

Lead Scoring Case Study



SUBMITTED BY :

1. Divyansh Parmathi
2. Sejal Agarwal
3. Vaibhav Shah

Contents

- Problem statement
- Problem approach
- Cleaning Data
- EDA
- Model Evaluation
- Observations
- Conclusion

Problem Statement

1. "A" Education, an education company, offers online courses to professionals in the industry. Each day, numerous professionals visit their website to explore available courses. After filling out a form on the website, these individuals are classified as leads by the company.
2. Following lead acquisition, the sales team initiates contact through calls, emails, and other methods. While some leads convert, many do not. The average lead conversion rate at A Education stands at approximately 30%. This implies that out of 100 acquired leads, only about 30 are successfully converted.
3. To enhance efficiency, the company aims to identify the most promising leads, referred to as Hot Leads. By pinpointing these leads, the conversion rate is expected to increase as the sales team can focus their efforts on communicating with high-potential leads rather than reaching out to everyone indiscriminately.

Business Objective

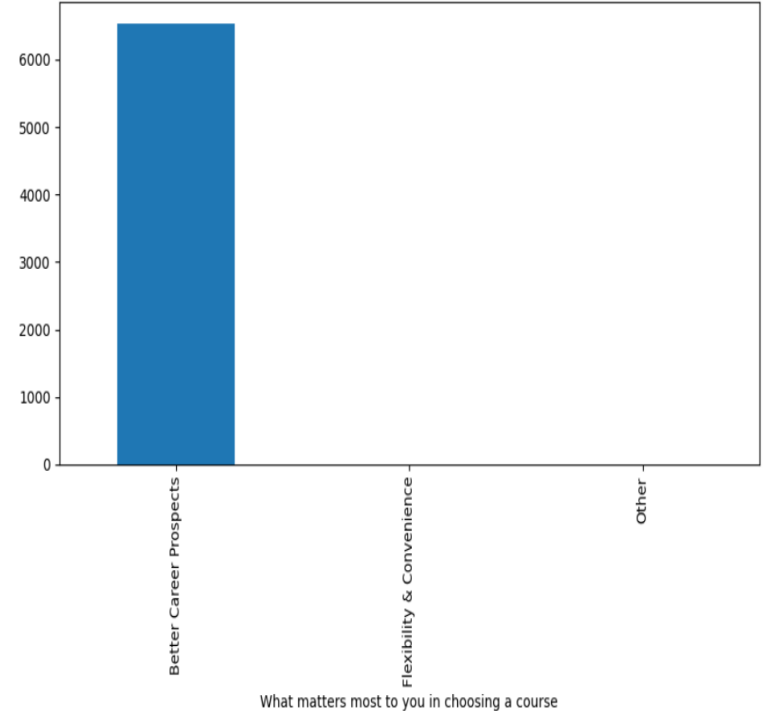
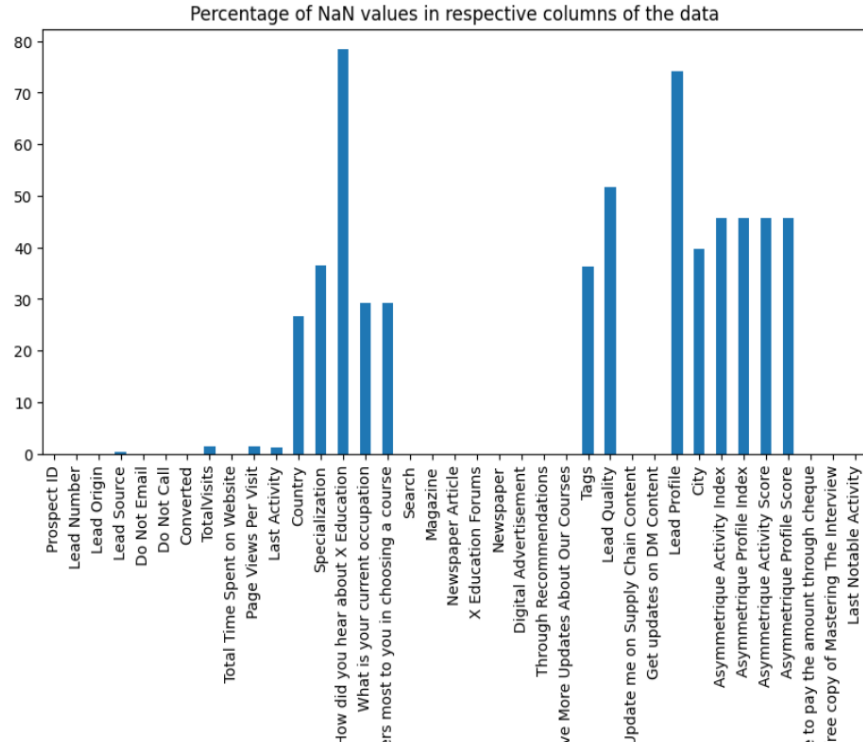
1. Lead X has assigned us the task of creating a model to assign a lead score ranging from 0 to 100 to each lead. This project aims to facilitate the identification of Hot Leads and ultimately enhance the conversion rate.
2. The CEO aims to attain an 80% lead conversion rate.
3. The plan involves deploying the model for future utilization, ensuring its availability and functionality for ongoing operations and decision-making processes. This includes implementing the necessary infrastructure, integrating with existing systems, conducting testing and validation procedures, and establishing monitoring and maintenance protocols to ensure the model's effectiveness and reliability over time.



