Lead Scoring Case Study

Krishna Mohan Sripada

Senior Data Engineer

11 years of Experience

EDA

- Exploratory Data Analysis (EDA) and Data Preprocessing
1. Removed columns with more than 30% null values
2. Removed columns having more No's than Yes's as there is no balance in the data
3. Removed Columns with more Select rather than actual values
4. Analyzed the Outliers and capped them at 95%
5. Dummified the variables accordingly

Model Building and Evaluation

1. Built the logistic regression model

2. Plotted the ROC curve

3. Model accuracy is 82%

4. Optimal cut-off point is around 0.35

Key Observations

1. Leads with noticeable activity in phone conversations or email have higher chances of converting

2. Leads from certain sources like Olark Chat, Referral Sites, and Organic Search tend to have higher conversion rates

3. Leads who have filled out the form on the website and engaged with multiple activities are more likely to convert

4. Leads from certain countries like India and specific cities like Mumbai have higher conversion probabilities

Recommendations

1. Focus sales efforts on leads with high scores based on engagement metrics and lead source

2. Implement targeted marketing strategies for specific demographics that have higher conversion rates

3. Enhance communication efforts, especially phone calls and personalized emails, to engage leads effectively

4. Continuously monitor and adjust strategies based on real-time feedback and performance metrics