Summary Report

I have made this Lead Score Prediction Model using Logistic Regression which is a supervised learning algorithm.

📌 Introduction:
Logistic regression, in contrast to linear regression, is utilized when the objective is to predict categorical outcomes, typically binary classification problems. Instead of predicting continuous values, logistic regression models the probability that an instance belongs to a particular class based on the relationship between the independent variables and the binary outcome.

📌 Objective:
A machine learning model is to be proposed to predict a Lead Score based on data related to this.

📌 Data:
The dataset consists of 9240 entries across 37 attributes, with the target column labeled as 'Converted'.

📌 Steps followed are:

🔑 Applied data preprocessing and preparation techniques in order to obtain clean data which includes following steps:

1. Importing and understanding of data
2. Missing value check
3. Correlation Check
4. Data Reduction
5. Data Cleaning/Wrangling
6. Feature Engineering
7. Univariate Analysis
8. Bivariate Analysis

🔑 Built machine learning model to be able to predict Lead Score based on the features using Logistic regression includes following steps:

1. Splitting data into training and testing data
2. Feature Scaling
3. Feature Selection
4. Building and training the model
5. Variance Inflation Factor
6. Making predictions from the model
7. Testing the performance of the model
a. ROC curve
b. Accuracy Score
c. Confusion Metrics
d. Precision and Recall

📌 Model Evaluation:
Logistic regression performed well giving a testing accuracy of almost 85.44 % and Precision of 86.93 % and Recall of 76.88 %