Problem Approach

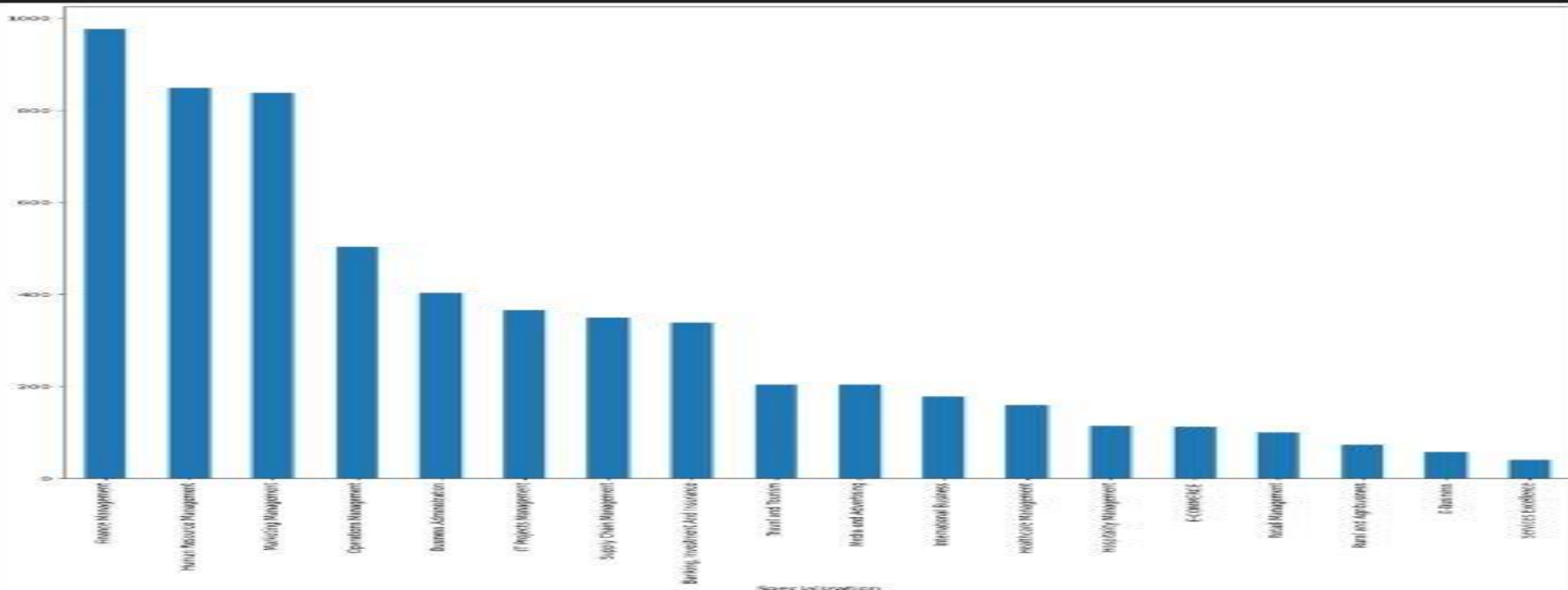


- Reading Data
- Cleaning Data
- EDA
- Creating Dummy
- Splitting data into train & test set
- Building Model
- Model Evaluation
- Making predictions
- Roc Curve
- Prediction on test set
