Enhancing Lead Conversion for X Education through Logistic Regression

X Education, an online course provider for professionals, faces a significant challenge: only 30 out of every 100 leads convert into paying customers. Despite generating a high volume of leads, the current conversion rate of 30% falls short of the desired 80%. To improve efficiency, the company must identify high-potential leads—referred to as 'Hot Leads'—so the sales team can concentrate their efforts on these prospects instead of targeting all leads indiscriminately. A logistic regression model provides valuable insights to help prioritize leads and optimize the sales process.

A historical dataset with around 9,000 leads has been provided, containing various features such as Lead Source, Total Time Spent on Website, Total Visits, and Last Activity. Some of these attributes may influence lead conversion, while others might have little to no impact. These attributes vary in importance; some strongly influence conversion, while others provide little value. This model analyses past lead data to help prioritize which leads are most likely to convert into paying customers. The regression model gives each lead a score based on their likelihood of converting. Leads with higher scores should be prioritized to maximize the sales team's efficiency and reduce time spent on unlikely conversions. After thorough Data Cleaning and conducting EDA, we have done data preprocessing and built a logistic model. The model's performance is evaluated with an accuracy of approximately 80.7%. We also examined metrics such as Sensitivity (approximately 80.2%), Specificity (around 81%), and the ROC curve. Based on these findings, we determined the optimal cutoff probability to be 0.344.

Based on the regression coefficients, we identified the key behavioural metrics as Lead Source, Total Time Spent on Website, Total Visits, Last Activity, Lead Origin, and Form Submission Time. By concentrating on these metrics, prioritizing high-potential leads, and personalizing the outreach strategy, X Education can enhance its sales efficiency. If executed effectively, this approach could bring the conversion rate closer to the desired 80%, ensuring the sales team focuses on what matters most.

With improved targeting, streamlined processes, and actionable insights, X Education is well-equipped to convert its Hot Leads into loyal customers. The logistic regression model provides valuable insights from past lead data, allowing the assignment of lead scores to potential customers. This scoring helps the sales team prioritize leads with the highest likelihood of conversion, ideally those with a predicted conversion probability above 70%, while deprioritizing colder leads. This targeted approach will increase efficiency, as the sales team can focus solely on promising prospects, reducing time wasted on unlikely conversions. By maintaining a clear focus on high-scoring leads and employing personalized outreach strategies, X Education can maximize its sales efficiency and work towards achieving its target conversion rate of 80%.