SUMMARY

Here, we have been provided the data of the leads dataset which consists of many different attributes such as total time spent on website, Lead origin, Lead source and others. Initially I have tried to read and understand the data. The provided data gives an understanding of the different attributes that can be analysed to get the probability of conversion of leads.

After reading the data, it is important to clean or preprocess the data. The columns with large number of nulls have values have been dropped. There are few columns which have select values which means that individual has not selected any option for that column. To get more information about the leads, the selection can be made necessary for different options. After removing the null values, the data has been visualized to perform exploratory data analysis, that gives an idea of the role of different attributes on the conversion rate.

The designed model shows that leads who spent high time on the website can be potential targets. Engagement of the customers can be increased through related emails and calls, as it was seen that leads which are opening email are hot leads. To add further, the unemployed people with background in management are the most hot leads, after which students can be next target.