LEAD SCORING ANALYSIS

Presented to: CHIEF DATA SCIENTIST

Prepared by: ASHWANI SINGH, ARNAB DAS & ARIJIT DAS

Date: 18TH MARCH-2025

INTRODUCTION

- **Objective**: Improve lead conversion rates through data-driven insights.
- **Approach**: Analyze past lead data to identify high-converting patterns.
- **Business** Impact: Optimized marketing efforts and increased sales efficiency.

DATA OVERVIEW

- **Dataset:** Leads.csv (Customer interactions & lead characteristics)
- **Features:** Demographics, engagement levels, marketing channels, etc.
- **Target Variable:** Conversion (Lead converted to a customer)

DATA CLEANING & PREPERATION

- Handled missing values & outliers.
- Encoded categorical variables.
- Feature engineering for better predictive power.

EXPLORATORY DATA ANALYSIS (EDA)

- Key insights from the dataset:

  - Top sources of high-converting leads.
  - Customer behavior patterns.
  - Impact of engagement on conversions.

- Visualizations (charts & graphs to support findings).
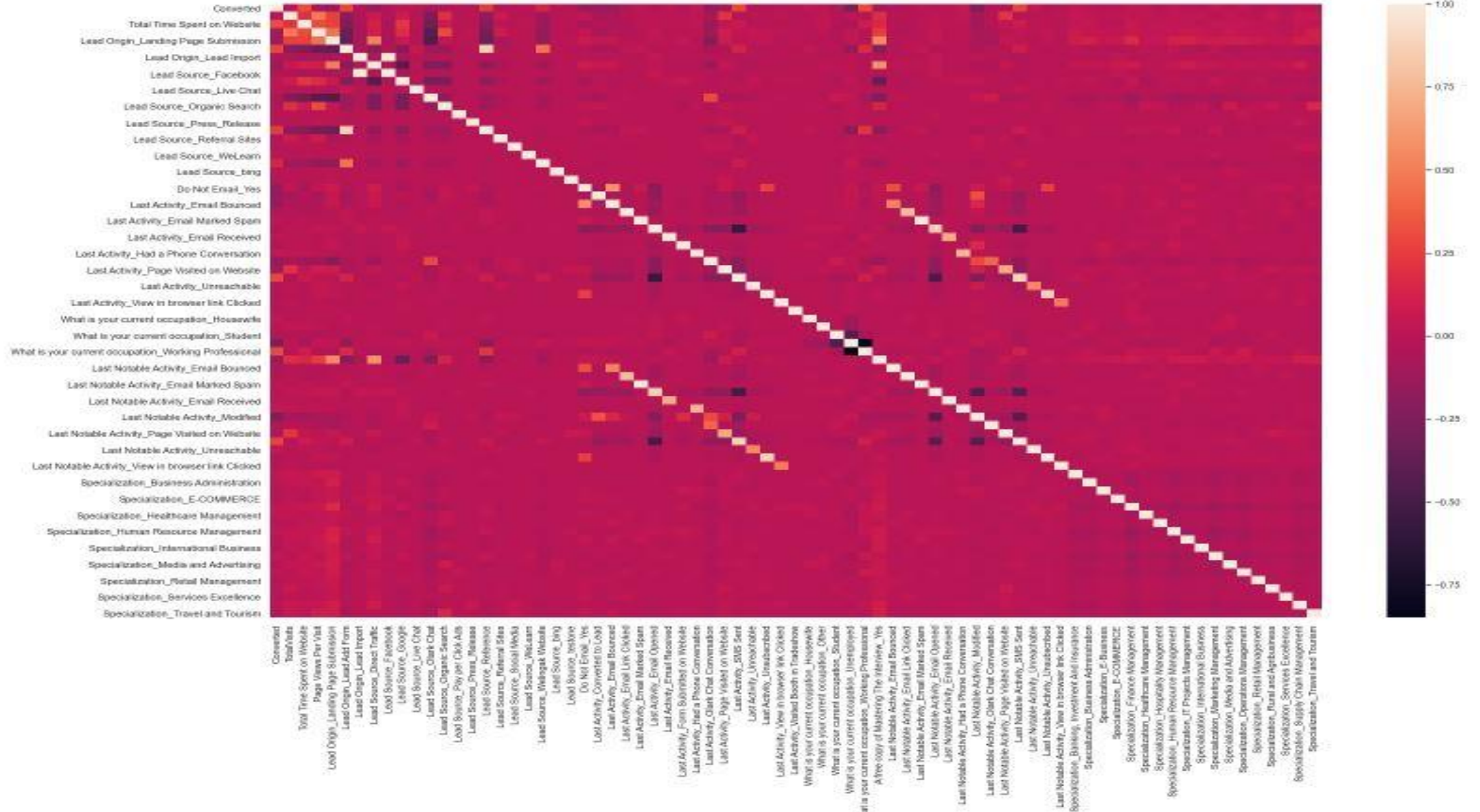
CORRELATION BETWEEN 3 VARIABLES
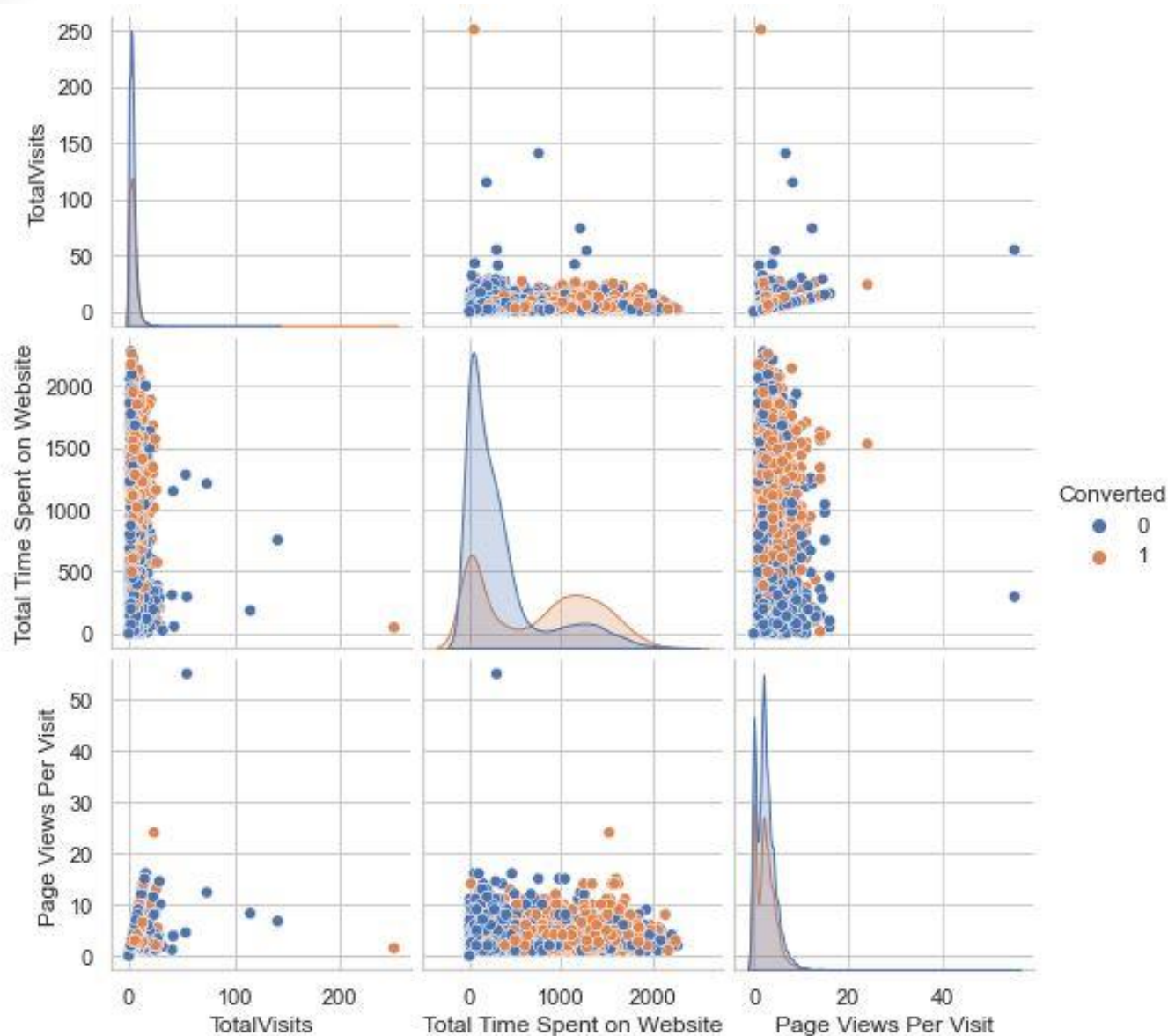
Graphical representation depicting the correlation between variables in terms of conversion rate:

