Summary Report

Steps Performed

▶ Data Import

    ▶ Reading and Visualizing the Data

▶ Data Preparation and Cleaning

    ▶ Treating Missing Values (Dropping columns with high missing percentage)

    ▶ Removing Highly Skewed Columns

    ▶ Treating Outliers

▶ Data Modelling

    ▶ Train Test Split

    ▶ Scaling of data

    ▶ Iterative Modelling using coarse tuning (RFE + Manual Tuning)

    ▶ Modelling using Logistic Regression's GLM

    ▶ Checking key parameters like Accuracy, Sensitivity, Specificity, ROC Curve, Optimal Cutoff, Precision and Recall

▶ Prediction using the derived model

    ▶ Running the model on Test Data Set

    ▶ Checking the key parameters whether model performs good on unseen data

Key Parameters Decoded:

Parameters	Train Data	Test Data
Accuracy	92.07%	91.78%
Sensitivity	93.12%	93.64%
Specificity	91.42%	90.64%
Precision	94.30%	86.03%

Feature Importance:

```
Tags_Closed_by_Horizzon                                    7.01
Tags_Lost_to_EINS                                          5.95
Lead_Source_Welingak_Website                               4.39
Tags_Will_revert_after_reading_the_email                   3.78
What_is_your_current_occupation_Working_Professional       2.15
Last_Activity_SMS_Sent                                     2.06
What_is_your_current_occupation_Unemployed                 1.80
Last_Notable_Activity_Modified                            -1.63
Tags_Graduation_in_progress                               -1.90
Tags_opp_hangup                                           -2.21
Tags_Interested__in_full_time_MBA                         -2.38
Tags_Interested_in_other_courses                          -2.61
Tags_Diploma_holder_(Not_Eligible)                        -3.21
Tags_Already_a_student                                    -3.78
Tags_Ringing                                              -4.25
Tags_switched_off                                         -4.64
Tags_invalid_number                                       -4.84
dtype: float64
```