

SUMMARY

Lead Conversion Analysis for X Education

Objective: The objective of this analysis is to understand which factors influence the probability of a lead getting converted into a paying customer for X Education. The goal is to focus on optimizing lead conversion rates based on historical data from website interactions and lead characteristics.

Steps Taken: Data Cleaning, EDA, Dummy Variable Creation, Train-Test Split, Feature Selection and Model Building, Model Evaluation and Prediction and Fine Tuning.

Key Things noted and executed:

- One Hot Encoding
- The numeric variables were scaled using the MinMaxScaler to standardize them.
- Recursive Feature Elimination (RFE) was used to select the top 15 most important variables.
- Variables with Variance Inflation Factor (VIF) greater than 5 were removed to reduce multicollinearity.
- Only variables with p-values less than 0.05 were kept in the final model.
- A logistic regression model was built using these selected features.
- The confusion matrix indicated strong performance, with balanced results for accuracy, sensitivity, and specificity, each at around 80%.
- The Receiver Operating Characteristic (ROC) curve was used to find an optimal probability cutoff, set at 0.35 for the model.
- Model metrics included:
 - **Accuracy:** 82.03%
 - **Precision:** 80.38%
 - **Recall:** 69.95%
 - **F1 Score:** 74.80%
 - **ROC-AUC Score:** 89.46%
- The model was applied to the test dataset with the optimum cutoff of 0.35, yielding similar results (accuracy, sensitivity, and specificity around 80%).
- Precision-recall analysis indicated a cut-off of 0.41 would lead to slightly higher precision (73%) and recall (75%) on the test set, balancing false positives and false negatives.

Key Factors Affecting Lead Conversion:

The following are the most influential factors contributing to lead conversion, ranked in order of importance:

1. Total Time Spent on Website:

- The longer a lead spends on the website, the higher the likelihood of conversion. This indicates high engagement with content and courses offered.

2. **Page Views Per Visit:**

- Leads with more page views per visit are likely browsing more courses and information, increasing conversion probability.

3. **Lead Origin:**

- Leads coming through the **Lead Add Form** show a strong intent to convert.

4. **Lead Source:**

- Specific lead sources contribute significantly to conversion probability:
- **Welingak Website:** Leads from this source are highly likely to convert.
- **Google Search:** Organic search results drive significant conversions.
- **Direct Traffic:** Returning visitors or those with direct URLs convert well.

5. **Email Communication:**

- Leads marked as "Do Not Email" show a negative impact on conversion, emphasizing the importance of email communication in conversion strategies.

6. **Last Activity:**

- Leads who received an **SMS** or had an **Olark chat conversation** as their last activity show higher conversion rates.
- Leads with **Phone Conversations** as their last activity also show a positive influence on conversion.

7. **Current Occupation:**

- Leads who are working professionals are more likely to convert, indicating interest from those seeking career advancement through education.

Recommendations for X Education:

- **Increase Time Spent on Website:** Focus on content that increases visitor engagement, such as course previews, testimonials, and interactive tools.
- **Optimize Lead Sources:** Continue focusing on high-performing lead sources like Google and Welingak Website, and invest in direct traffic campaigns.
- **Leverage Communication Channels:** Focus on SMS and chat as effective follow-up methods. Be cautious with leads marked as "Do Not Email" as they are less likely to convert.
- **Target Working Professionals:** Marketing strategies should focus on working professionals seeking to upskill, as they have the highest conversion rates.

By focusing on these key factors, X Education can improve its lead conversion rate, making better use of marketing efforts and outreach to potential customers.