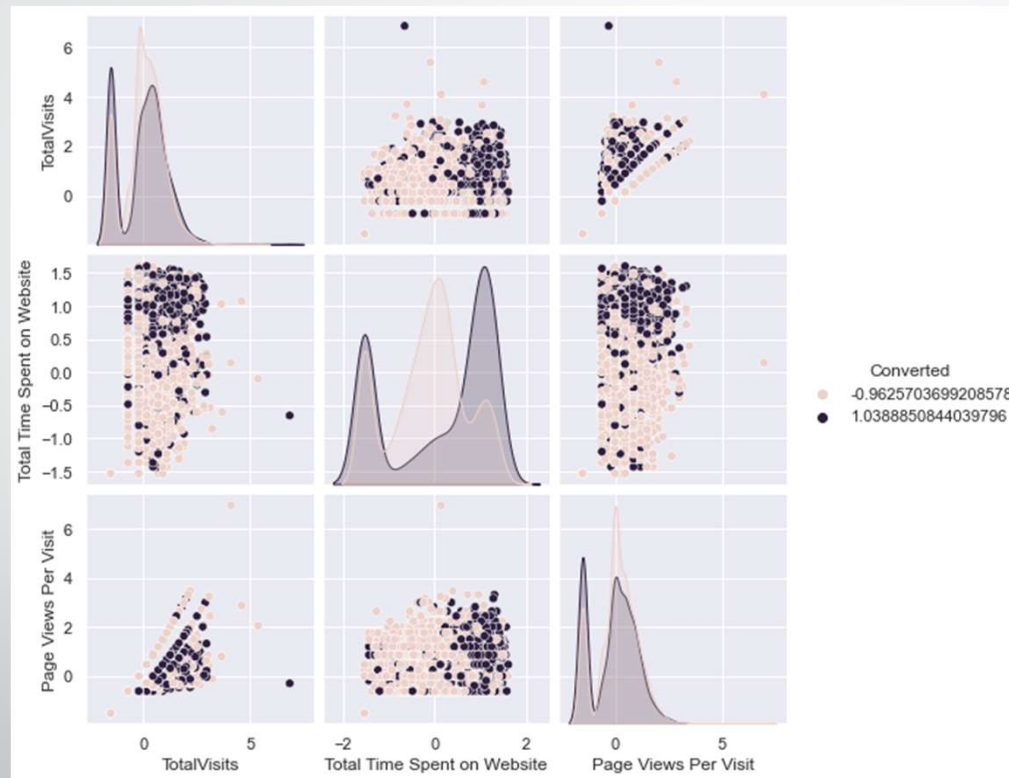
# ANALYSIS

1. Read and understand the data
2. Clean the data
3. Prepare the data for Model Building
4. Model Building
5. Model Evaluation
6. Making Predictions on the Test Set

