Lead Scoring Case Study

Harshvardhan Singh Chauhan

December 11, 2023

Abstract

This report encapsulates the comprehensive findings and insights derived from a meticulous Lead Scoring case study that employed a logistic regression model. The primary objective of this study was to meticulously address the specific business requirements pertaining to lead conversion, shedding light on critical factors that contribute to the transformation of leads into paying customers.

1. Introduction

The initial stages of the customer acquisition process witness an influx of leads; however, the conversion into paying customers remains a challenging feat. The study, therefore, focused on delving deep into the nuances of lead conversion dynamics. The dataset under scrutiny predominantly features leads originating from India, with a keen emphasis on analyzing the city-wise distribution, highlighting Mumbai as the epicenter of lead concentration.

2. Methodology

The study hinged on the strategic deployment of a logistic regression model, a powerful tool in predictive modeling. The model underwent rigorous training on a dataset brimming with diverse lead-related information, encompassing demographic details, platform activities, and

interactions with various features. This methodological approach ensured a robust analysis, capable of providing nuanced insights into the lead conversion landscape.

3. Code

The implementation of the logistic regression model and the subsequent analysis is encapsulated in the github link mentioned below. The code meticulously handles the intricacies of data preprocessing, model training, and evaluation, ensuring transparency and reproducibility in the analysis.

https://github.com/HS-Chauhan/Lead-Scoring-Case-Study

4. Key Findings

4.1 Geographical Analysis

The geographical distribution of leads revealed a preponderance of leads originating from India, with Mumbai emerging as the city with the highest lead count. This insight aids in customizing marketing strategies to align with regional preferences and trends.

4.2 Incomplete Data Handling

Certain columns exhibited a peculiar level termed 'Select,' denoting that the lead failed to make a selection. To extract meaningful insights, mandatory selections were enforced for crucial columns such as Customer Occupation, Specialization, etc. This process significantly enhanced the quality of the dataset.

4.3 Activity Metrics

An intriguing correlation emerged between higher total visits and increased time spent on the platform, signifying a positive influence on lead conversion chances. This revelation emphasizes the importance of user engagement metrics in predicting conversion outcomes.

4.4 Specialization Insights

Leads enrolling in courses for Better Career Prospects, especially those specializing in Finance Management, demonstrated a heightened probability of conversion. Additionally, specializations in HR, Finance, and Marketing Management surfaced as strong indicators of conversion potential.

4.5 Customer Engagement

Elevating customer engagement, particularly through email and calls, emerged as a critical factor in driving lead conversion. Leads exhibiting a propensity to open emails demonstrated a significantly higher likelihood of conversion. Additionally, leveraging SMS communication was identified as an untapped avenue for potential benefits.

4.6 Occupation Analysis

A substantial portion of leads currently identifies as unemployed, highlighting an opportune segment for targeted efforts. Focusing marketing and engagement strategies on this demographic could yield tangible improvements in overall lead conversion rates.

5. Conclusion

In summation, the logistic regression model has proven instrumental in unraveling the intricacies of lead scoring, offering actionable insights for strategic decision-making. The amalgamation of geographical, activity, specialization, and engagement patterns equips the business with a tailored approach to target leads with the highest likelihood of conversion. The nuanced focus on unemployed leads and the implementation of proactive customer

engagement strategies stand out as key recommendations for enhancing overall conversion rates.

6. Submission Instructions

The complete codebase and dataset are appended herewith, ensuring transparency and facilitating further discussions on the study's methodologies and findings.