Lead Scoring Case Study

Problem Statement An X Education need help to select the most promising leads, ie.,the leads that are most likely to convert into paying customers. The company requires us to build a model wherein you need to assign a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance. The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.

Goals and Objectives There are quite a few goals for this case study.

Build a logistic regression model to assign a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads. A higher score would mean that the lead is hot, i.e. is most likely to convert whereas a lower score would mean that the lead is cold and will mostly not get converted. There are some more problems presented by the company which your model should be able to adjust to if the company's requirement changes in the future so you will need to handle these as well. These problems are provided in a separate doc file. Please fill it based on the logistic regression model you got in the first step. Also, make sure you include this in your final PPT where you'll make recommendations.

These are the steps that will be followed:

1. Data Preparation

1.1 Data Loading 1.2 Data Inspection 1.3 Data Imbalance 1.4 Data Cleaning 1.5 Datatype Conversion

2: EDA and Data Viz 2.1 Univariate Analysis of Boolean Variables 2.2 Univariate Analysis of Categorical Variables 2.3 Univariate Analysis of Numerical Variables 2.4 HeatMap Multi-Colinearity Check

3: Model Building 3.1 Dummy Variables 3.2 Train-Test Split and LR Model 3.3 Scaling Numeric Features 3.4 Feature Reduction Using RFE and Manual Methods 3.5 Fit Training Set and Make Predictions

4: Model Validation 4.1 Making Predictions on Test Dataset 4.2 Final Evaluation Metrics for The Test Dataset

5: Measuring The Precision Recall Trade-off 6: Assigning The Lead Score 7: Determining Feature Importance 8: Final Observations and Recommendations

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js