Summary

This case study is all about an Education company X which is struggling to get more candidates join their courses. The conversion rate was 30% earlier now company wants to increase their conversion rate to 80%.

Steps which is used to deal with data: -

1. **Cleaning data:** The data was having non values; few extra columns were needed to removed. Data preparation was done just to make data set operation as it was partially clean.

2. **EDA –**EDA was done to check the data and make it more operational, and did Bivariate and univariate analysis to find out the useful aspects regarding whole dataset.

3. **Dummy Variables**: The dummy variables were created and later on the dummies with 'not provided' elements were removed.

4. **Model Evaluation:** A confusion matrix was made. Later on the optimum cut off value (using ROC curve) was used to find the accuracy, sensitivity and specificity which came to be around 80% each.

5. **Prediction:** Prediction was done on the test data frame and with an optimum cut off as 0.35 with accuracy, sensitivity and specificity of 80%.

Conclusion

Lead Origin Lead Add Form are the ones the to whom the calls should be made as conversion rate is high

The company should make calls to the leads who are the "working professionals" as they are more likely to get converted.

The company should make calls to the leads who spent "more time on the websites" as these are more likely to get converted.

The company should make calls to the leads coming from the lead sources "Olark Chat" as these are more likely to get converted.

The company should not make calls to the leads who chose the option of "Do not Email" as "yes" as they are not likely to get converted.