- Precision - Recall



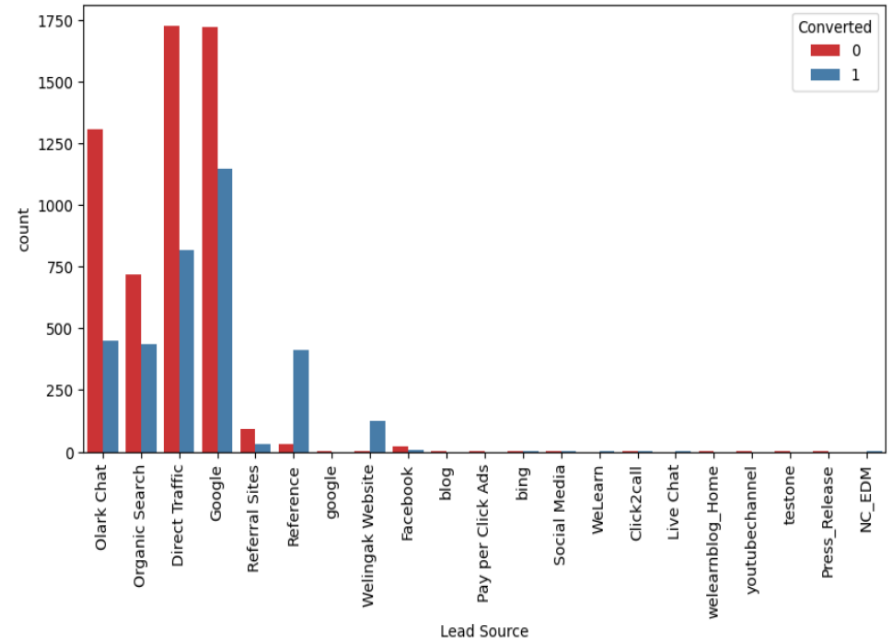
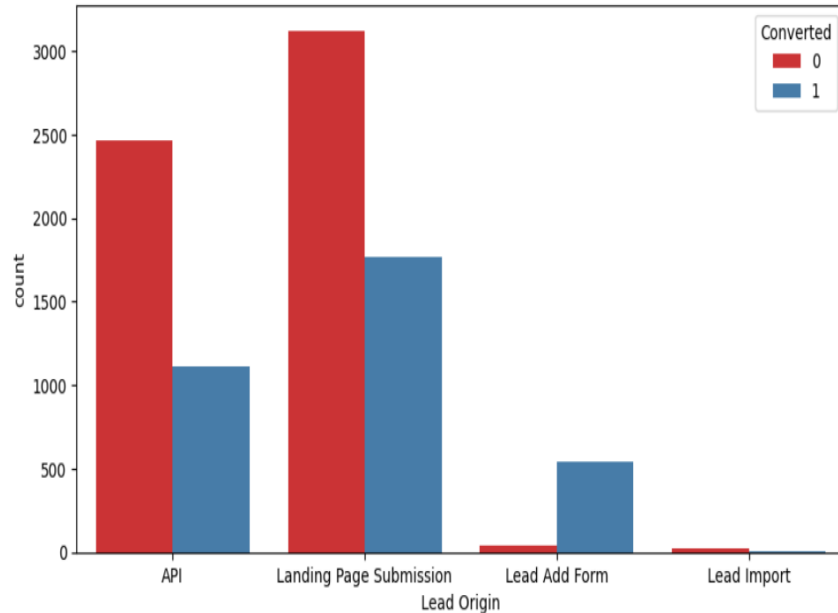
Specialization

Leads from Human Resources, Finance Management & Marketing management specializations have high probability to convert.