➢ Total Visits
➢ Total Time spent on Website
➢ Page Views per Visit
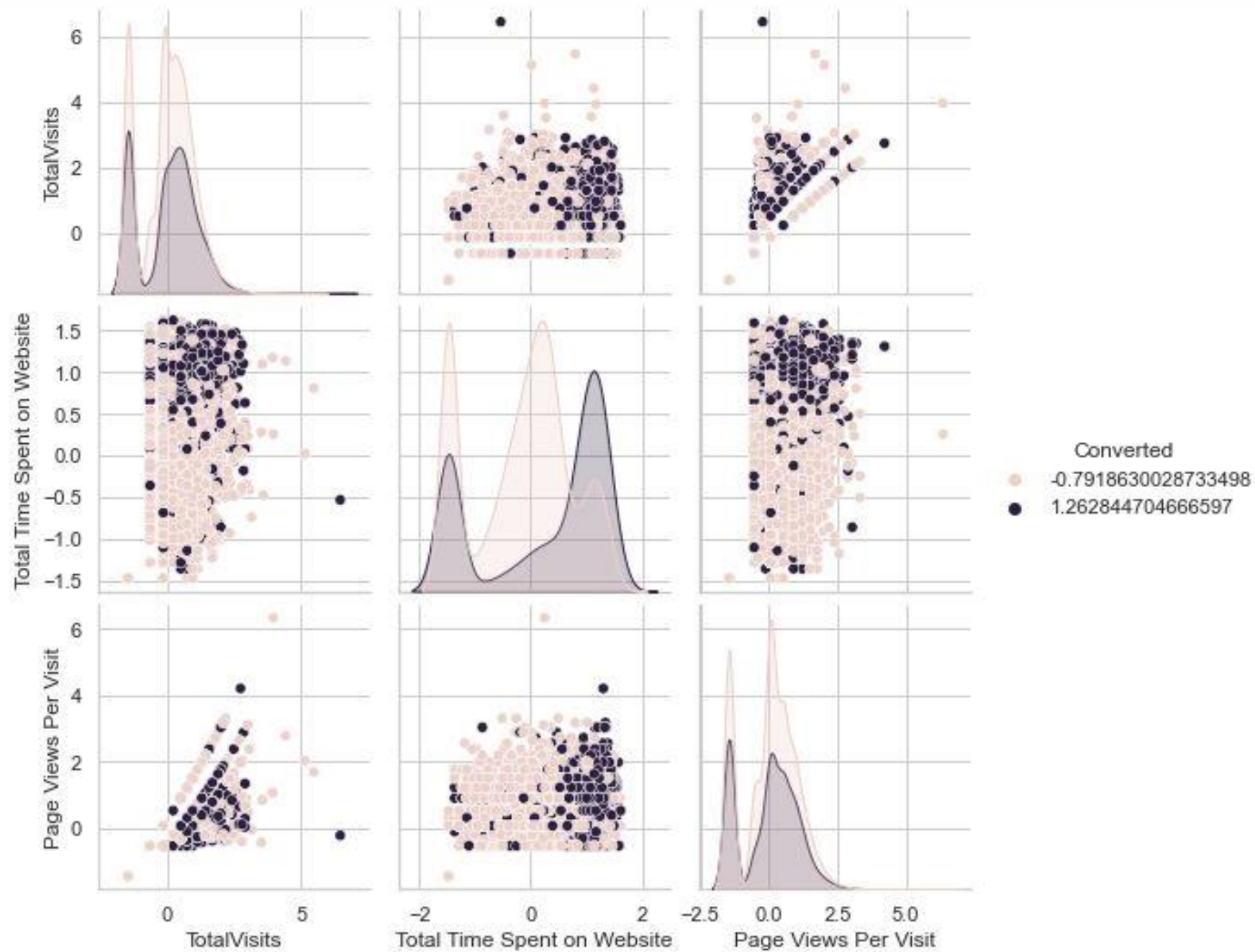
Correlation between all Variables

MODEL BUILDING & EVALUATION

- **Models Used:** Logistic Regression, Decision Trees, Random Forest, etc.
- **Performance Metrics:** Accuracy, Precision, Recall, F1-score, AUC-ROC.
- **Best Model:** Recursive Feature Elimination.

