Summary Report on Predictive Analysis for X-education   Vivek Sharma

<u>Executive Summary</u>

The provided data had 9240 entries across 37 different variables that was collected by executives of X-education company. The pursuit was to find out predictive power of this data and come out with conclusions that can increase the leads for X-education company, as they were struggling with very low lead conversion.

The task was taken up with a mandate to give meaningful outcomes that can enable X-education to work on any improved marketing efforts or increased resource allocation in certain areas which can churn out good lead conversion and better business as a whole. To give meaningful data driven outcomes we have performed the cleaning of data which removed lot of variables with more than 30% missing or null values. Then wherever the null values were less than 30% they were adjusted by imputation with median value or sometimes replacing with the frequently occurring value.

After cleaning, the exploratory data analysis was done to understand the cause and effect relationship of different independent variables on the dependent variable. It gave some interesting results which clearly indicates that X-education should continue to focus on Lead origin (particularly the landing page submissions), Google and direct website traffic as maximum of potential leads coming through these bucket got converted. As there is conversion happening for people who opt for having emails, a special focus should be given to floating very precise and meaningful emails to specific potential leads who can use these emails to clarify their doubts and can get more interested in the courses offered by X-education. Similar concept applies to making specific calls as well. Furthermore, the people who are currently unemployed are more interested in courses from X-education, so focus on this bucket should continue along with some new courses which can be introduced for working professionals as there is a traction in that bucket as well, however it needs to be properly nurtured. Recommendations are working big time for lead conversion, hence company should continue to focus on learner's experience, so that the word-of-mouth effect continues and the business conversion happening through recommendations gets more strengthened. Last but not the least, another insight that is easily visible from this analysis is that the Total visits and the Page views per visit are almost directly proportional to the lead conversion, hence some business conclusions can be drawn from this parameter as well.

Prediction analysis was further done by making Logistics Regression models using the provided data and the 5th Logistics regression model showed that prediction of conversions (using the parameters in the model, coinciding with few insights drawn from exploratory analysis) gives us 80% (79.8%) conversion of leads which was the expectation of CEO, X-education. In other words, if the parameters included in the 5th model are taken care and business decisions are made to strengthen these attributes, the lead conversion can reach up to 80% or more. This 5th model is having a predictive power with accuracy of 81%, precision of 75%, specificity of 83% and sensitivity of 79%.

This summary is meant for business conclusions