



# Lead Scoring Case Study

## SUBMITTED BY:

-DIKSHA MISHRA

-TANVI PRADHAN

# Problem Statement

X Education sells online courses to industry professionals. 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

# Business Goal

X Education needs help in selecting the most promising leads, i.e. the leads that are most likely to convert into paying customers.

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 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%.

# Strategy

- Source the data for analysis
- Clean and prepare the data
- Exploratory Data Analysis
- Feature Scaling
- Splitting the data into Test and Train dataset
- Building a logistic Regression model and calculating Lead Score
- Evaluating the model by using different metrics - Specificity and Sensitivity or Precision and Recall
- Applying the best model in Test data based on the Sensitivity and Specificity Metrics

# Variables Impacting the Conversion Rate

- Do Not Email
- Total Visits
- Total Time Spent On the Website
- Lead Origin
- Lead Source
- Last Activity
- Current Occupation
- Last Notable Activity

# Recommendations

- The sales team of the X-Education should focus on the leads having lead origin - lead add form, occupation - Working Professional, Lead source - Wellingak website
- Hot Leads are identified as 'Customers having a lead score above 35. The Sales Team of the company should first focus on the 'Hot Leads'
- The 'Cold Leads' (Customers having lead score  $\leq 35$ ) should be focused on after the Sales Team is done with the 'Hot Leads'
- There are many important variables like city, specialization, and occupation which can potentially explain Conversion better. It is important for the management to make few of this information mandatory to fill so that we can use them in our model and build important decisions for the business
- We have high recall scores than precision scores. Hence this model has the ability to adjust to the company's requirements in coming future
- Sensitivity will ensure that almost all leads who are likely to Convert are correctly predicted whereas high Specificity will ensure that leads that are on the brink of the probability of getting Converted or not are selected
- It's better to focus least on customers who do not want to be called about the course
- If the Last Notable Activity is Modified, he/she may not be the potential lead



**Thank You**