Lead Case Study

Sudhir Singh

DS-C59

First Step Analyzing the Data

The company is taking the leads from the different social media platforms as well as customers who are coming to the platform

We are taking the leads and putting logistic regression to build the Machine Learning Model. Some of the important factors that we need to take care of the data quality

We need to remove all the data from the Table that is not changing throughout the table, remove all the null values, remove some of the categorical columns which will not be affecting the result

Checking for the Outliners

- **Now you can see that you have all variables as numeric.**
- Now you can see that you have all variables as numeric.
- **Check for Outliers and treat then**

Splitting the test and train data

• We have split the test and train data in a 30:70 ratio

Feature Scaling

We have done the feature scaling of the data

Checked the correlation between them, and removed all the highly correlated data from the table.

Used the regression Model

- Used the regression Model to build the Model and train it
- During the Model Building, we used the RFE for feature selection
- Test the accuracy
- Checked the VIF and then removed all the attributes which are having high VIF
- Checked the Accuracy Again

Test the final Model on the test data

- We have finally tested the data for the test data

- We have checked the accuracy

And amazed to check the results that accuracy has not changed in the final test as well.