

Summary

1. The given Lead scoring case study will be done using a logistic regression model to meet the business requirements.
2. At the start of the case study, it can be observed that there are many leads but only a few of them are converted into paying customers.
3. It can be observed that most numbers of leads belong to "India" and Mumbai is the city that trends highest in terms of lead.
4. "Select" values were observed in a few of the columns in the dataset. It generally means that the person did not select any of the options and left it as blank. It is best to treat "Select" as a null value
5. The increasing chances of lead to be converted can be due to high number of total visits by a person, Lead Origin, Lead Add Form, total time spent by a person on the platform.
6. It can be observed that last notable activity, improve customer engagement through emails & calls and for the one which are opening email have high probability to convert. It is suggested that sending SMS will help in conversion.