



Lead Score Assignment

Rajat Pandey



Business Case

Who: The X Education Online Course edutech company

What: They provide higher education and technical courses to the experienced professionals.

Issue: The company does a lot of email marketing, cold-calls, SEO optimization and does get a lot of leads and opportunities, however, the lead conversion rate is not upto mark and the company wants to look into the options and recommendations to find out the root cause and upgrade and enhance their approach for the same.

Current State: The current lead conversion rate is 30%.



Solution and Actions

We have performed the data analysis on the details provided by X company and created a model for the same to provide a detailed analysis of the trend and major features that are affecting the leads.

- 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.



Model Development & Evaluation

- Model Building:

Firstly, RFE was done to attain the top 15 relevant variables. Later the rest of the variables were removed manually depending on the VIF values and p-value (The variables with $VIF < 5$ and $p\text{-value} < 0.05$ were kept).

- Model Evaluation:

A confusion matrix was made. Later on the optimum cut off value (using ROC curve) was used to find the accuracy, sensitivity and specificity which came to be around 80% each.



Prediction and Recall

Prediction:

Prediction was done on the test data frame and with an optimum cut off as 0.35 with accuracy, sensitivity and specificity of 80%.

Precision – Recall:

This method was also used to recheck and a cut off of 0.41 was found with Precision around 73% and recall around 75% on the test data frame.



Observations and Findings

It was found that the variables that mattered the most in the potential buyers are (In descending order):

1. The total time spent on the Website.

2. Total number of visits.

3. When the lead source was:

- a. Google
- b. Direct traffic
- c. Organic search
- d. Welingak website

4. When the last activity was:

- a. SMS
- b. Olark chat conversation

5. When the lead origin is Lead add format.

6. When their current occupation is as a working professional.

Keeping these in mind the X Education can flourish as they have a very high chance to get almost all the potential buyers to change their mind and buy their courses.

Few Visualizations

