LEAD SCORING

CASE STUDY

BY ADARSH

PROBLEM STATEMENT

- An organization called X Education offers online training for business professionals. The company reviews its courses on a number of well-known websites, including Google.

- The most promising leads that can become paying clients should be chosen by X Education. Even while the business generates a large number of leads, only a small percentage of those leads end up becoming paying customers.

- Numerous channels, including email, website adverts, Google searches, etc., are used to generate leads. By engaging leads who are known to be interested in taking the course.
- The organization has experienced a 30% conversion rate for the entire process of converting leads into clients. Implementing lead-generating qualities is ineffective at assisting conversions.

BUSINESS GOAL

- The business needs a model to be created for choosing the most promising prospects.

- Each lead will be assigned a lead score that will show how promising they may be. The lead's likelihood of conversion increases with lead score; the smaller the lead score, the lower the likelihood of conversion

- The lead conversion rate in the model should be at least 80%.

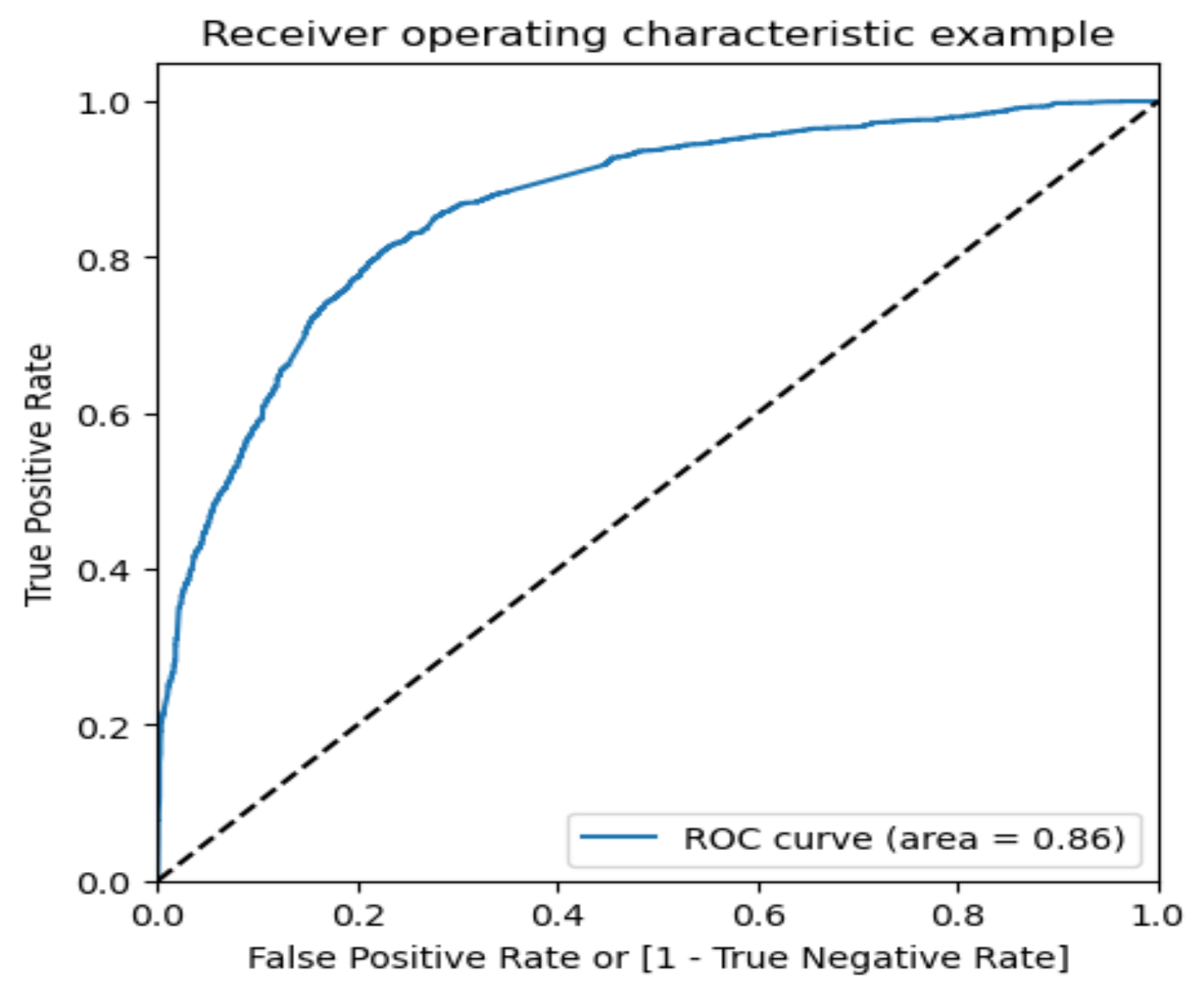
STRATEGY

- Upload data Clean up and get the data obtained ready for more analysis.
- Exploratory data analysis to identify the conversion-enhancing characteristics
- Scaling capabilities
- Get the data ready for model construction.
- Create a model of logistic regression.
- Give each lead a lead score.Model testing on a train set
- Model evaluation using various metrics and measurements.
- Model testing on a test set.
- Check the model's accuracy and other performance indicators.

