

LEAD SCORING OPTIMIZATION

Presentation for X
Education Lead Scoring
By –

- Vaishnavi Muthane
- Deepak Narayanan
- Deepak Goyal

PROBLEM STATEMENT

- X Education, an online course provider for professionals, faces a low lead conversion rate of around 30%. Despite generating many leads through marketing on websites and search engines, only a few convert into paying customers. To improve this, the company aims to identify "Hot Leads"—those most likely to convert—so the sales team can focus their efforts more efficiently. The goal is to boost the conversion rate to around 80%. To achieve this, a model will be built to assign a lead score to each lead, indicating their likelihood of conversion, thereby enhancing the effectiveness of the sales team's efforts.

BUSINESS GOAL OF THE CASE STUDY

Increase Conversion Rate: Boost from 30% to 80% by identifying high-potential leads.

Optimize Sales Efforts: Focus on leads with higher conversion probabilities.

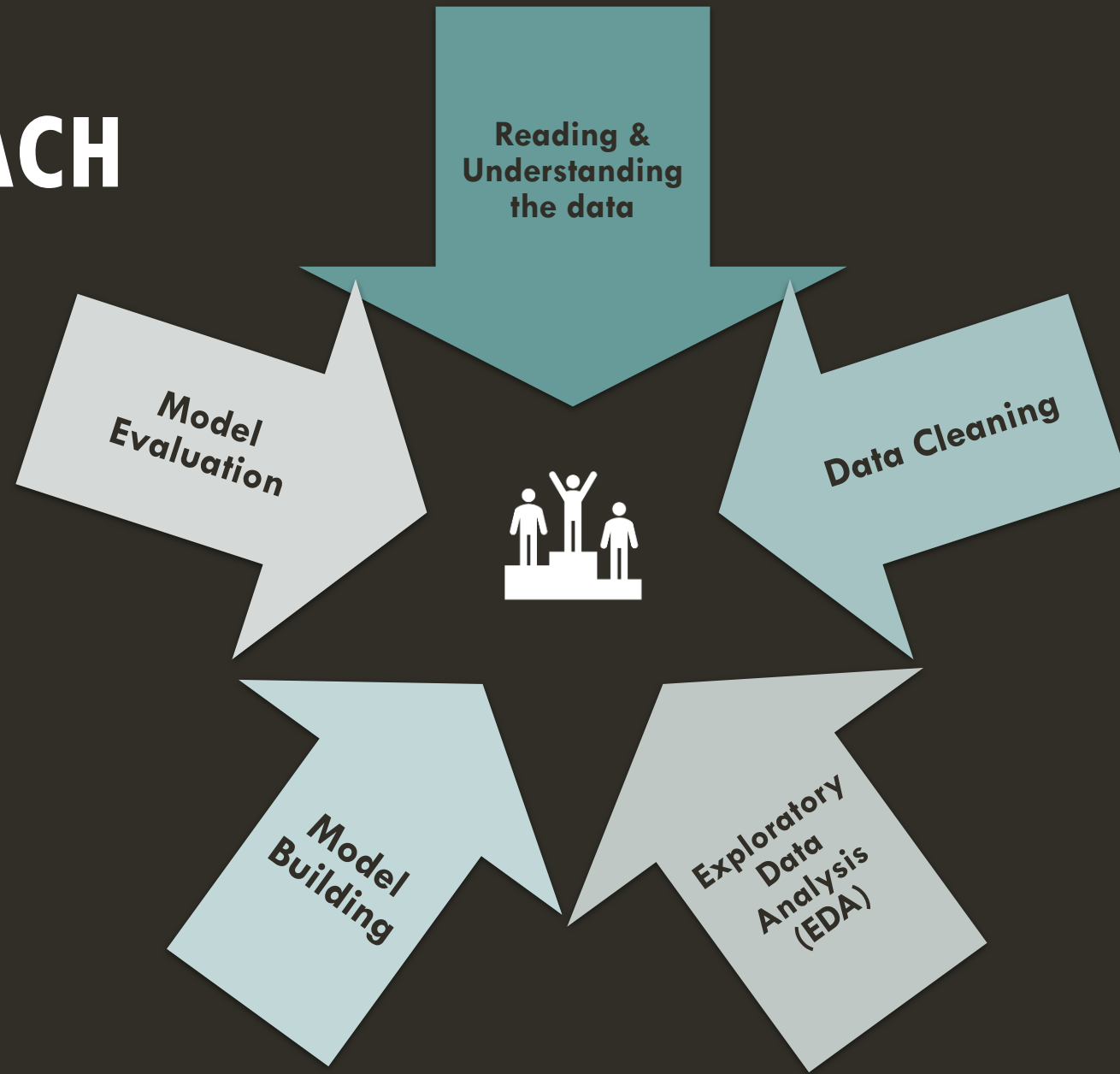
Personalize Outreach: Use lead scores (0-100) to target potential customers.

