# Lead Score Case Study

GROUP MEMBERS

1. **Kruti Dhabalia**

2. **Kuntal Dutta**

3. **Kundan Kumar**

# Problem Statement

❑X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted.

❑To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'.

❑If they successfully identify this set of leads, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone.
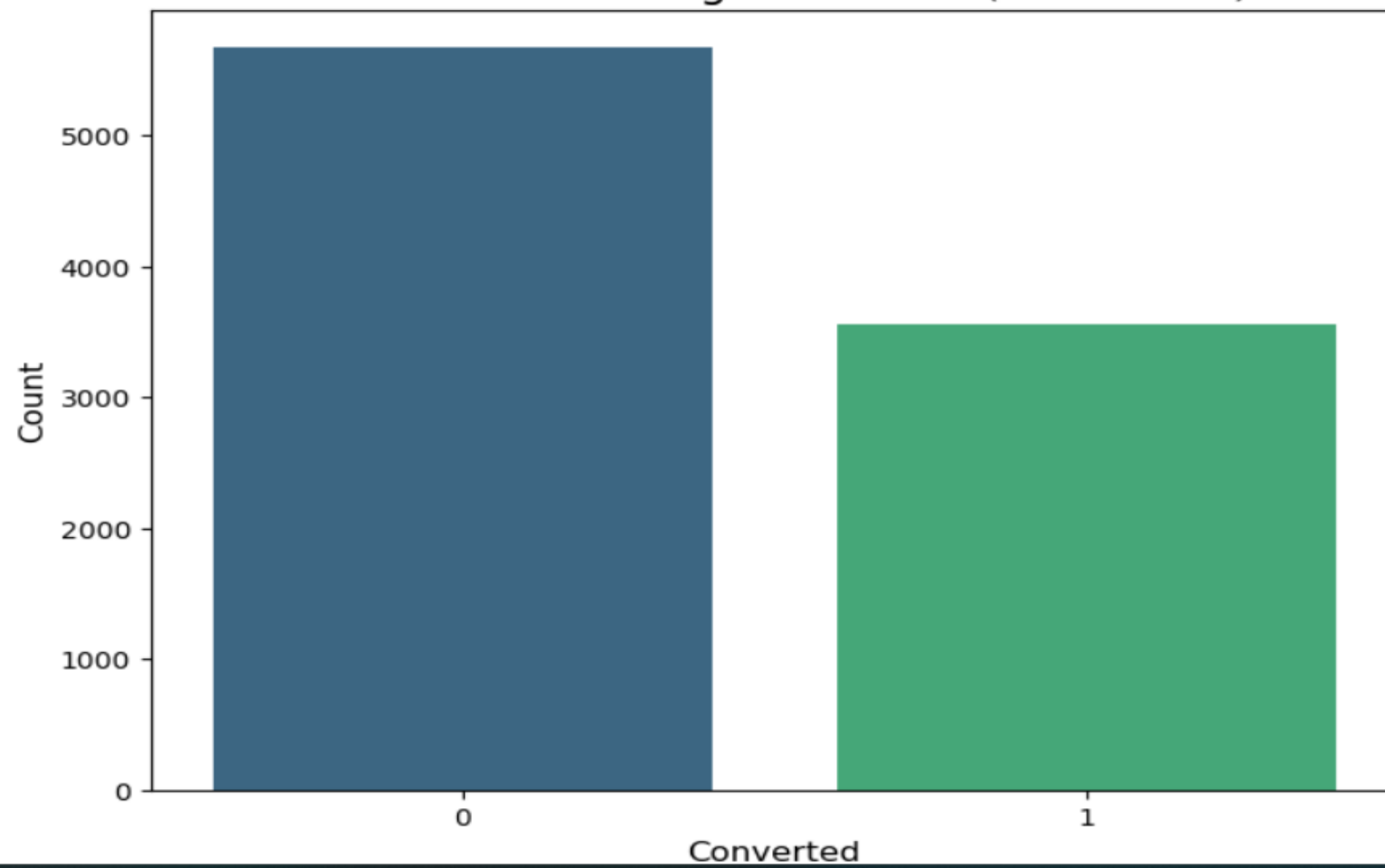
# Data cleaning and data manipulation

