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 paid 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 column which is why it shows 'Select'. To get some useful data we have to make compulsory selection. Example : Specialization, etc.

4) The company should make calls to the leads who spent "more time on the websites" and high number of "total visits" as these are more likely to get 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 high probability to convert.

6) Most of leads current occupation is Unemployed, which means gave more focus on unemployed leads.

7) Talking to last notable Activity. making improvement in customer engagement through email and calls would help to convert leads. As the leads which are opening email have high probability to convert, same as sending SMS will also benefit.