Lead Scoring Case study

X Education sells online courses to industry professionals. In recent times the conversion rate for the company was around 30% indicating of 100 queries about the course, around 30% get enrolled with the course. The objective for the company was to increase the conversion rate for which it is necessary to identify the influencing factors that help to convert the potential lead to leads. A data of around 9,420 observations was used to analyze and identify the factors. Overall, there were around 37 parameters captured during every query. From the data it seems that unemployed people take the courses more often compared to working professionals. However, working professionals had a higher conversion rate compared to unemployed people. Further, compared to visiting on website, people searching on the search engines such as Google were more frequent.

Logistic regression was used to address the objective. After data was cleaned: removing variables that does not have lot of variations, lot of missing values (> 30%), removing observations with missing values, and handling outliers; it was processed. The correlation matrix indicates that the conversion depends on last notable activity, total time spent, lead source, and the current occupation. Besides, it also indicated a high correlation across dummy variables, which was identified and corrected using variance inflation factor (VIF). The final model had an accuracy of around 79% which was tested with receiver operating characteristic (ROC) curve having area under the curve of around 86% and precision of around 78%.

The best model indicates that *lead source*, *total time spent on website*, and *current occupation* were the most influencing variables which was based on coefficients for the scaled variables. To increase the conversion rate, it can be suggested to connect people spending more time (738 units of time) on website and are a working professional. Besides, the person that has been referred shall be prioritized and a good cashback or bonus should be introduced to the people introducing the new leads. In addition, an advertisement campaigns can be run at the Welingak website, to grab more leads as a source.

Further depending on the availability of interns for making calls, false positive rates can be increased or decreased. False positives can be increased where abundant interns are available, so as more calls can be made while, it can be decreased when the resources are limited with the company. Finally, it can be concluded, addressing the factors highlighted would support in increasing the conversion rate from 30% to a higher limit.