

LEAD SCORING ANALYSIS FOR X EDUCATION



- DRIVING CONVERSATION WITH DATA

• BY YASHASWINI KC (GCP C3)

CONTENTS

- Overview
- Problem Statement
- Data Exploration
- Model Building
- Recommendations
- Conclusion



OVERVIEW

- The goal is to assist X Education, in improving its lead conversion rate by building a predictive model to identify potential leads.
- The dataset provided contained various attributes related to leads, including lead source, total visits, last activity, and conversion status.
- Our task is to develop a logistic regression model to assign lead scores and optimize the lead conversion process.

PROBLEM STATEMENT

Problem Statement: Maximizing Lead Conversion

1. *Challenges Faced* - X Education encounters inefficiencies in lead conversion despite a steady influx of leads. With only a 30% conversion rate, there's a pressing need to enhance the process.
2. *Objective* - Our goal is to develop a predictive model that assigns lead scores, identifying hot leads with a higher likelihood of conversion. This targeted approach aims to boost the conversion rate to 80%.
3. *Dataset Overview* - We possess a dataset of 9000 leads, comprising various attributes such as lead source, website engagement, and past conversions. Through this dataset, we seek insights to refine our lead prioritization strategy.



PROBLEM STATEMENT

4. *Approach* - Leveraging logistic regression, we aim to assign lead scores between 0 and 100 to prioritize potential leads. Additionally, our model should adapt to future requirements, ensuring scalability and flexibility.
5. *Expected Outcomes* - By implementing an efficient lead scoring system, X Education anticipates increased conversion rates, optimized resource allocation, and enhanced customer acquisition strateg



DATA EXPLORATION

Understanding the Dataset: The dataset contains 9240 entries and 37 features, offering a comprehensive view of our leads' interactions and characteristics.

Peek into the Data: Through summary statistics and visualization, we'll explore key variables like lead sources and website engagement to understand our leads' behavior and preferences.

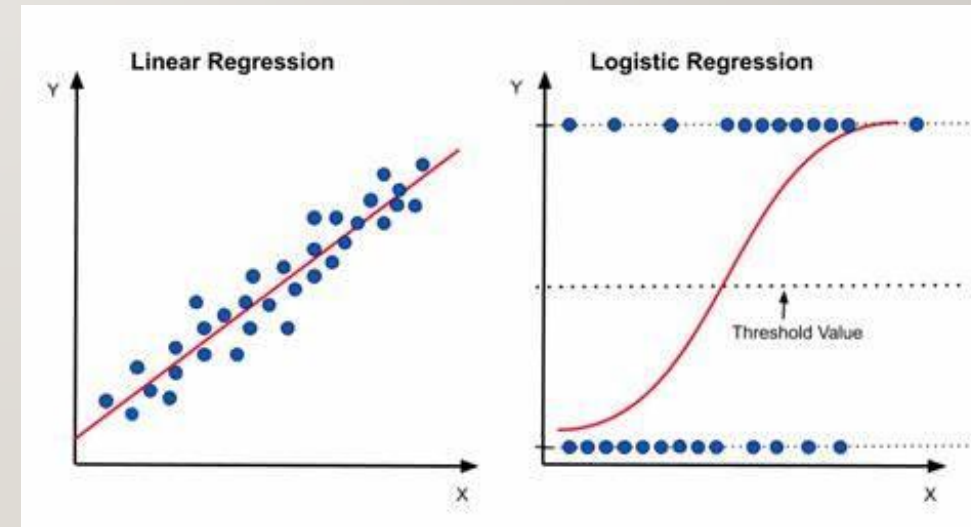
Visualization: Visual aids such as histograms and scatter plots will guide us, revealing patterns and correlations crucial for effective decision-making.

MODEL BUILDING

- Two models were trained and evaluated:
 1. Linear Regression and
 2. Logistic Regression.
- In the case of Linear Regression, the Ridge regression variant was utilized due to its ability to handle multicollinearity.
- The mean squared error (MSE) and R-squared were used as evaluation metrics, resulting in an **MSE of 0.0986 and an R-squared of 0.5896.**

MODEL BUILDING

- For Logistic Regression, the coefficients of the model were extracted and analysed to identify the top three variables contributing most towards lead conversion probability.
- The top three variables were identified as follows:
 1. Total Time Spent on Website
 2. Lead Number
 3. Tags indicating the lead will revert after reading the email



RECOMMENDATIONS

- Based on the analysis, we recommend the following:
 - a) Improve Lead Scoring
 - b) Use Automated Emails and Ads
 - c) Share Helpful Content
 - d) Be Active on Social Media
 - e) Analyse Data to Improve
 - f) Work with Other (like influencers)
- With these changes, X Education can keep sales going strong without making too many phone calls when they're ahead of schedule.

CONCLUSION

- In conclusion, the analysis demonstrates the importance of various factors in predicting lead conversion probability for X Education.
- By implementing the recommendations outlined in this report, X Education can enhance its marketing and sales strategies, ultimately leading to increased customer acquisition and revenue generation.

A close-up photograph of two hands shaking in a firm grip. The hands are positioned in the center of the frame, with fingers interlaced. The person on the left is wearing a dark blue suit jacket and a white shirt with a visible cuff. The person on the right is wearing a dark blue suit jacket. The background is blurred, showing out-of-focus lights in shades of blue and white, suggesting an indoor setting with ambient lighting. A thin white vertical line is positioned to the left of the text.

THANK YOU