Flexible Model: Ensure adaptability to future needs.

Data-Driven Insights: Inform decisions with key metrics and historical data.

Sustainable Growth: Improve conversion rates and customer satisfaction for long-term success.

APPROACH



DATA PREPARATION AND CLEANING

1. Handling Missing Values:

- Columns with more than 35% missing values were dropped.
- Imputation of missing values for other columns.
- 'Not Specified' was used for some missing categorical values to retain data integrity.

2. Categorical Variables:

- Conversion of categorical variables to dummy variables.
- Removal of 'Select' levels as they represent missing values.

3. Numeric Variables:

- Applied MinMaxScaler to scale numeric columns for consistency.

4. Data Splitting:

- Split the dataset into training (70%) and testing (30%) sets.

EXPLORATORY DATA ANALYSIS (EDA)

Distribution of Converted Leads:

- Visualized the distribution of the target variable 'Converted'.
- Approximately 30% of leads were converted.

Lead Source Analysis:

- Identified top lead sources with significant differences in conversion rates.

Time Spent on Website:

- Leads spending more time on the website had higher conversion rates, highlighting the importance of engaging website content.

Lead Origin:

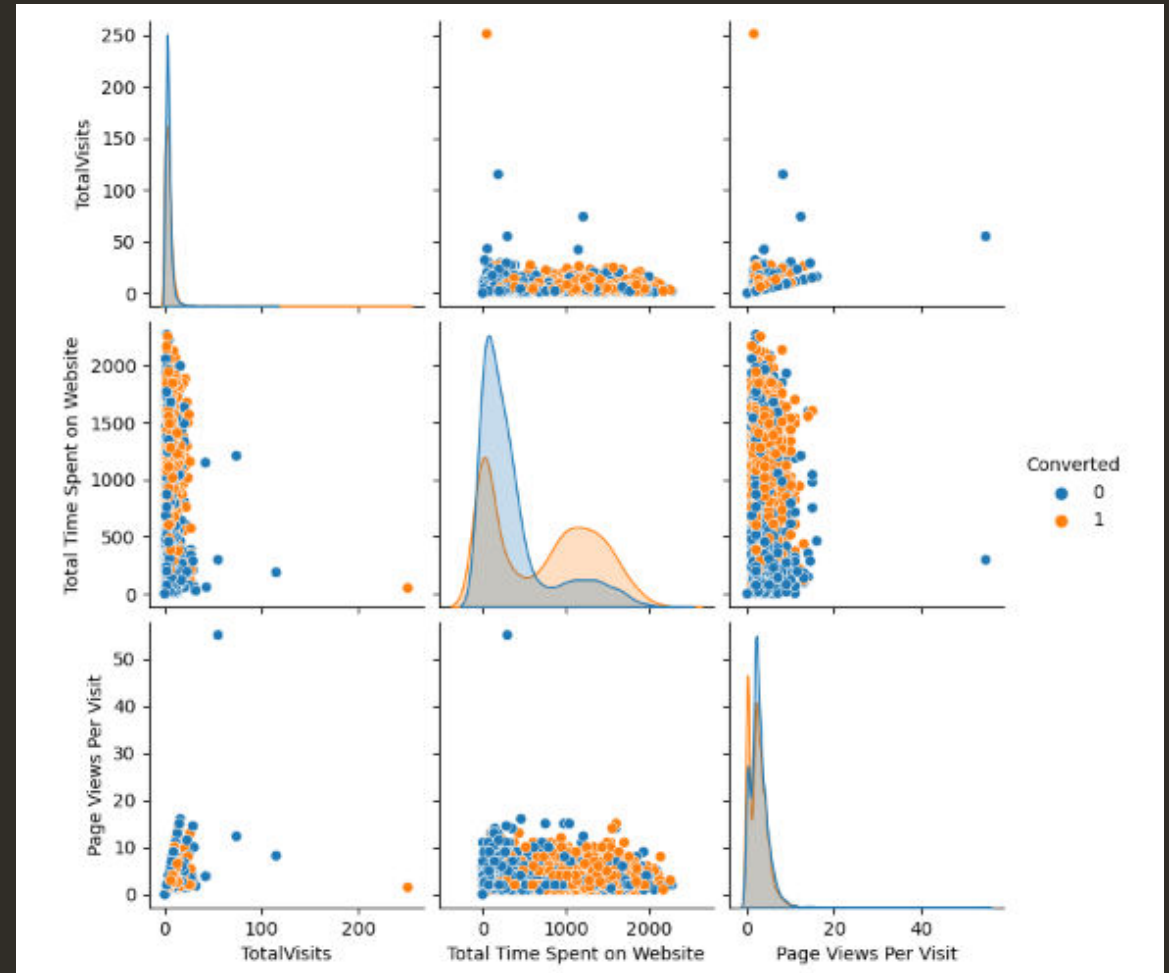
- Certain origins, like 'Lead Add Form', had higher conversion probabilities.