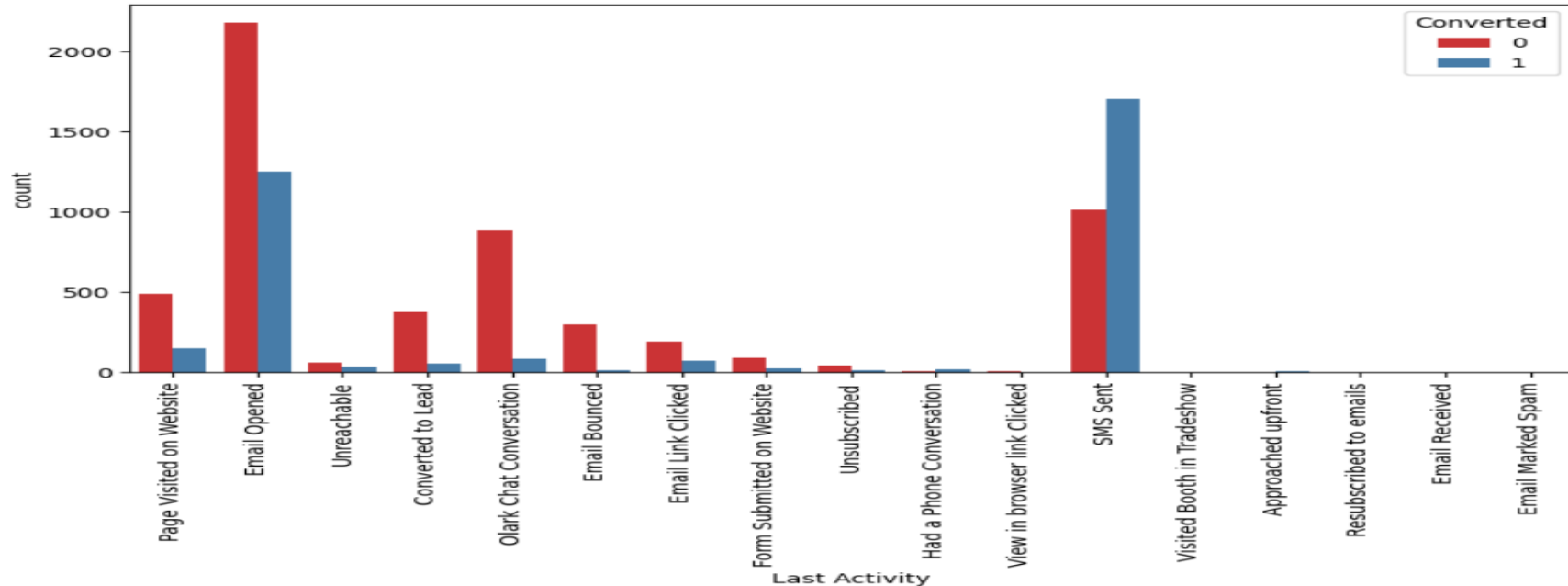
Lead Source & Lead origin

In the lead source, leads generated through Google and direct traffic exhibit a higher probability of conversion. Conversely, in Lead origin, the majority of leads are generated through submissions.



