

Lead Score Analysis and Recommendations

Lead Score Analysis: Evaluating and assigning scores to potential sales leads based on diverse criteria and behaviors.

Objective: Prioritize and identify leads with the highest likelihood of conversion for targeted efforts.

Dataset Overview

- **Prospect ID:** Unique customer identifier.
- **Lead Number:** Assigned lead number.
- **Lead Origin:** Identification method (API, Landing Page Submission, etc.).
- **Lead Source:** Source of the lead (Google, Organic Search, Olark Chat, etc.).
- **Do Not Email/Do Not Call:** Customer's preference for communication.
- **Converted:** Target variable indicating successful conversion.
- **TotalVisits, Total Time Spent on Website, Page Views Per Visit:** Website interaction metrics.
- **Last Activity:** Last customer activity (Email Opened, Olark Chat Conversation, etc.).
- **Country, Specialization, How did you hear about X Education, etc.:** Customer information.

Steps for Lead Score Analysis Project

Define Objectives: Clearly outline project objectives.

Data Gathering: Collect comprehensive dataset on leads.

Data Inspection: Understand dataset structure, features, and initial patterns.

Data Cleaning: Address missing values, outliers, and preprocess data.

EDA: Visualize data to gain insights into distribution and relationships.

Feature Selection: Identify key features impacting lead scores.

Lead Scoring Model: Choose suitable model or methodology.

Data Preprocessing: Encode categorical variables, scale data, and split into sets.

Model Training: Train lead scoring model on prepared dataset.

Model Evaluation: Assess model performance using metrics.

Adjust and Optimize: Fine-tune model based on evaluation results.

Interpret Results: Analyze factors influencing lead scores for actionable insights.

Documentation: Document entire process, including code and parameters.

Communication: Communicate findings and insights to stakeholders.

Iterate and Improve: Continuous improvement based on feedback or new information.

Implementation: Integrate lead scoring system into workflow.

Monitoring: Regularly monitor and update lead scoring model.

Feedback Loop: Establish feedback loop for continuous improvement.

Scaling Data

- from **sklearn.preprocessing** import **StandardScaler**
- scaler = **StandardScaler()**
- X_train[['TotalVisits','Page Views Per Visit','Asymmetrique Activity Score','Asymmetrique Profile Score']] = scaler.fit_transform(X_train[['TotalVisits','Page Views Per Visit','Asymmetrique Activity Score','Asymmetrique Profile Score']])
- X_train.head()
- **Importance:** Ensures all features contribute equally to model training.
- **Methods:** *Min-Max Scaling*, Standard Scaling, Robust Scaling.

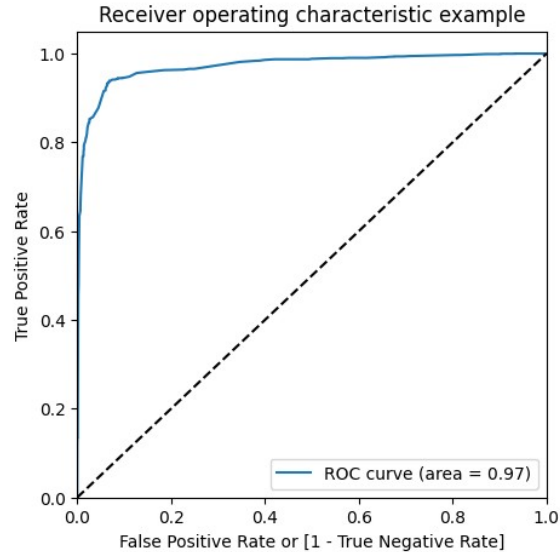
Feature Analysis and Decision

Decision: Drop '*Tags_Unknown*' column based on P-value and VIF analysis.

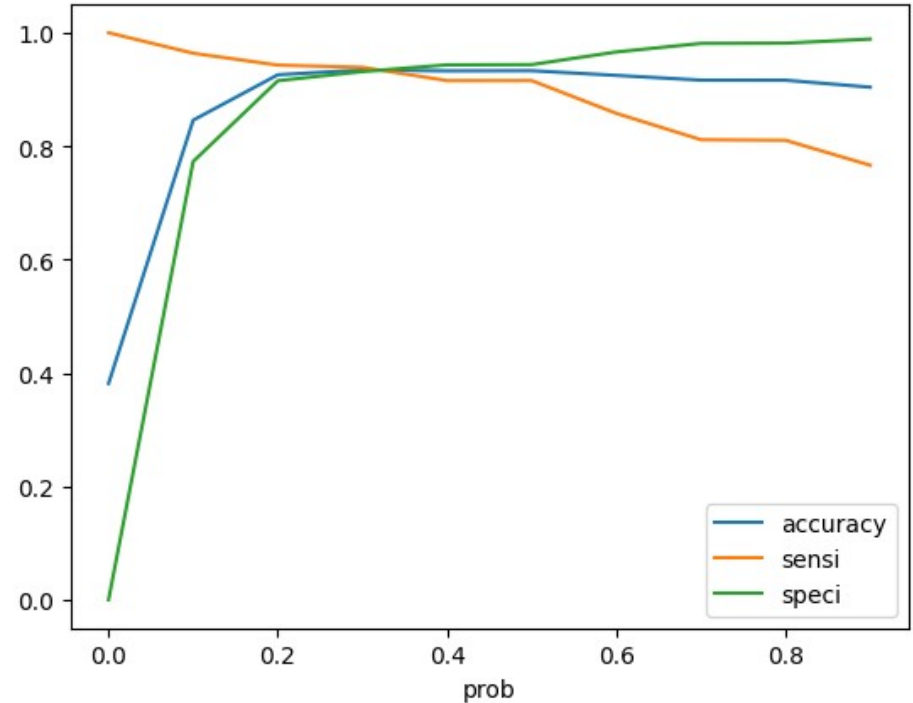
Reasons: Insignificant contribution, potential multicollinearity.

- `X_train[col1].drop('Tags_Unknown', axis=1,inplace=True)`

ROC Curve Analysis



- **Tool:** Utilized ROC curve for model performance evaluation.
- **Optimal Cutoff:** Precision and recall trade-off.



Results

Train Data:

- Accuracy: 93.3%
- Sensitivity: 92.6%
- Specificity: 93.7%

Test Data:

- Accuracy: 92.9%
- Sensitivity: 91.9%
- Specificity: 93.6%

Business Recommendations

Tags Recommendations:

- Focus on **"Lost to EINS"** leads.
- Engage with **"Closed by Horizzon"** and **"Will revert after reading the email"** leads.
- Address and monitor **"Unknown"** tag leads.

Last Notable Activity:

- Focus on leads with **"SMS Sent"** as the last notable activity.
- Address **"Unsubscribed"** leads.

Lead Source:

- Pay attention to leads from **"Welingak Website."**

Other Features:

- Higher **"Asymmetrique Activity Score"** leads are more likely to convert.
- Engage with **"Olark Chat Conversation"** leads.
- Evaluate impact of **"Do Not Email"** on lead conversion.
- Address leads with **"Worst"** Lead Quality.
- Verify and correct contact details for **"invalid number"** and **"switched off"** leads.
- Address leads with **"Unknown"** lead profiles.