Last Activity:

- Activities like 'SMS Sent' and 'Olark Chat Conversation' correlated with higher conversion rates.

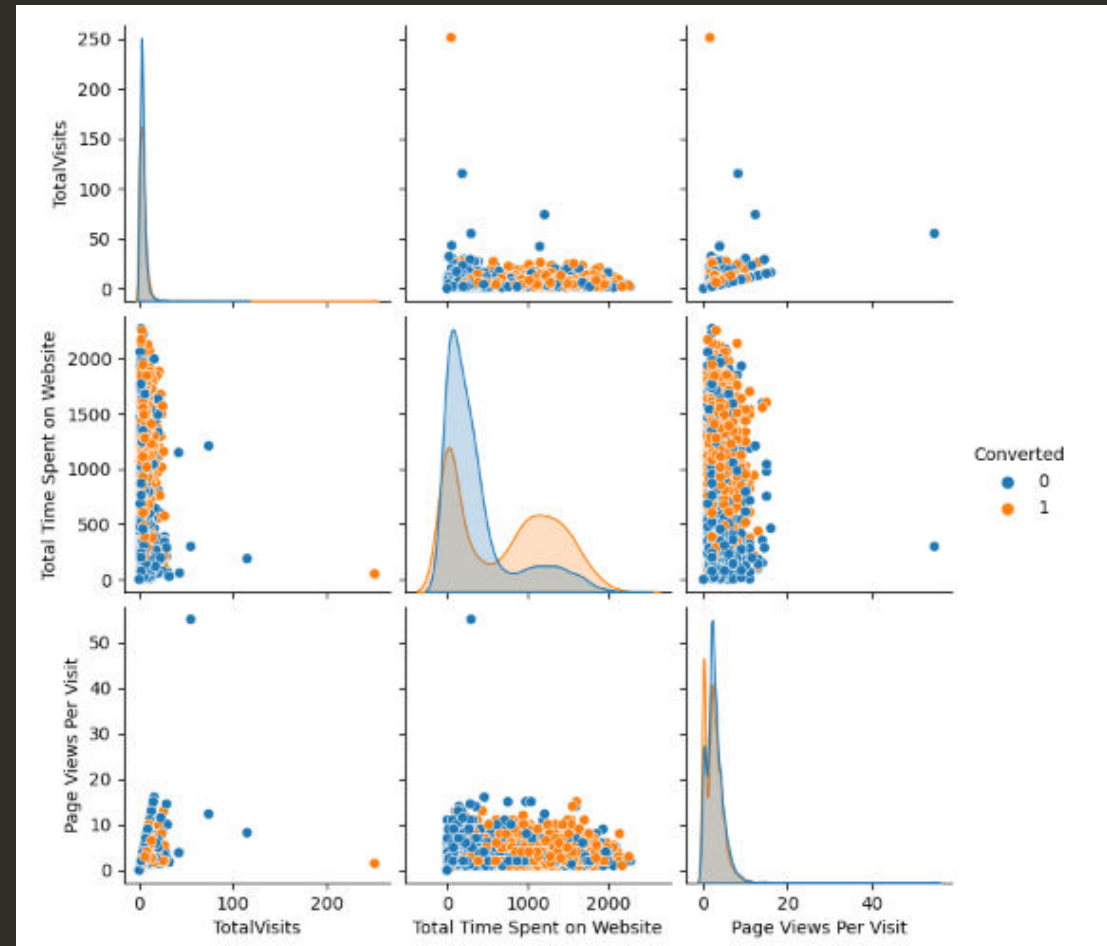
VISUALIZING THE DATA RELATIONSHIPS BY CONVERSION

**ANALYSIS - IT SHOWS THE IMPACT
ON PAIR FEATURE RELATIONSHIPS**