1. Remove columns with more than 50% missing values.

2. Filling missing values for numerical columns with the median.

3. Filling missing values for categorical columns with the mode.

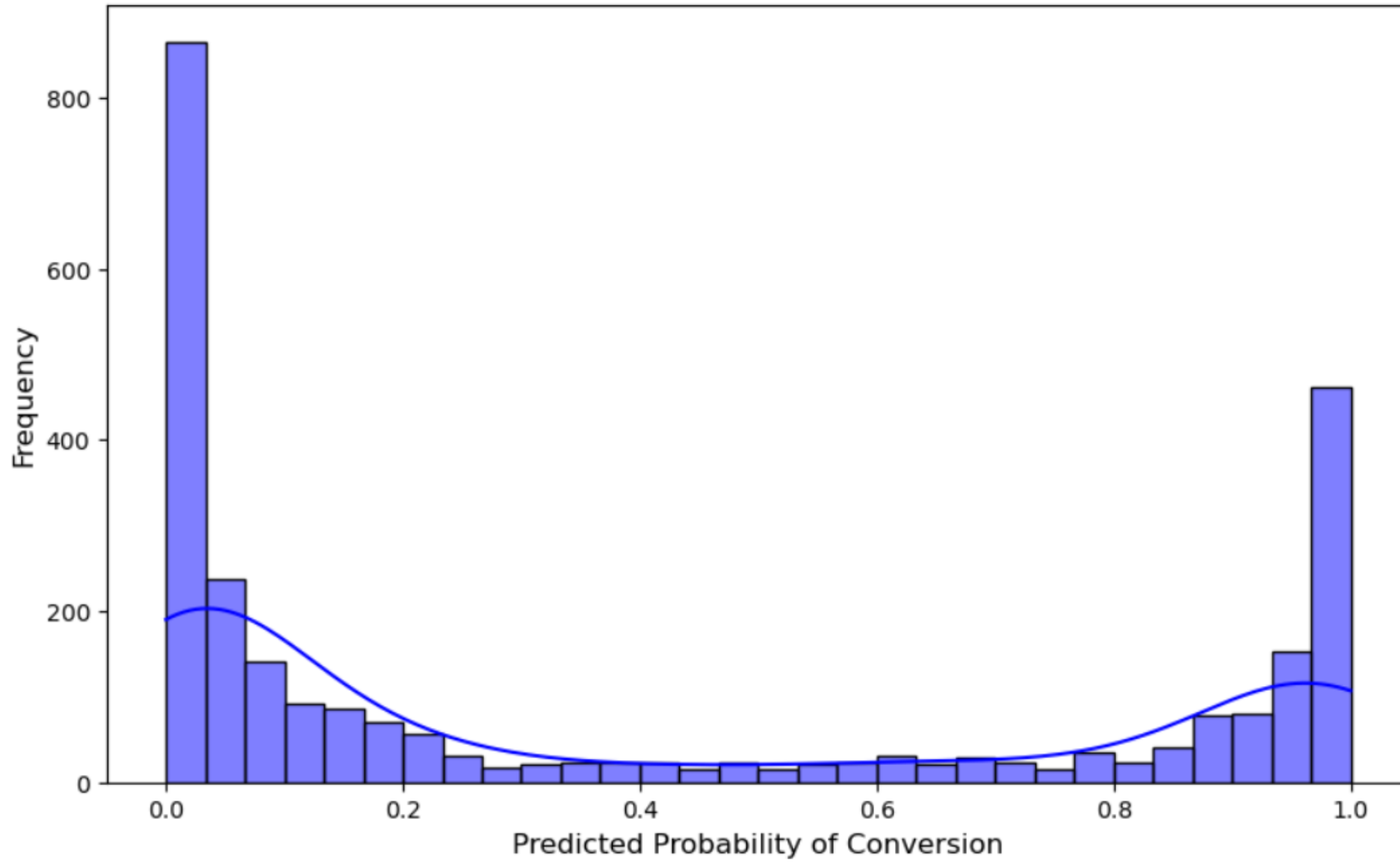4. Dropping columns that are irrelevant for modeling.
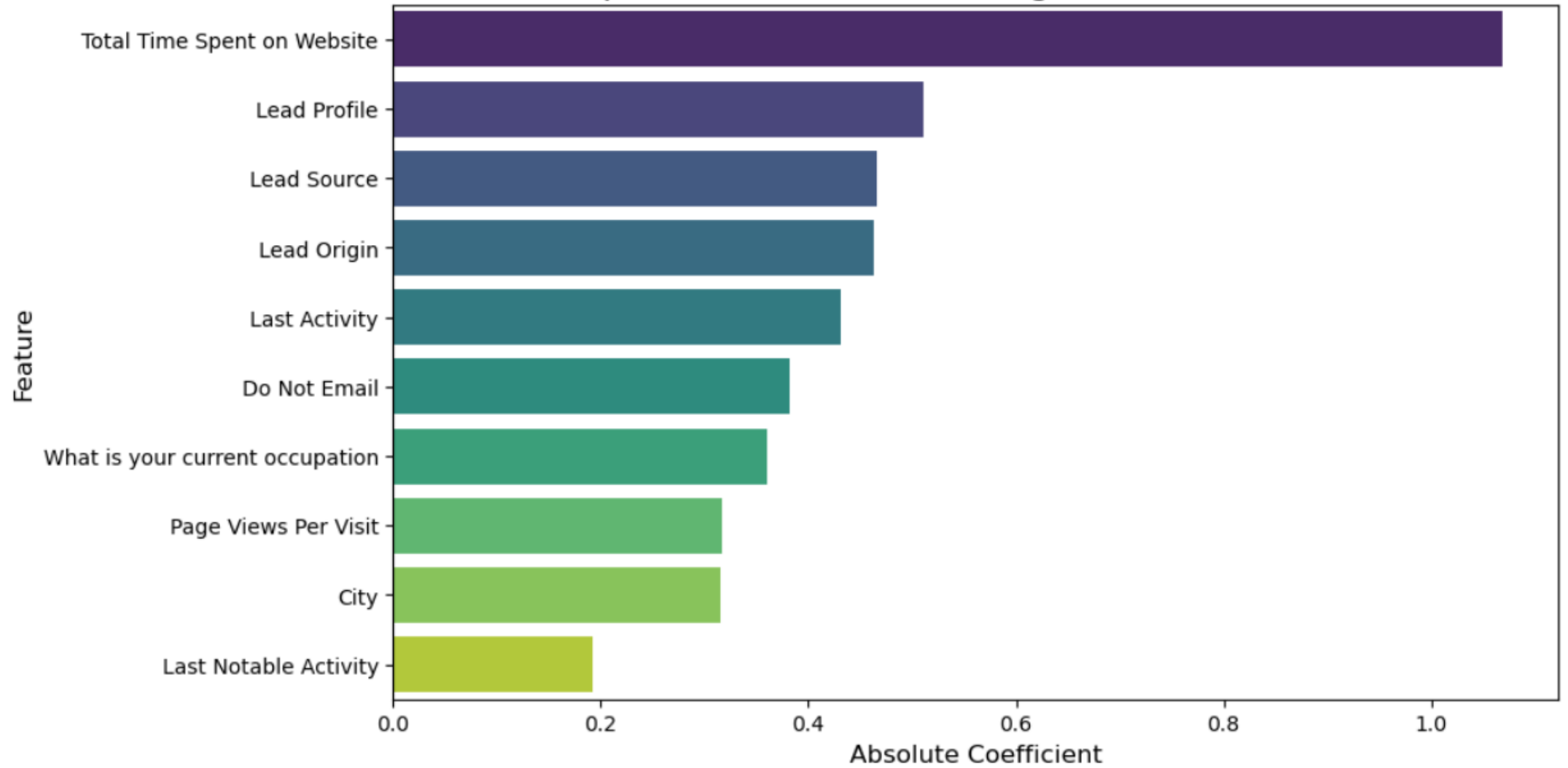
# ROC Curve
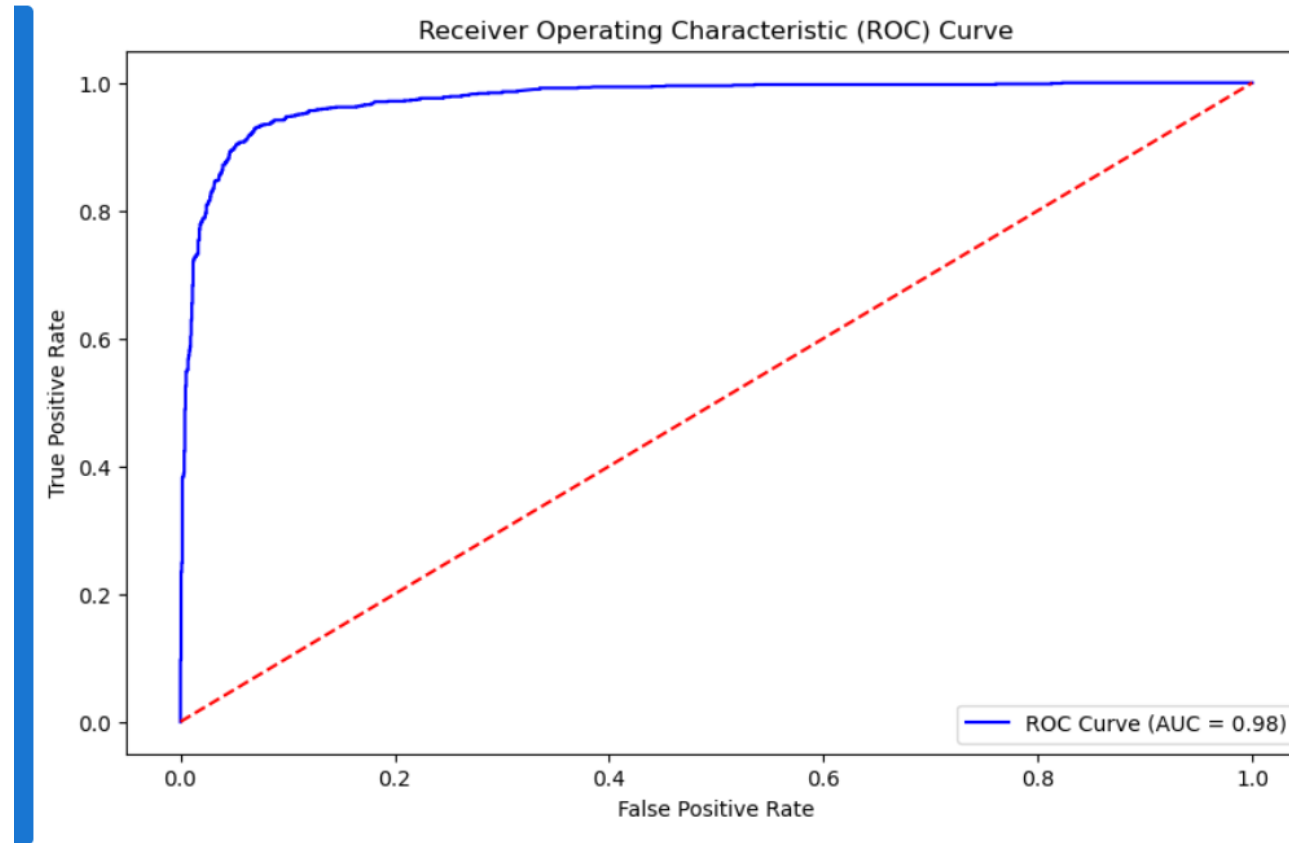
