

Lead Scoring Case Study: Enhancing Conversion Rates with Predictive Analytics.



Problem Statement

- The objective is to optimize lead conversion rates by identifying key factors influencing conversion and predicting the likelihood of success for each lead.
- Challenges addressed: handling missing data, categorical encoding, and model interpretation.



Objectives

- Develop a logistic regression model for lead scoring.
- Understand key variables influencing lead conversion.
- Provide actionable insights for targeted marketing.

Data Understanding and Preparation

- **Dataset:** Leads.csv, initial shape, and exploration.
- **Cleaning steps:** Removed columns with >3000 missing values, handled Select entries as missing, and created dummy variables for categorical features.



EDA

- Top correlated features influencing conversion rates (e.g., Total Time Spent on Website and Lead Score).
- Visuals: Heatmap of correlations, boxplots for numerical variables, and trends segmented by target class.

Model Approach

- Used Logistic Regression as the primary model.
- Feature selection: Recursive Feature Elimination (RFE).
- Data preprocessing: Scaling with **MinMaxScaler** and splitting into training/testing sets.



Model Evaluation

- Accuracy, Precision, Recall.
- Confusion matrix for classification results.
- Insights into model performance and adjustments made.



Key Variables and Business Insights

- **Top Predictors of Conversion:** Features like Total Time Spent on Website, Last Activity, and Lead Origin showed strong influence on conversion rates.
- **Behavioral Insights:** Leads who spent more time engaging with the website had higher conversion probabilities.
- **Categorical Impact:** Certain categories within Lead Source and Last Notable Activity significantly differentiated converted from non-converted leads.
- **Actionable Insight:** Targeting and nurturing leads with high website engagement and specific activity patterns can enhance conversion efficiency.



Strategies for Aggressive and Low-Priority Periods

- Aggressive Periods: Focus on leads with the highest predicted conversion probabilities by tailoring marketing efforts and offering personalized incentives.
- Low-Priority Periods: Engage lower-probability leads through automated campaigns, nurturing them with educational content to improve future conversion chances.
- Resource Allocation: Use lead scores to optimize sales team efforts, ensuring high-value leads receive priority during busy periods while maintaining touchpoints with lower-priority leads.



Recommendations and Action Plan

- Implement the logistic regression model in the CRM system for real-time lead scoring and dynamic segmentation to target high-potential leads effectively.
- Focus on improving engagement metrics like Total Time Spent on Website to enhance user interaction and boost conversion rates.
- Allocate sales efforts based on predicted conversion probabilities, prioritizing high-scoring leads for immediate follow-up.
- Regularly monitor model performance and re-train it with updated data to maintain accuracy and adapt to market changes.



Conclusion

- The logistic regression model successfully identified key drivers of lead conversion, enabling targeted and efficient marketing strategies.
- The case study highlights the importance of data preparation and feature selection in building scalable, actionable predictive solutions.



Thank you!