

Summary

1. Lead scoring case study has been done using logistic regression model to meet the constraints as per business requirements.
2. There are a lot of leads in the initial stage but only a few of them are converted into paying customers. The most numbers of leads are from **INDIA** and in terms of city highest number are from **Mumbai**.
3. There are a few columns in which there is a level called 'Select' which basically means that the student had not selected the option for that particular column which is why it shows 'Select'. To get some useful data we have to convert it to NULL and then make compulsory selection at later steps according to the nature of data in that column. Likewise, Customer Occupation, Specialization, etc.
4. The high number of Total Visits & Total Time Spent on platform may be increasing the chances of lead to be converted.
5. The leads are joined course for Better Career Prospects, most of having Specialization from Finance Management. Leads from HR, Finance & Marketing Management specializations are of high probability to convert.
6. Talking to last notable Activity, making improvement in customer engagement through email & calls will help to convert leads. As the leads which are opening email have high probability to convert, same as sending SMS will also benefit.
7. Most of leads current occupation is Unemployed, which means gave more focus on unemployed leads.