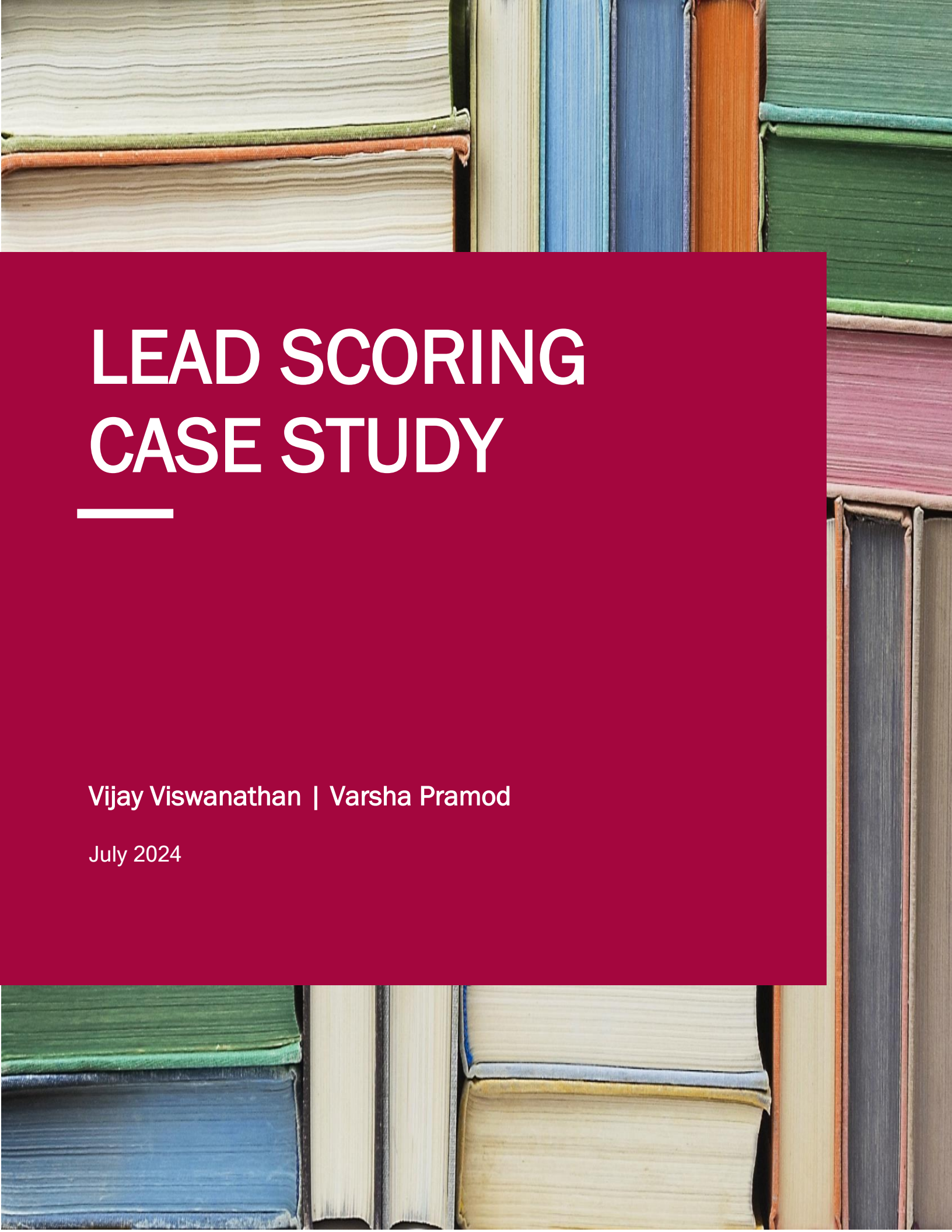

LEAD SCORING CASE STUDY

Vijay Viswanathan | Varsha Pramod

July 2024

Our Understanding

X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses. The company generates leads via

- Filling form on their website
- Referrals

Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not. The typical lead conversion rate at X education currently is around 30%. The CEO expects the target lead conversion rate to be around 80%.

To get a higher lead conversion, the company wishes to identify the most potential leads, also known as 'Hot Leads', so that the sales team can focus more on communicating with them rather than making calls to everyone.

Objective

X Education aims to optimize lead conversion using predictive models. They want to identify the most promising leads (referred to as "Hot Leads") to focus their efforts effectively.

Analysis Approach

1. Data understanding
   a. Import Necessary Libraries
   b. Load the data and Data dictionary.
2. Data Preprocessing and Preparation
   a. Handle missing values.
   b. Treat outliers.
   c. Create dummy variables for categorical features.
3. Exploratory data Analysis
   a. Conduct univariate, bivariate, and multivariate analyses.
   b. Analyze correlations.
4. Model building and Evaluation.
   a. Split data into training and test sets.
   b. Scale features.
   c. Select relevant features.
   d. Evaluate the model.

Conclusion

- The final model identifies potential leads with an accuracy of approximately 80%.
- Key contributing variables: “Total Visits,” “Total Time spent on the website,” and “Lead Origin Lead Add Form.”
- Focus on leads originating from Lead Add Form or Welingak Website and Total visits or total time spent on website for better conversion rates
- Adjust threshold during intern periods for aggressive outreach.
- Increase threshold to minimize unnecessary calls.

Learning

- In this case study, we applied logistic regression for binary classification by predicting conversion probabilities.
- The model can also be adjusted based on business needs by changing the cut off (e.g., during intern periods or when minimizing calls)