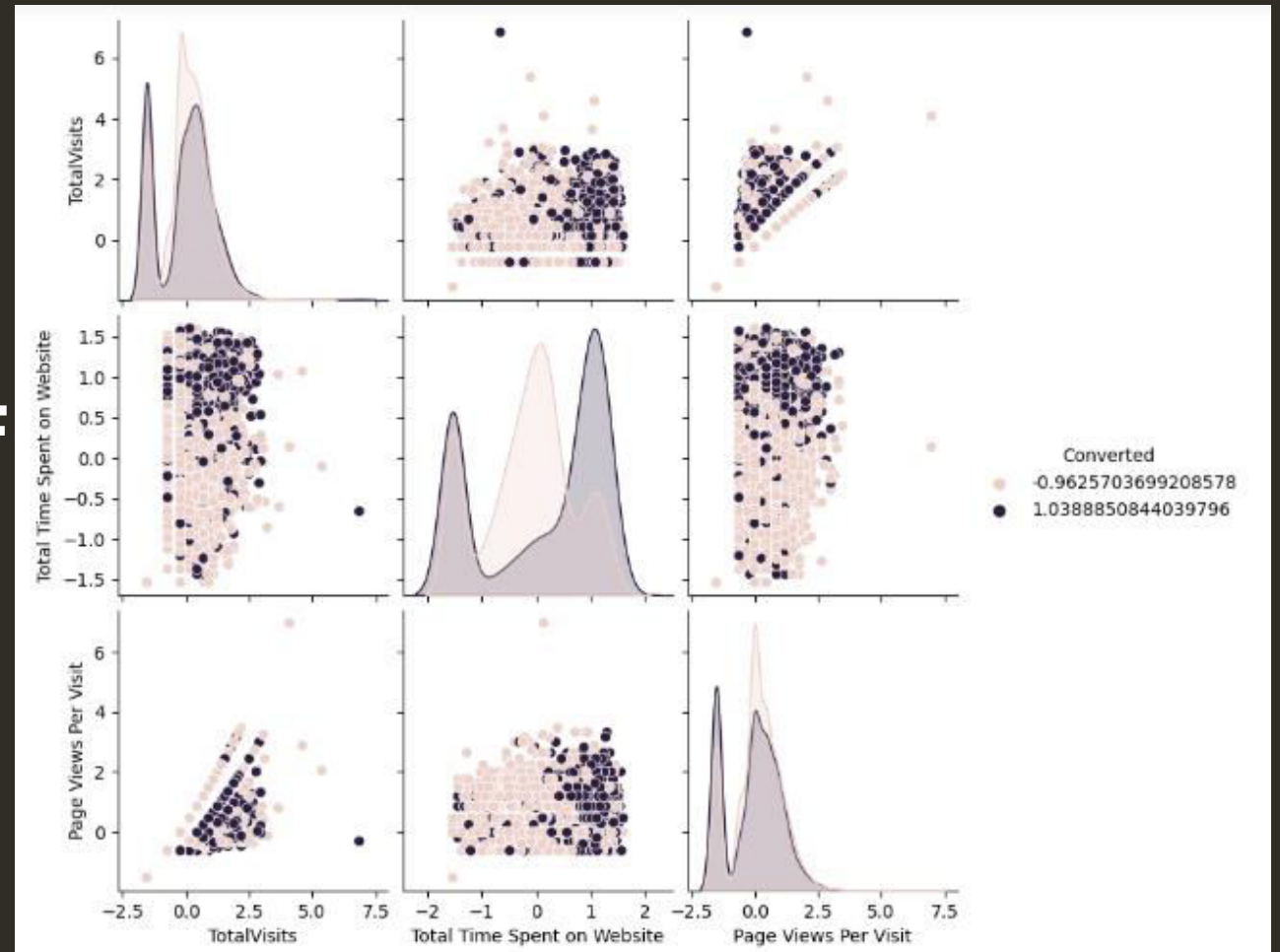
ANALYZING LEAD CONVERSION FEATURE RELATIONSHIPS

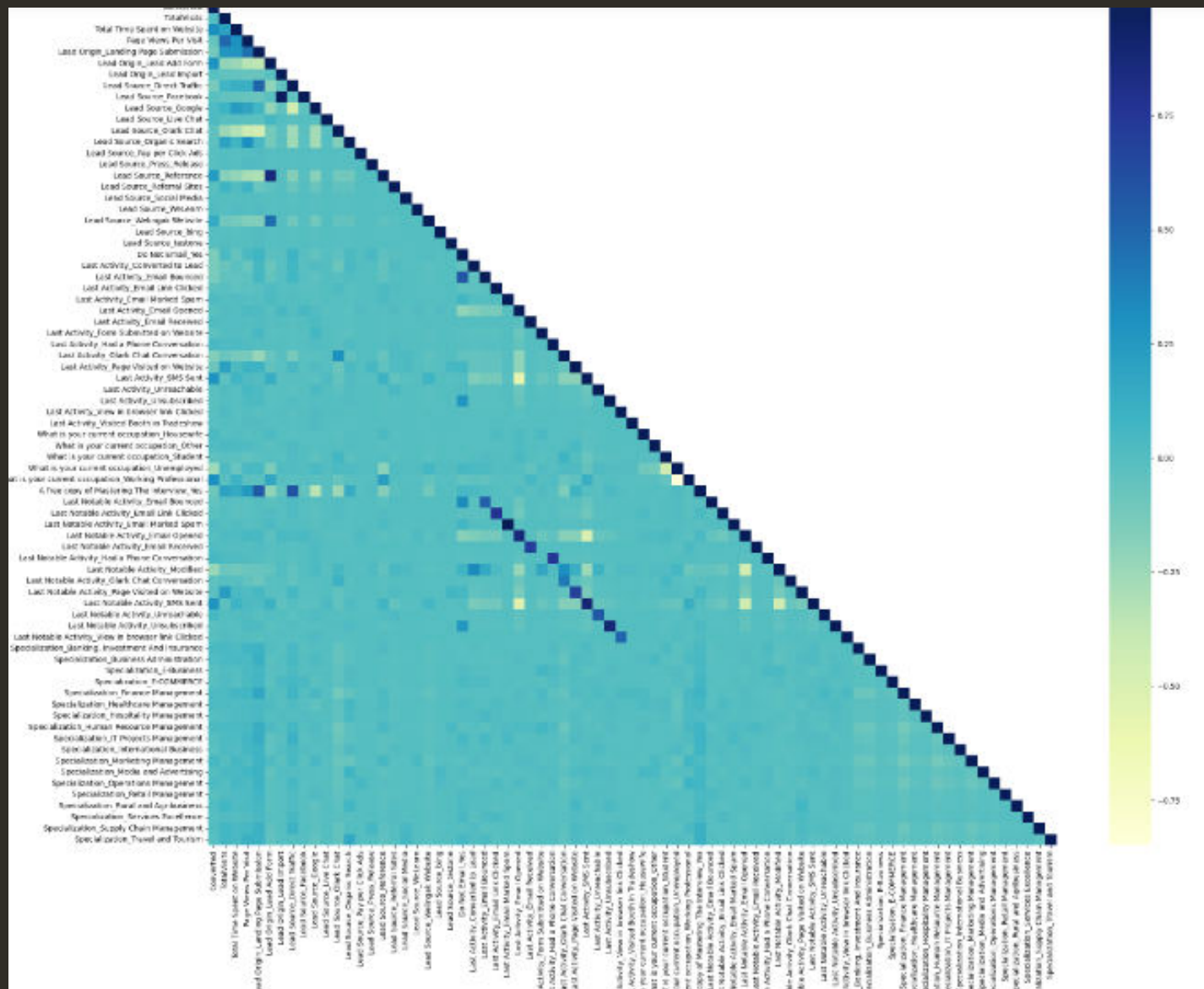
ANALYSIS - FEATURE INTERACTIONS AFFECTING LEAD CONVERSION PROBABILITY



VISUALIZING TRANSFORMED LEAD CONVERSION DATA

ANALYSIS - DISTRIBUTION OF TRANSFORMED LEAD CONVERSION DATA





MODEL EVALUATION AND PERFORMANCE METRICS

Evaluation Process:

- Assess model performance using various metrics.
- Compare predicted outcomes with actual data.

Key Metrics:

- Accuracy, precision, recall, and F1-score.
- ROC curve and AUC for binary classification.