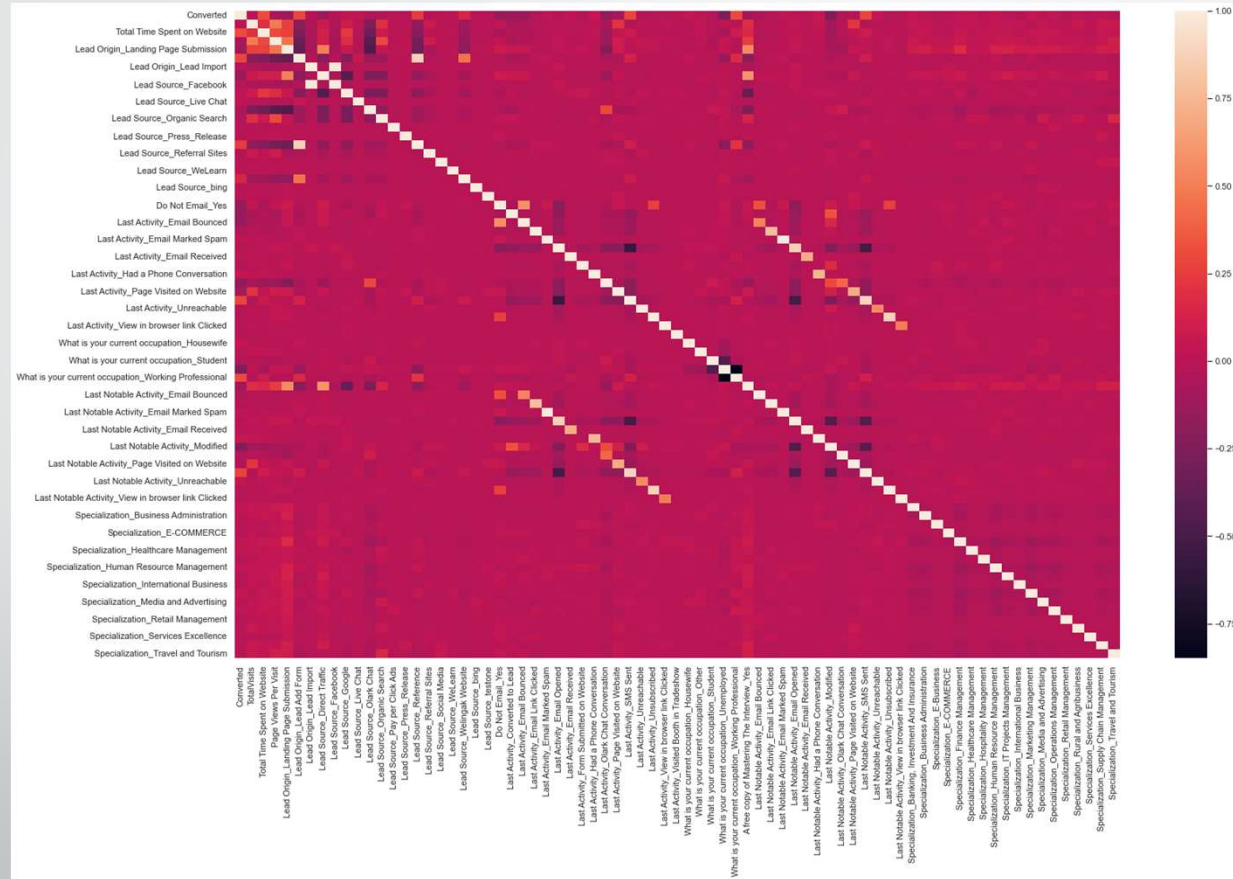
# Preparing the data for modelling



# Preparing the data for modelling(contd)

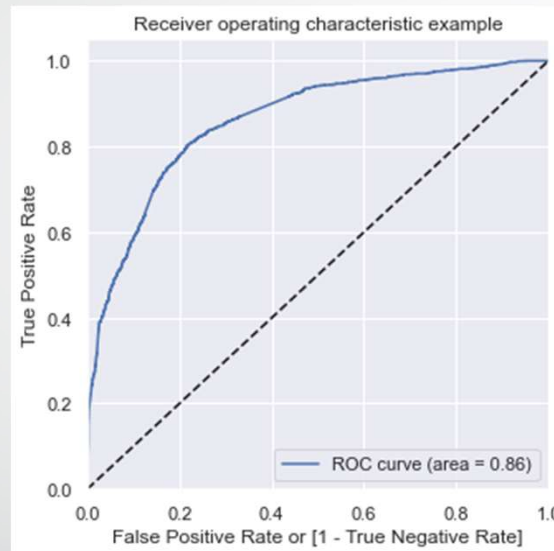


# Correlations



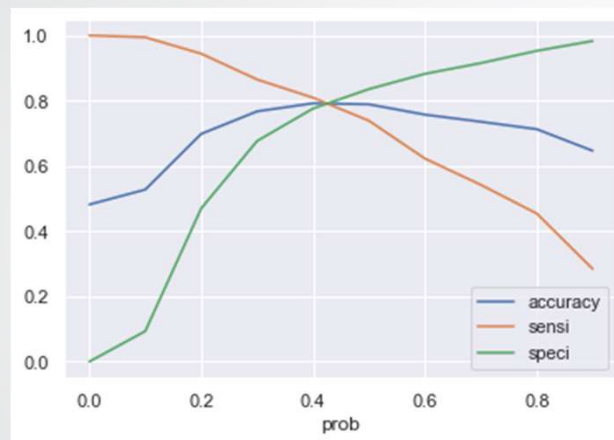


# Finding the Optimal Cutoff



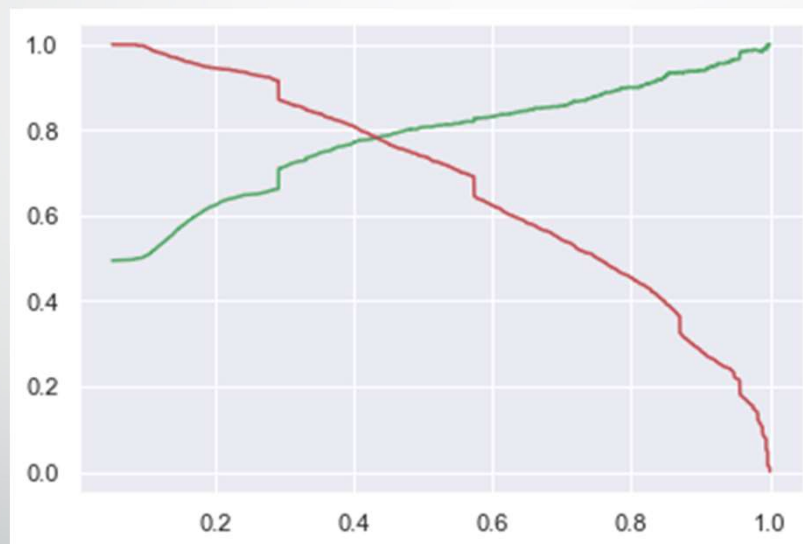
The area under the curve of the ROC is 0.86 which is quite good. So we seem to have a good model. Let's also check the sensitivity and specificity tradeoff to find the optimal cutoff point

## Plotting the values of accuracy, sensitivity, and specificity at different values of probability cutoffs



As we can see that around 0.42, you get the optimal values of the three metrics. So let's choose 0.42 as our cutoff now

# Precision and Recall trade-off



# Final Observation

Let us compare the values obtained for Train & Test:

## Train Data:

- Accuracy : 92.29%
- Sensitivity : 91.70%
- Specificity : 92.66%

## Test Data:

- Accuracy : 76.1%
- Sensitivity : 71.8%
- Specificity : 82.3%

The Model seems to predict the Conversion Rate very well and we should be able to give the CEO confidence in making good calls based on this model.

Thus we have achieved our goal of getting a ballpark of the target lead conversion rate to be around 80% . The Model seems to predict the Conversion Rate very well and we should be able to give the CEO confidence in making good calls based on this model to get a higher lead conversion rate of 80%.

# Recommendations

- ✓ The company **should make calls** to the leads coming from the lead sources "Welingak Websites" and "Reference" as these are more likely to get converted.
- ✓ The company **should make calls** to the leads who are the "working professionals" as they are more likely to get converted.
- ✓ The company **should make calls** to the leads who spent "more time on the websites" as these are more likely to get converted.
- ✓ The company **should make calls** to the leads coming from the lead sources "Olark Chat" as these are more likely to get converted.
