

Summary

The X Education provides online courses for industry professionals. The company get leads through various sources but the conversion rate is around 30%. To raise the conversion rate to 80% as per the CEO's inclination we have done effective analysis with the variables which are the means to converting the leads.

The following steps have been taken:

- Reading and understanding the data with the help of data dictionary.
- Importing the required libraries and inspecting the data.
- Replacing the 'Select' with Nan for not effecting the analysis.
- Dropped the columns with only one unique class and columns with index numbering which doesn't help for the analysis.
- Dropped the columns which have missing values more than 35% but managed to hold few columns though having more missing values as they were seem to be important for the analysis.
- The columns still having missing values were imputed by necessary apt values either mean, mode etc.
- Outlier treatment was done for necessary continuous variables and few variables were bucketed into 'others' to control the data imbalance.
- Done Univariate analysis and Bivariate analysis wherever required.
- Converted required variables into binary classification and created dummy variables for categorical data.
- Divided the data into train set and test set by 70:30 ratio and scaled the data using Standard Scaler.
- Selected significant features using RFE technique.
- Built models using logistic regression method and predicted.
- With Confusion matrix got the Accuracy, Sensitivity, Specificity values and plotted required curves to get the optimal probability cut-off value.
- With Recall and Precision plotted another curve to get the probability cut-off.
- With the obtained values predicted on the test set and attained good performance model. And marked the lead score for the leads accordingly.

Results Obtained from the above model are:

On Train Set:

- Accuracy: 85.45%
- Recall: 82.54%
- Precision: 80.24%

On Test Set:

- Accuracy: 84.61%
- Recall: 80.58%
- Precision: 77.83%

- **The X education can target on the following to convert the leads into hot leads**
 Leads sourced from Welingak Website.
 Leads originated from Lead Add Form.
 Leads whose current occupation is Working Professional.
 Leads sourced from Olark Chat.
 Whose last activity is SMS sent.
 Leads who spend most time on the website.
- **The X education can focus and improvise on the following things to strengthen the lead quality**
 Leads marked by the assigned employee as Worst Quality.
 Leads marked by the assigned employee as Not Sure.
 Leads marked by the assigned employee as Might be.
 Customers opted not to be emailed about the course (Do Not Email).
 Leads marked by the assigned employee as Low in Relevance.
 If the Last Notable Activity by the student is Modified.
 If the Last activity performed by the customer is Olark Chat Conversation.

The X Education can increase the lead conversion rate by focusing the above factors and improving the negative impacting factors towards the positive approach. The company can increase the target lead conversion rate to be around 80% by using the model with threshold of 0.4 as per the CEO's inclination.