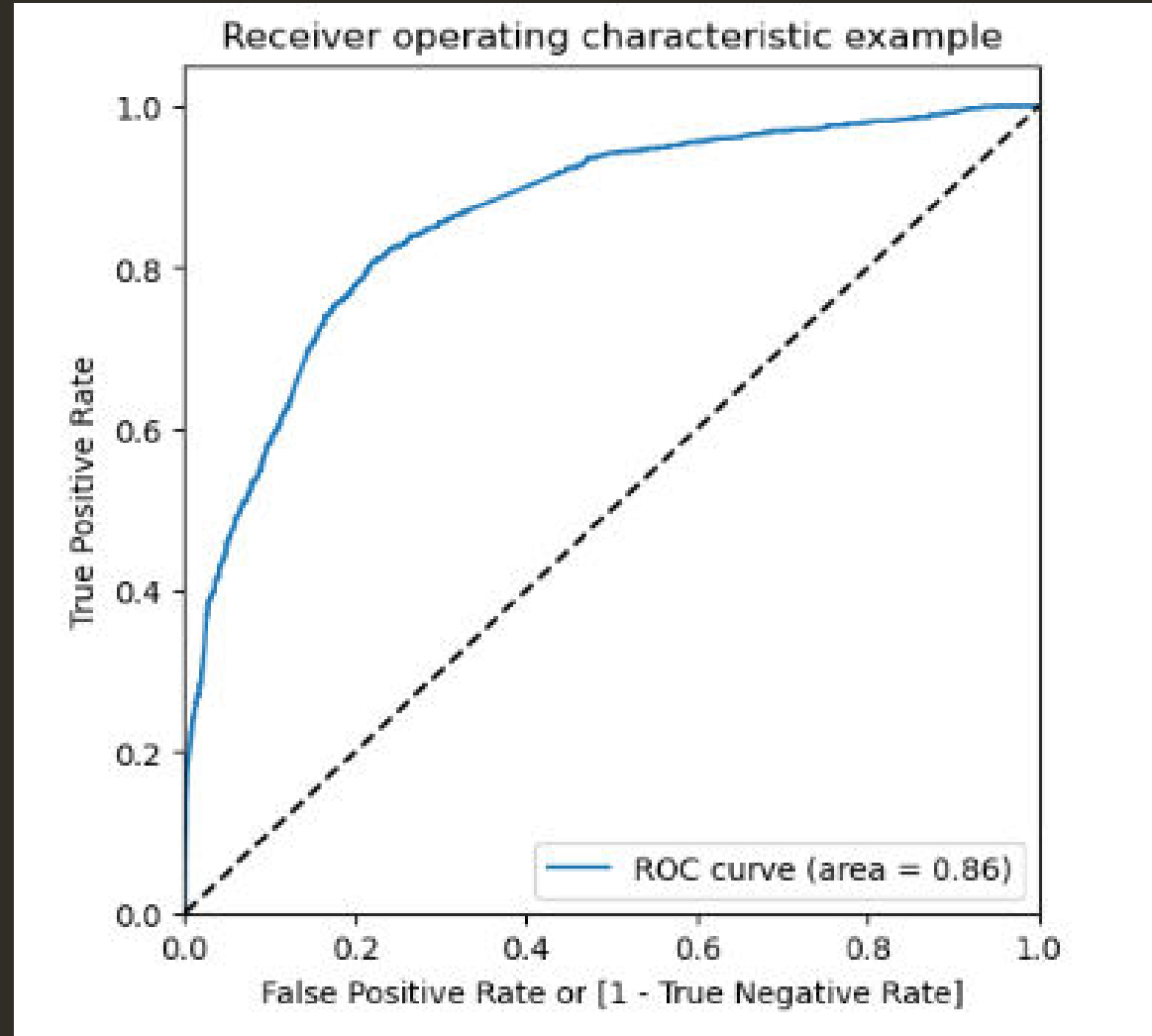
Interpretation:

- Understand strengths and weaknesses of the model.
- Choose metrics aligned with project goals.