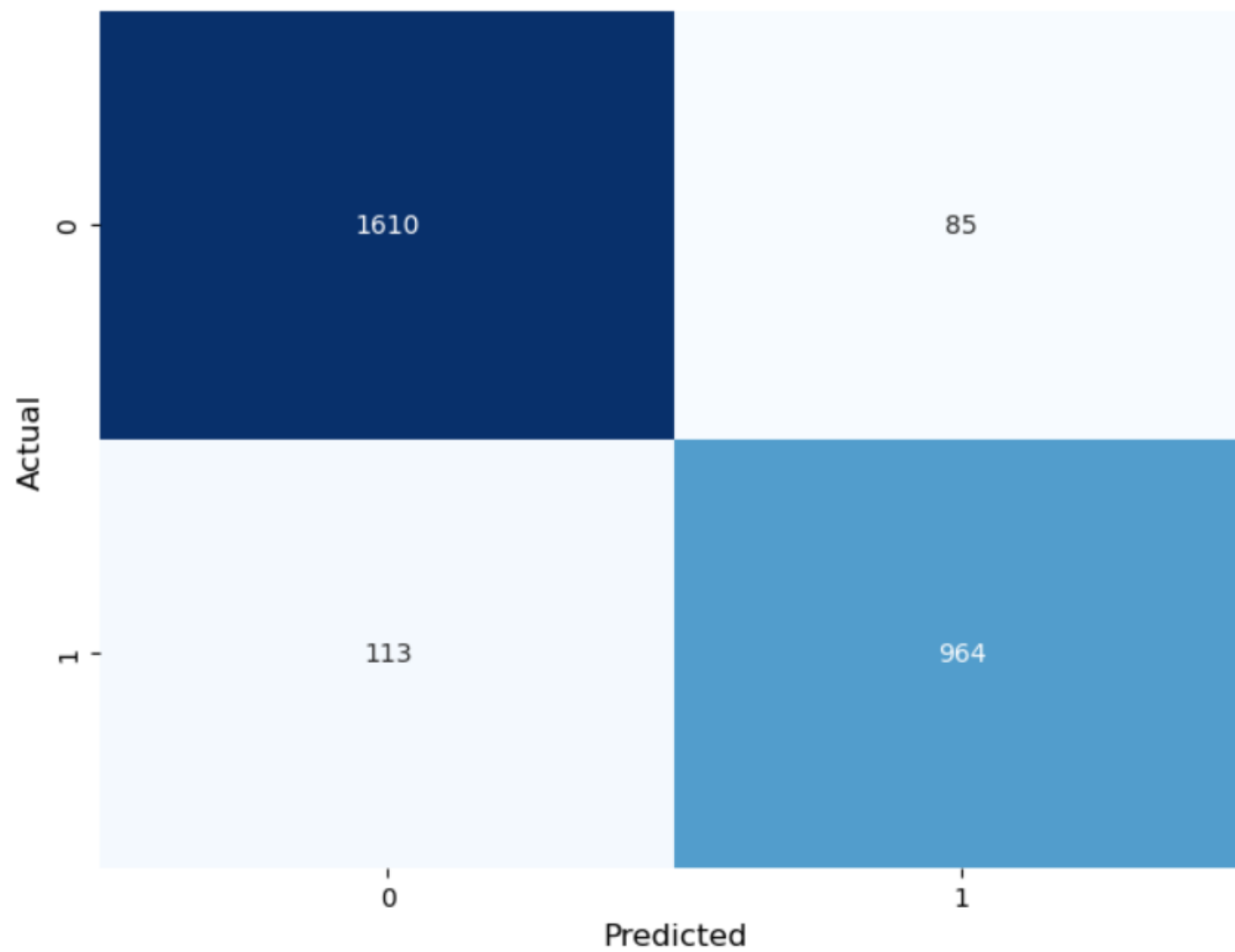Distribution of Target Variable (Converted)

Distribution of Predicted Probabilities

Top 10 Features Contributing to Lead Conversion

Confusion Matrix

# Conclusion

Based on the model's coefficients, the top three variables contributing the most towards the probabilitxy of a lead being converted are:

Total Time Spent on Website

Lead Source

Last Activity