

Case Study Summary: Lead Scoring for X Education

“The goal of this analysis was to develop a predictive lead scoring model for X Education to enhance lead conversion rates by identifying high-priority leads and optimizing the sales process.”

Steps Followed

1. Data Cleaning:

- The dataset contained missing values, with features like Total Visits and Page Views per Visit were dropped and the outliers were capped at 99 percentile.
- Columns such as Tags and Lead Quality with over 35% missing data were dropped.
- Dummy variables were created for categorical features.

2. Exploratory Data Analysis (EDA):

- Patterns revealed that **68.5% of leads** spending over 3 minutes on the website converted.
- Categories like Lead Source: Direct Traffic, Google and Last Activity: Email Opened exhibited strong conversion tendencies.
- Leads with occupation as Working progression also showed good conversion rate.

3. Model Development:

- A logistic regression model was built to predict the likelihood of lead conversion.
- Recursive Feature Elimination (RFE) was used for feature selection identifying the top 15 variables.
- Variables with VIF < 5 and p-value < 0.05 were retained to ensure multicollinearity was addressed.

4. Model Evaluation:

Performance metrics included:

- **Accuracy:** 79%
- **Precision:** 76%
- **Recall:** 80%

An optimum cutoff probability of 0.4 was determined using an ROC curve and PR for balanced performance.

Key Insights

1. Top Variables:

- Total Time Spent on Website: Is a significant predictor (Coefficient:1.165).
- Occupation: Working professional: A strong indicator of lead intent (Coefficient:2.57).
- Last Activity: Had a phone conversation: Highlights the importance of phone engagement (Coefficient: 2.12).

2. Categorical Variables to Focus On:

Lead Source: Google, Direct Traffic, Organic Search

Strategies for Implementation

1. Aggressive Conversion Phase:

- Assign high-probability leads (e.g., >0.8) to interns during peak periods.
- Focus on personalized outreach to maximize engagement.

2. Low-Priority Phase:

- Reduce phone calls and engage medium-probability leads through email campaigns.
- Limit calls to leads with probabilities above 0.75 for efficiency.

“The analysis demonstrated that prioritizing high-conversion leads and segmenting by behavior metrics, such as website engagement and email, phone interaction, can enhance conversion rates by over 20%. By applying these insights.