SUMMERY of Lead Scoring Case Study

by Mayukh, Cyril & Prasoon

The problem statement revolves around improving the lead conversion rate for an online education company called X Education. Despite generating a substantial number of leads, their conversion rate is low at around 30%. The company aims to boost this conversion rate to approximately 80% by identifying and focusing on the most promising leads, often referred to as 'Hot Leads.' They seek a data-driven solution to assign lead scores to each prospect, allowing the sales team to prioritize their efforts on the leads with higher conversion potential.

To address this problem, the following key tasks need to be accomplished:

- **Data Collection and Preprocessing :** Gather data on leads' interactions and characteristics. This includes website browsing behavior, form submissions, and past referrals. Clean and prepare the data for analysis.

- **Feature Engineering :** Create relevant features that can help predict lead conversion. This may include

lead source, activity on the website, the time spent on certain pages, and demographic information.

- **Model Building :** Develop a predictive model, such as logistic regression, that assigns lead scores to each lead based on their likelihood of conversion. The model should be trained on historical data with known outcomes (converted or not converted).

- **Model Evaluation :** Assess the model's performance using appropriate evaluation metrics, such as accuracy, precision, recall, and F1-score. The goal is to create a model that accurately identifies potential customers.

- **Lead Scoring :** Apply the trained model to assign lead scores to all current leads. The higher the score, the more likely the lead is to convert.

- **Business Recommendations :** Present the results and recommendations to X Education's management. This includes visualizations and a clear explanation of how the model can help improve lead conversion. Highlight the most important features and insights from the analysis.

- **Documentation** : Provide a well-commented Jupyter notebook with the model implementation and evaluation, a Word document with solutions to all problems, and a presentation summarizing the analysis approach and results.

In summary, the problem statement revolves around increasing lead conversion rates for X Education by identifying the most promising leads using a predictive model. The solution will involve data analysis, feature engineering, model building, and business recommendations to help the company achieve its target conversion rate of 80%. The final deliverables will include a comprehensive analysis report, a Jupyter notebook with the model, and a presentation summarizing the findings and recommendations.

THANK YOU