Summary Report for Lead Case Study

Problem Statement

An education company named 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. Now, although X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted. What you need to do?

- 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.

Report

In order to increase the conversation rate of the leads, we have prepared a logistic regression model to target only the genuine potential leads so that the sales representatives can focus more on those leads and came up to below observations

On inspecting the data, we get to know that below data attributes have high null percentage which indicates that most of the people are not trying to fill these data attributes and hence it is less significant for deciding the lead score .So we decided to drop these.

- Tags
- Lead Profile
- Lead Quality
- Asymmetrique Activity Index
- Asymmetrique Activity Score
- Asymmetrique Profile Index
- Asymmetrique Profile Score
- How did you hear about X Education

On further analysis, we get to know that below data attributes have majority of similar responses and will not help in deciding which lead is more probable.

- What matters most to you in choosing a course
- Country
- 'I agree to pay the amount through cheque

- Get updates on DM Content
- Receive More Updates About Our Courses
- Update me on Supply Chain Content
- X Education Forum
- Through Recommendations
- Digital Advertisement
- Newspaper Article
- Magazine
- Do Not Call

Out of all attributes of leads data, the major attributes which we need to target to assign the lead score, based on our model are:

- **Total Time Spent on Website:** Person who are spending more time on the website are more probable to be genuine leads who can convert
- **Lead Source:** With lead source as Direct Traffic, Google, Olark Chat and Organic Search are not potential for taking the education program but source who are coming with reference are needs to be targeted more.
- **What is your current occupation:** This is very important factor for deciding lead score, person who are unemployed along with working professionals and student are likely to convert and join the courses offered
- **Last Activity:** This also will help us deciding the lead score as person who has email opened and sms sent can be used for assigning the lead score.
- **Specialization:** This also help in deciding the lead score. Based on our model, people with hospitality management are less likely to be potential leads and people from Finance Management are likely to be potential leads.

Our model is providing the lead score for all lead with sensitivity and Accuracy around 80% and can help in increasing lead conversion rate up to 80% which is way more than previous lead accuracy rate as 30%.

The Lead Scores are also calculated to help the CEO and the Sales team to target potential leads accordingly.