Summary for Lead Score Case Study

X Education gets lot of leads, it has conversion rate of around 30%. Requirement is to build a model where a lead score is to be assigned to each of the leads such that the customer with a higher lead score have higher chance of conversion. CEO's target for lead conversion rate is around 80%.

The steps involved while making a model for analysis are:

Data Cleaning:

1. Columns with more than 3000 values as null were dropped.
2. Dropping not so important columns
3. Dropping columns with high 'Select' values

EDA:

1. Performed univariate and bivariate analysis for categorical and numerical variables.

Data Preparation and Modelling:

1. Creating dummy variables
2. Splitting the train – test data in ratio of 70:30.
3. Feature Scaling
4. Checking the correlation.

Model Building:

1. Used RFE to reduce variables from 48 to 15 variables.
2. Manual Feature Reduction process was used to build models by dropping variables with p value > 0.05
3. All the variables have VIF < 5 and p-values are low.

Model Evaluation:

1. **Train Dataset**
   A. Cut-off value is 0.42
   B. Accuracy: 78.85%
   C. Sensitivity: 78.86%
   D. Specificity: 78.84%
   E. Precision: 80.40%

F. Recall: 73.58%

2. **Test Dataset**
   A. Cut-off value is 0.42
   B. Accuracy: 78.83%
   C. Sensitivity: 78.63%
   D. Specificity: 79.02%
   E. Precision: 78.31%
   F. Recall: 78.64%

Recommendation:

Feature responsible for good conversion rate are:

4. The total time spend on the Website
5. Total number of visits
6. Lead Origin_Lead Add Form
7. Current occupation_Working Professional
8. Last Notable Activity_Had a Phone Conversation