

Lead Scoring Case Study

Initially, there are a lot of leads generated but only a few of them convert as paying customers. The conversion rate for a particular company depends on various factors which can be used as well as derived from the data we have at hand. Understanding the data and making it semantically clean has been the first step towards our goal. Further the process of exploratory data analysis helps us to nurture the potential leads well. In order to get a higher lead conversion, first we need to remove all the noise from our data and sort out the best prospects from the leads that have been generated.

'TotalVisits', 'Total Time Spent on Website', 'Page Views Per Visit' are features which contribute most towards the probability of a lead getting converted. Before we move towards building a model to uplift the current business scenario, we made sure to clean the data, verify the data types, handle missing values, detect outliers and perform univariate and bivariate analysis.

While building the model, we were sure to use those features that we selected using Recursive feature elimination on the train set. Checking the variables p-value and multicollinearity index has been one of the crucial aspects for feature selection. After this, the regression model is built and studied to make observations. Accuracy, sensitivity and specificity graph is used as the base to select our target variable cut off. Once the target variable is set, metrics like ROC, confusion matrix, precision, recall have been used to test our training set. The accuracy achieved by our model comes out to be eighty-four percent. Based on the overall study by the algorithm, we perform a similar transformation on our test set and get our top scoring leads. Using the regression model, we got around a one to two percent difference between train and test data's performance metrics. This means, we did not overfit our training data and it is performing well. Keeping in mind the approach and end goal, we can make great use of the data transformed by the model and hence provide a well defined and informed solution to the business.

We must keep this list of leads handy so that they can be informed about new courses, services, job offers and future higher studies with a maximum probability of conversion. Building our model shows us the accuracy and percentage of leads that have the maximum possibility of conversion. This helps us in monitoring and tailoring the information sent to them on a daily basis. By carefully providing job offerings, information or courses that suit best according to the interest of the leads, the company can yield a greater deal of profit. A proper plan will be needed to chart the requirements of each lead in order to capture them as prospects. Focusing on converted leads, holding question-answer sessions, being in constant contact with leads might get us the right information about them. Make further inquiries and appointments with the leads to determine their intention and mentality to join online courses.

Reported By - Jital Engineer, Mukul Yadav, Siddhi Kadam