

# **Lead Scoring Assignment**

## **Summary**

A data analyst project was undertaken to predict which customers are likely to convert on an education portal. The portal belonged to 'X' education institute. We are attempting to look for ways to increase sales and efficiency while keeping costs in mind. The data given is most about digital activity and some other details. For example, we have a parameter to check how much a customer spends on the website?, what is his/her occupational background?, How did the user hear about 'X'? Whether they want to opt for email notifications? Etc.

From the given data exploratory data analysis and logistic regression was done in order to gather insights for business decisions. Only these results business analysis and domain analysis was done as far as possible to present a holistic picture. From the analysis stated we have found a certain target audience to contact who have higher chances of generating better conversion rates.

## **Target Audience**

Customers who are Unemployed and Working Professional as they have a very high lead conversion rate. Customers whose last activity is 'SMS sent'. We do not know as of now what information this activity is conveying to the customer however, this is a very important strategy to keep in mind as this is generating almost 50% lead conversion rate. Customers whose Last activity has been Olark chat conversation should not be contacted, very poor leads. This can be researched further and if required investment in this activity can be removed. There is a strong negative relation to this variable against lead conversion both from EDA and ML insights. In order to address the volume, start filtering out the cold leads i.e. people with bounced email, people who do not spend much time on the website etc.