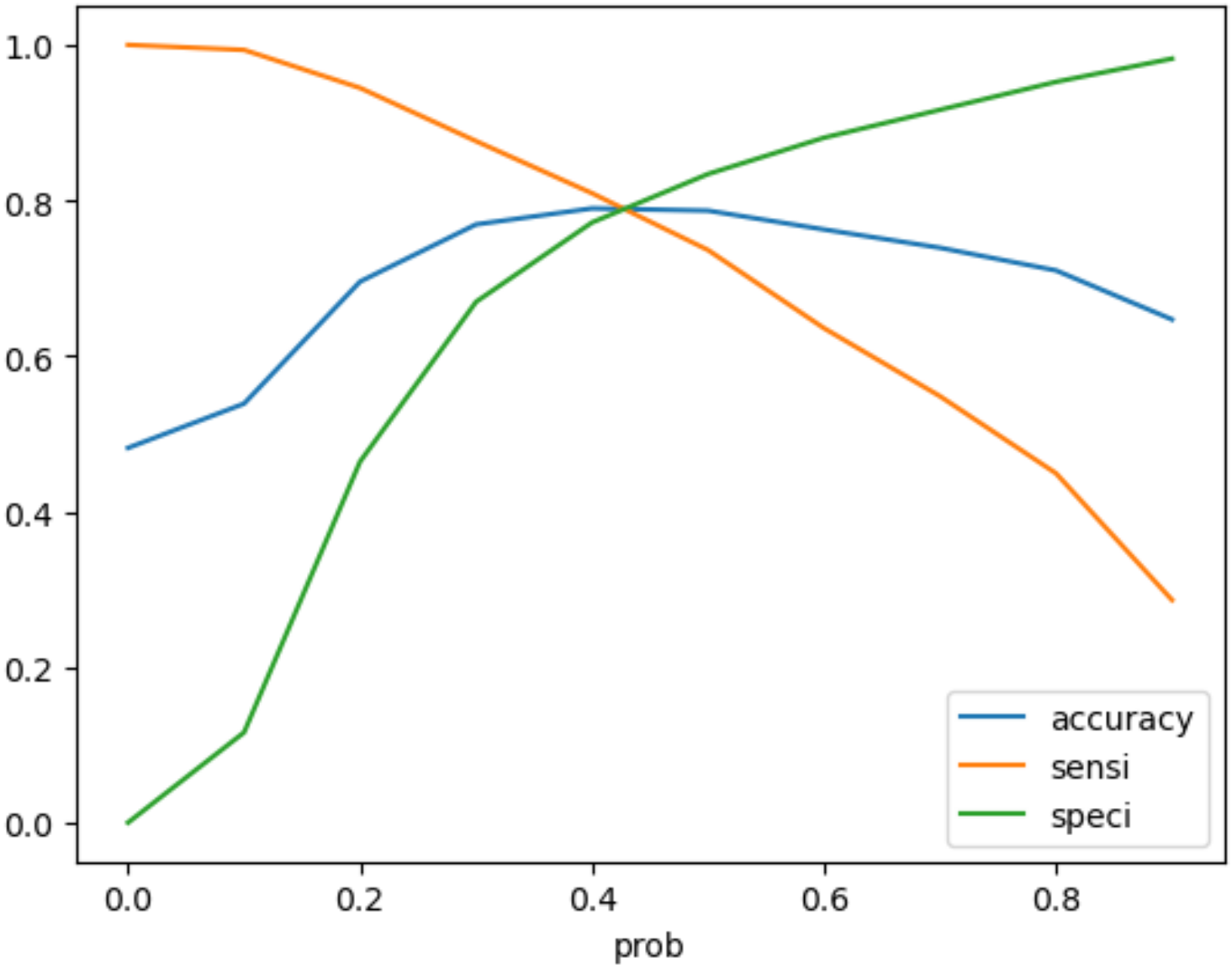
MODEL BUILDING

- Dividing a test set into a train set.
- Variables of scale in the train set.
- Create the initial model.
- RFE can be used to get rid of irrelevant variables.
- Build the next iteration.
- Remove variables with high p-values.
- Verify the VIF value for each of the current columns. Using a train set, predict.
- Analyze precision and other metrics with the test set, predict.
- Analysis of test predictions' precision and recall.

ROC CURVE


Receiver operating characteristic example
True Positive Rate
False Positive Rate or [1 - True Negative Rate]
0.0
0.2
0.4
0.6
0.8
1.0
ROC curve (area = 0.86)

MODEL EVALUATION



ACCURACY SENSTIVITY AND SPECFICITY

- 76.91 % ACCURACY
- 87.62 % SENSITIVITY
- 66.95 % SPECIFICITY

CONCLCUSION

- People that spend more time than usual are likely to convert, therefore reaching out to them and targeting them can be beneficial.
- Lead conversion can be significantly impacted by SMS messaging.
- Submissions to landing pages can assist in identifying new leads.
- High conversion rates exist in the administration of marketing and human resources. These specialties can produce promising leads.
- References and incentives for recommending leads can help increase conversion rates.
- A significant lead conversion rate has been observed for alert messages or information.
- Model for Logistic Regression:-

- The model exhibits good accuracy of around 81%.Accuracy.
- Sensitivity, Specificity Measures, and Precision, Recall Curves were used to choose the threshold.76% sensitivity and 83% specificity are displayed by the model.
- Overall model identifies the most promising leads and accurate