

Lead Score Case Study

By:

Kundan Yadav

Business Objective

- X Education sells online courses to industry professionals.
- Despite acquiring numerous leads, their conversion rate is low.
- Objective: Identify 'Hot Leads' to improve conversion rates.
- Goal: Build and deploy a model to identify these promising leads.

Solution Methodology

1. Data Cleaning & Manipulation:

- Handle duplicates and missing values.
- Drop unnecessary columns.
- Impute values where necessary.
- Handle outliers.

2. Exploratory Data Analysis (EDA):

- Univariate analysis: value counts, distributions.
- Bivariate analysis: correlations and patterns.

3. Feature Engineering:

- Scaling and encoding variables.
- Creating dummy variables.

4. Model Building:

- Use logistic regression for classification.
- Validate the model.
- Present conclusions and recommendations.

Data Manipulation

- - Total Rows: 9240, Total Columns: 37.
- - Removed single-value features and non-essential columns.
- - Dropped columns with over 35% missing values.
- - Handled outliers and normalized numerical variables.

Exploratory Data Analysis (EDA)

- Conducted univariate and bivariate analyses.
- Identified key relationships and patterns in the data.

Data Conversion

- Normalized numerical variables.
- Created dummy variables for categorical data.
- Analysis-ready dataset: 8792 rows, 43 columns.

Model Building

- Split data into training (70%) and testing (30%) sets.
- Used Recursive Feature Elimination (RFE) for feature selection.
- Built the model by iteratively removing variables with high p-values and VIF values.
- Achieved an overall accuracy of 81% on the test set.

ROC Curve

- Identified the optimal cut-off point for balanced sensitivity and specificity.
- Optimal cut-off at 0.35.

Conclusion

Key factors for potential buyers (in descending order):

- 1. Total time spent on the website.
- 2. Total number of visits.
- 3. Lead source: Google, direct traffic, organic search, Welingak website.
- 4. Last activity: SMS, Olark chat conversation.
- 5. Lead origin: Lead add format.
- 6. Current occupation: Working professional.

These insights can help X Education target potential buyers more effectively and increase conversion rates.