- ✓ The company **should make calls** to the leads whose last activity was SMS Sent as they are more likely to get converted.
- ✓ The company **should not make calls** to the leads whose last activity was "Olark Chat Conversation" as they are not likely to get converted.
- ✓ The company **should not make calls** to the leads whose lead origin is "Landing Page Submission" as they are not likely to get converted.
- ✓ The company **should not make calls** to the leads whose Specialization was "Others" as they are not likely to get converted.
- ✓ The company **should not make calls** to the leads who chose the option of "Do not Email" as "yes" as they are not likely to get converted.

# Summary

- There are a lot of leads generated in the initial stage (top) but only a few of them come out as paying customers from the bottom. In the middle stage, we would need to nurture the potential leads well (i.e. educating the leads about the product, constantly communicating etc.) in order to get a higher lead conversion.
- First of all, sort out the best prospects from the leads you have generated. '**TotalVisits**', '**Total Time Spent on Website**', '**Page Views Per Visit**' which contribute most towards the probability of a lead getting converted.
- Then, We must keep a list of leads handy so that you can inform them about new courses, services, job offers and future higher studies. Monitor each lead carefully so that you can tailor the information you send to them. Carefully provide job offerings, information or courses that suits best according to the interest of the leads.
- A proper plan to chart the needs of each lead will go a long way to capture the leads as prospects.
- Focus on converted leads. Hold question-answer sessions with leads to extract the right information we need about them.
- Make further inquiries and appointments with the leads to determine their intention and mentality to join online courses.