

Summary Report:

Steps:

- Firstly , imported necessary libraries like NumPy and scikit-learn.
- Loaded and preprocess the dataset.
- Cleaned the dataset by removing or imputing missing values.
- Exploratory data analysis to find the important attributes.
- Splited the data into training and testing sets.
- Performed feature scaling if required.
- Created a logistic regression model using the LogisticRegression class.
- Trained the model on the training data using the fit() method.
- Made predictions on the test set.
- Evaluated the model's performance using metrics such as accuracy, precision, and recall.
- Additionally, visualize results using appropriate plots.

Results:

Train Data:

- Accuracy : 81.0 %
- Sensitivity : 81.7 %
- Specificity : 80.6 %

Test Data:

- Accuracy : 80.4 %
- Sensitivity : 80.4 %
- Specificity : 80.5 %
-

Thus, we have achieved our goal of getting a ballpark of the target lead conversion rate to be around 80% . The Model seems to predict the Conversion Rate very well and we should be able to give the CEO confidence in making good calls based on this model to get a higher lead conversion rate of 80%.

The customers which should be contacted are the customers whose "Lead Score" is equal to or greater than 85. They can be termed as 'Hot Leads'. So there are 368 leads which can be contacted and have a high chance of getting converted.

Important Features of Final Model:

Lead Source_Welingak Website
Lead Source_Reference
What is your current occupation_Working Professional
Last Activity_Other_Activity
Last Activity_SMS Sent
Total Time Spent on Website
Lead Source_Olark Chat
const
Last Notable Activity_Modified
Last Activity_Olark Chat Conversation
Lead Origin_Landing Page Submission
Specialization_Others
Do Not Email

Recommendations:

- Prioritize making calls to leads from "Welingak Websites" and "Reference" sources, as they exhibit a higher conversion likelihood.
- Target calls towards "working professionals" among leads, as they demonstrate a higher likelihood of conversion.
- Focus on leads who spend more time on the websites, as they are more likely to convert.
- Prioritize calls to leads from the "Olark Chat" source, given their higher conversion potential.
- Target leads whose last activity was "SMS Sent," as they are more likely to convert.
- Avoid making calls to leads with the last activity as "Olark Chat Conversation," as they show lower conversion likelihood.
- Refrain from calling leads from the "Landing Page Submission" source, as their conversion likelihood is lower.
- Avoid contacting leads with Specialization marked as "Others," as they are less likely to convert.
- Do not make calls to leads who have selected "Do not Email" as "yes," as their conversion likelihood is lower.