Improvement:

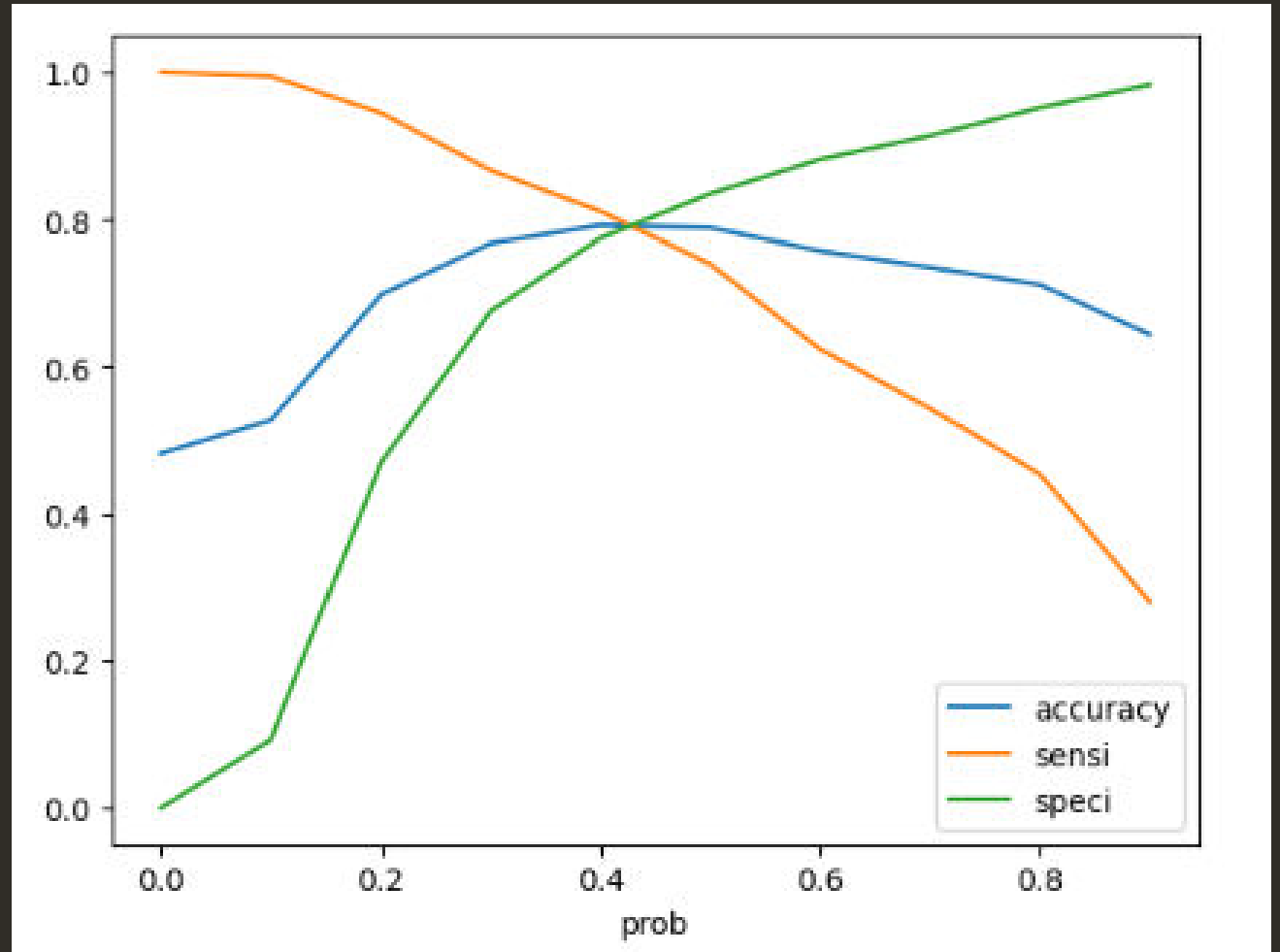
- Iteratively refine models based on evaluation results.
- Optimize hyperparameters for better performance.

