

Lead Scoring Case Study

A brief summary report that how we approaches to the assignment

Understanding of data

- Inspection

Data Cleaning

- The first step is to clean the data and remove unwanted variables.
- After removing we found that columns are having 'Select' as their labels and we need to replace those with the null values.
- Removed columns having more than 35% of null values.
- For remaining missing values, we have replaced with maximum number of occurrences of the column.

Data Visualization

Visualising Numerical Variables and Outlier Treatment

Visualising Categorical Variables

Data Transformation:

- Changed the multicategory labels into binary variables in the form of '0' and '1'.
- Created dummy variables for some variables.
- Checked the outliers and removed some of the numbers using 0.99 - 0.1% analysis.

Data Preparation:

- Splitting the dataset into train and test dataset.
- Scaled the dataset using the StandardScaler().

Model Building

It didn't work due to this below error

TypeError: `__init__()` takes 2 positional arguments but 3 were given