KEY BUSINESS INSIGHTS

- High-converting leads come from [Top Channels].
- Customer engagement level is a strong predictor of conversion.
- Optimized marketing strategies can improve conversion rates by X%.

TOP IMPACTINGFACTORS FOR LEAD GENERATION

	Features	VIF
9	What is your current occupation_Unemployed	2.82
1	Total Time Spent on Website	2.00
0	TotalVisits	1.54
7	Last Activity_SMS Sent	1.51
2	Lead Origin_Lead Add Form	1.45
3	Lead Source_Olark Chat	1.33
4	Lead Source_Welingak Website	1.30
5	Do Not Email_Yes	1.08
8	What is your current occupation_Student	1.06
6	Last Activity_Had a Phone Conversation	1.01
10	Last Notable Activity_Unreachable	1.01

ROC
(Area under curve shown as 0.86)

Receiver operating characteristic example
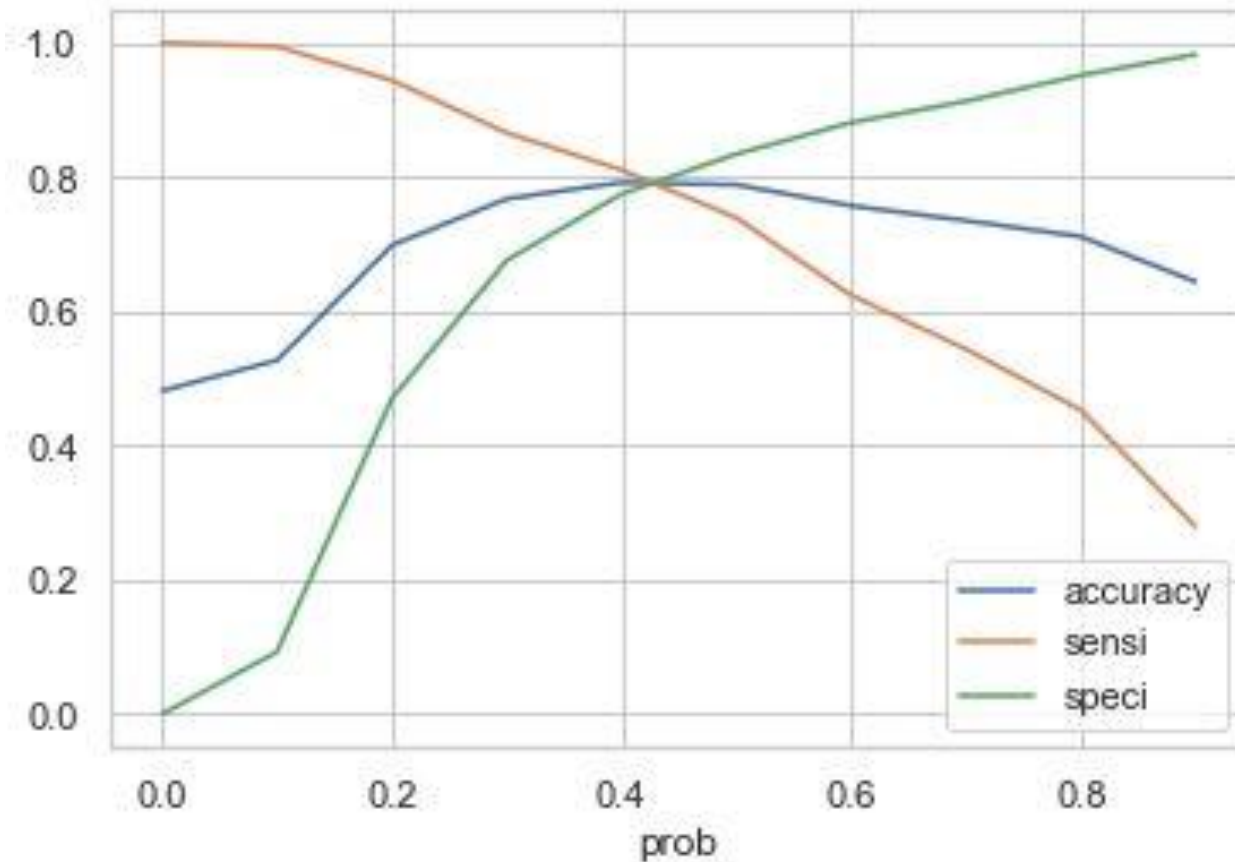
ROC curve (area = 0.86)
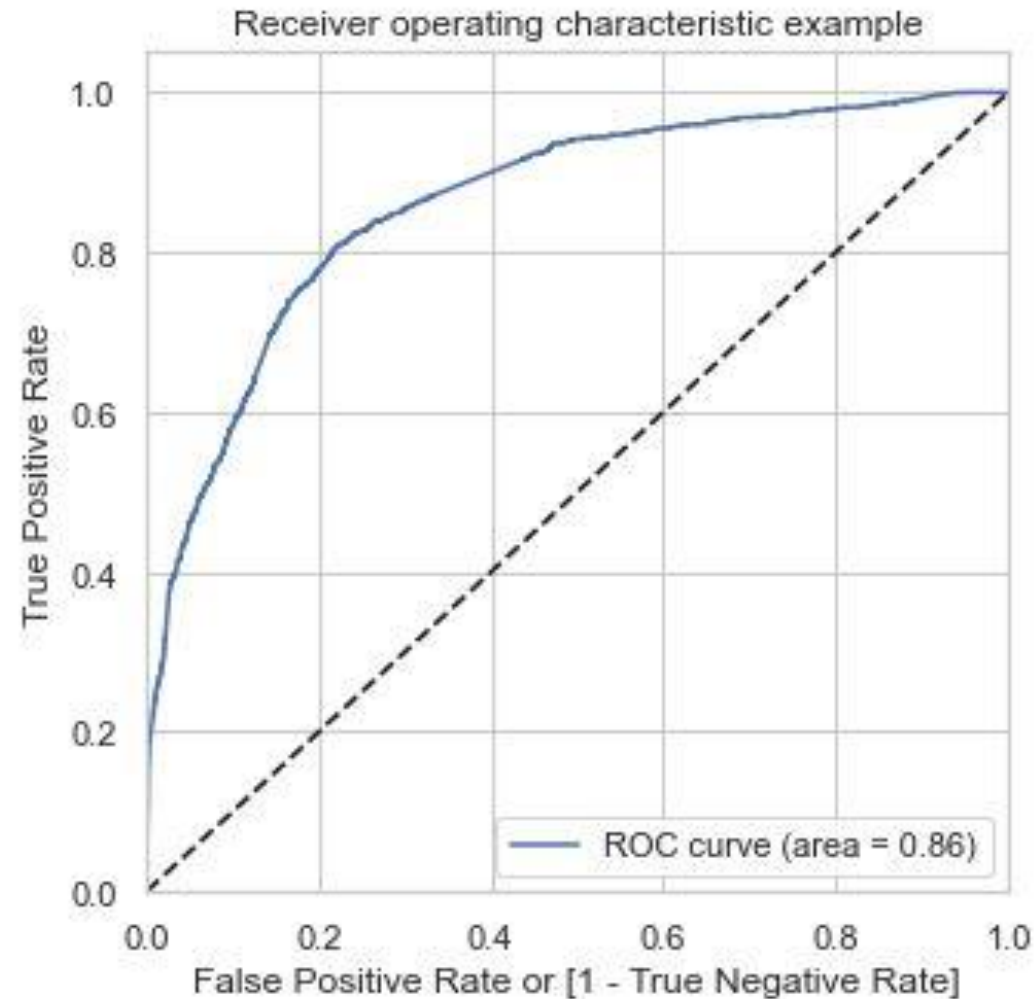
Optimal Cut-Off
(Cut-off came out to be 0.42)

RECOMMENDATIONS & NEXT STEPS

- Focus marketing budget on high-converting sources.
- Personalize engagement strategies based on lead behavior.
- Continuous monitoring & model retraining for improved accuracy.

CONCLUSION

- Data-driven lead scoring can significantly boost sales efficiency.
- Implementing the insights can drive higher conversion rates.
- Future scope: AI-driven automation for dynamic lead prioritization.

THANK YOU !