**ROC CURVE
ANALYSIS - ROC
EVALUATES MODEL'S
CLASSIFICATION
PERFORMANCE.**

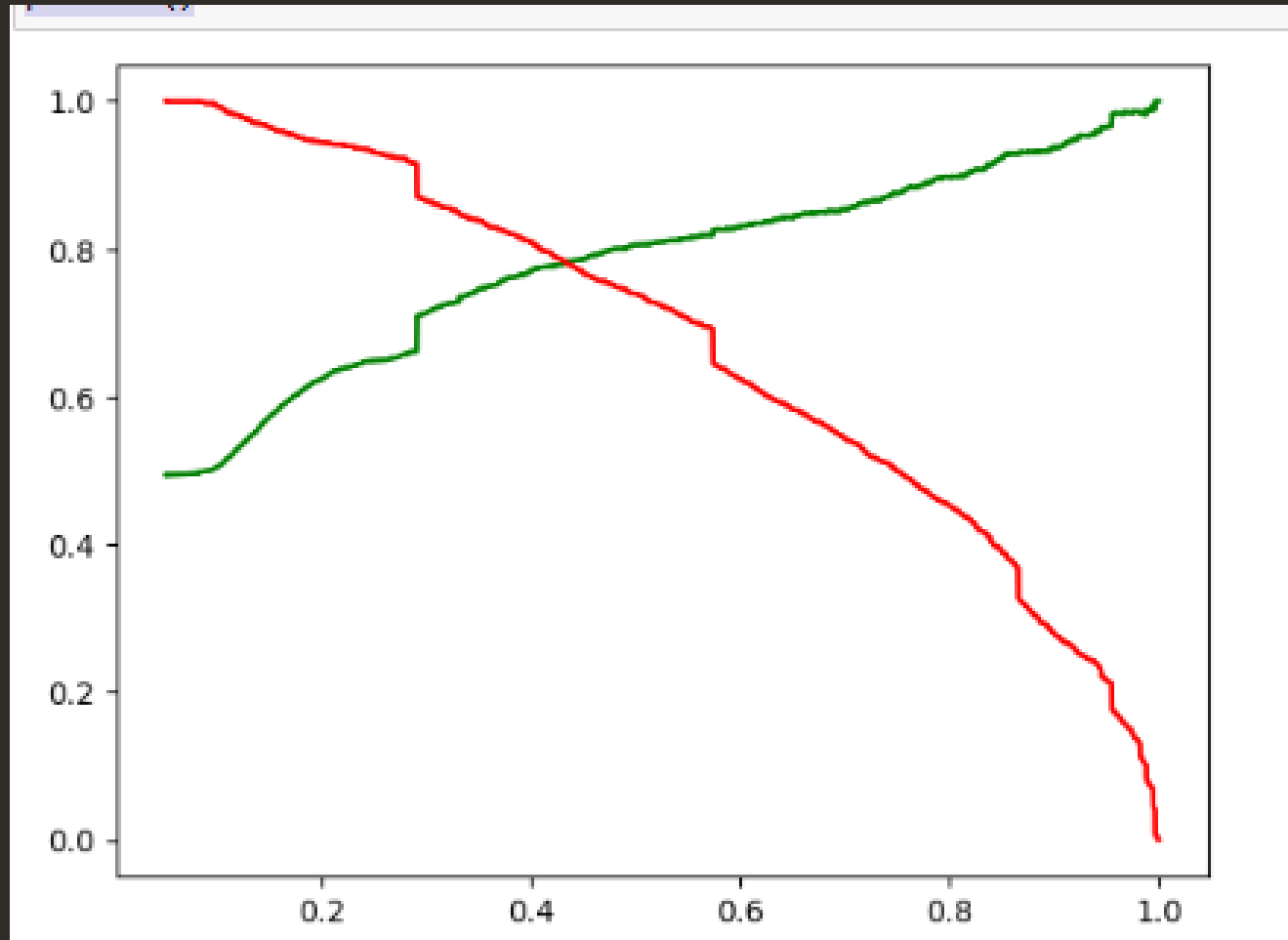


VISUALIZING CUTOFF THRESHOLDS FOR CLASSIFICATION

ANALYSIS - EVALUATION OF CLASSIFICATION THRESHOLD PERFORMANCE.



PLOTTING PRECISION AND RECALL CURVES.



**Analysis -
Precision and
recall vary with
threshold.**

FINDING THE OPTIMAL CUTOFF

- **Accuracy of the model is 79.09 %**
- **Sensitivity score of the model: 79.34%**
- **Specificity score of the model: 78.85%**

MAKING PREDICTIONS ON THE TEST SET

- Overall accuracy score of the model: 78.45%
- Sensitivity score of the model: 77.95%
- Specificity score of the model: 78.92%
- Precision score of the model: 78.40%
- Recall score of the model: 77.71 %

SUMMARY

To improve lead conversion, identify top prospects through metrics like 'TotalVisits', 'Total Time Spent on Website', and 'Page Views Per Visit'. Nurture leads with consistent communication and education about offerings. Maintain a comprehensive list to keep leads informed about relevant courses and opportunities. Prioritize converted leads by engaging in Q&A sessions and understanding their needs through appointments. Tailor offerings to align with their interests and maximize conversion rates. This approach ensures efficient lead management and personalized engagement, resulting in higher conversion rates and increased business success.



THANK YOU