Summary report:

Initially started with understanding the given data set and data preparation. Removed columns having more than 40% of null values and records having lesser null values. With the help of mapping function and label encoder, converted all categorical variables to numerical format. After this, dataset got splitted into train and test sets. On train data set, scaling is applied where values very large as compared to other variables. Plotted correlation matrix to check multicollinearity, and then removed such variables from train and test data set. After all these steps, started with model building. Initially, feature selection was performed using RFE (Recursive Feature Elimination) to reduce efforts, then based on VIF values remaining variables were removed. We computed some metrics to determine the model stability. At the end, the same model was applied to test data set. As mentioned in the problem statement, lead score (0-100) was assigned to each lead based on probabilities, which then depicts hot leads and cold leads.