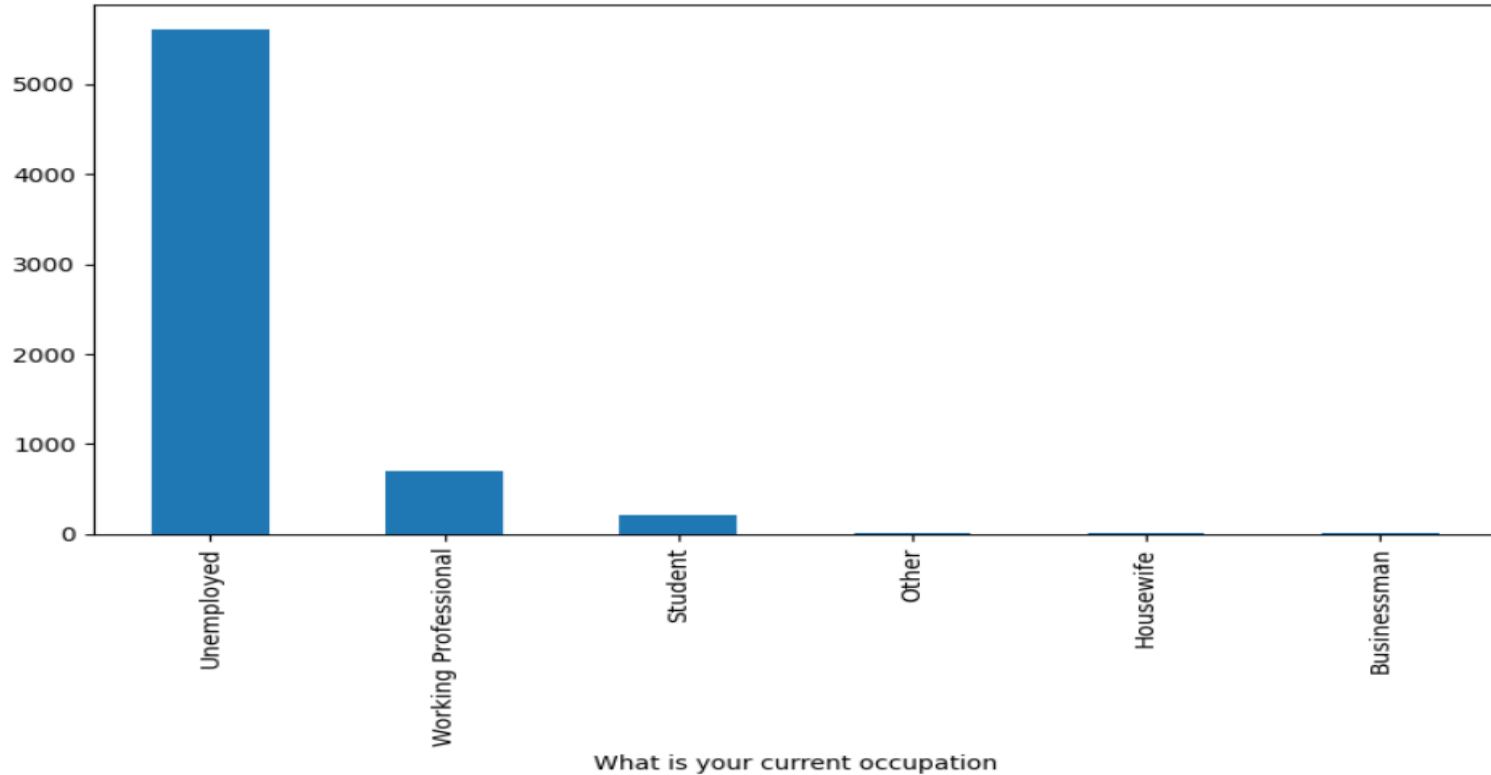
Last lead Activity

Leads which are Email Opened have high probability to convert, same as sending SMS will also benefit.



