Lead Scoring Case study - Executive Summary Report

By Ashwin Hariharan

Introduction

X Education, an online education company catering to industry professionals, has been experiencing a low lead conversion rate. In an effort to enhance the lead conversion rate, the objective of this project was to develop a machine learning model that could identify high-potential leads, also known as "Hot Leads." This would enable the sales team to focus on communicating with these hot leads rather than contacting everyone indiscriminately, thereby increasing efficiency.

The process

To achieve this goal, we undertook a comprehensive analysis of the available data and built a logistic regression model to assign a lead score between 0 and 100 to each lead. The higher the score, the more likely the lead is to convert, and vice versa. Our model achieved an accuracy score of **80%**, with a sensitivity of **81.3%** and a specificity of **79%**.

During the data preprocessing stage, we dealt with missing values by dropping columns with more than 40% null values, removing columns with a single value or highly skewed data, and eliminating rows with very few null values. We then employed **one-hot encoding** for categorical columns and applied feature scaling on numerical columns using a standard scaler. **RFE** was utilized to reduce the number of initial features to 15. Subsequently, we conducted manual feature selection based on **VIF** scores and p-values, arriving at a logistic regression model after four iterations.

Model evaluation required metrics like sensitivity and specificity due to the skewed nature of the data. We used the ROC curve to determine an optimal balance between these metrics. Ultimately, we calculated the lead scores using the conversion rate probability predictions, allowing us to rank leads in order of "hottest leads" first.

Learnings

Our analysis revealed that the following variables were strong indicators of lead conversion success or failure:

- Email preference
- Total time spent on the website
- Leads originating from landing page form submissions
- Leads who have had conversations via phone
- Leads that are working professionals

Based on these findings, we recommend that X Education invest their resources in the following strategies to improve lead conversion rates:

1. **Enhance website engagement**: Develop strategies for engaging leads on the website, such as personalized content and user-friendly navigation. This will encourage leads to spend more time on the site, increasing the likelihood of conversion.
2. **Improve phone and chat communication**: Train staff to better engage leads over phone and chat conversations, ensuring prompt responses and tailored solutions. This will help create a strong rapport with potential customers.
3. **Target working professionals**: Sales efforts should prioritize working professionals, as they have been identified as a key demographic with higher conversion rates.
4. **Optimize email campaigns**: Invest in long-term strategies to make email campaigns more effective. This includes segmenting email lists and personalizing email content.

By implementing these recommendations, X Education can expect to see a significant improvement in lead conversion rates, allowing for more efficient use of sales and marketing resources and contributing to overall business growth.