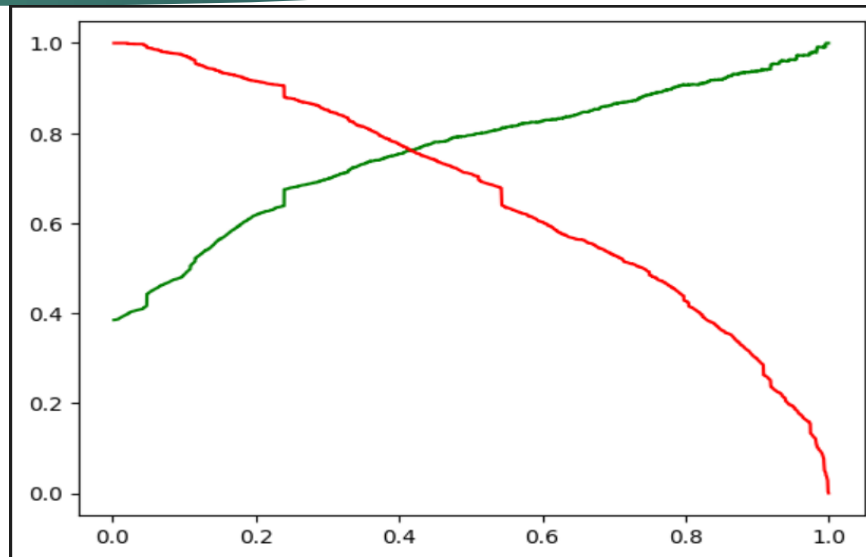
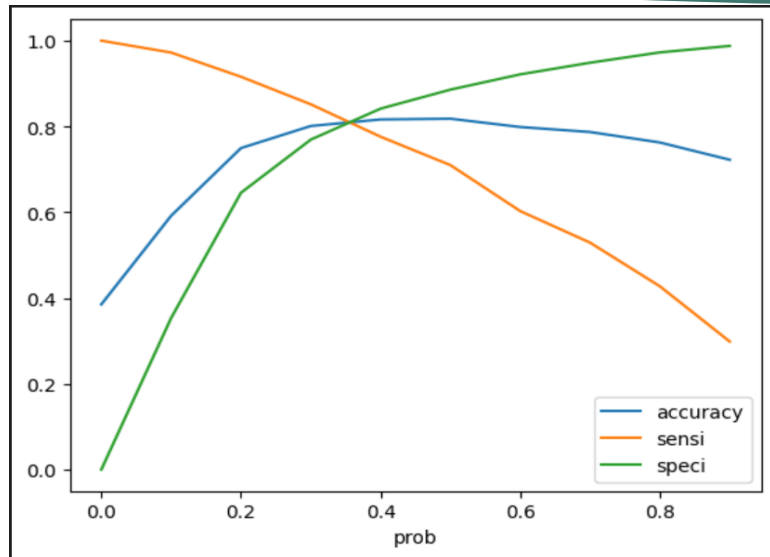
Your Occupation

Leads which are Unemployed are more interested to join the course than others.



Model Evaluation

ROC curve



Observations

Train Data:

Accuracy : 81.32%

Sensitivity : 81.43%

Specificity : 81.25%

Test Data:

Accuracy : 80.7%

Sensitivity : 79.87%

Specificity : 81.11%

- Lead Source_Olark Chat
- Specialization_Others
- Lead Origin_Lead Add Form
- Lead Source_Welingak Website
- Total Time Spent on Website
- Lead Origin_Landing Page Submission
- What is your current occupation_Working Professionals
- Do Not Email

Conclusion

- Conversion rates for API and Landing page submissions fall within the range of 30% to 35%, which closely aligns with the average. However, the conversion rates for Lead Add form and Lead import are notably lower. This observation highlights the need for greater attention to be focused on leads originating from API and Landing page submissions.
- The bulk of leads are sourced from Google/direct traffic, with the highest conversion ratio observed for leads from references and the Welingak website.
- Leads who spend more time on the website demonstrate a higher likelihood of conversion.
- Opening emails is the most common last activity among leads, while the highest conversion rate is associated with SMS sent. Additionally, the majority of leads are unemployed, although the highest conversion rates are seen among working professionals.