

# BUSINESS SUMMARY REPORT

## **Data Preparation, Modelling and Evaluation:**

The data is imported to a notebook, handled outliers, imputed missing values, and removed columns with missing data and redundancy to prepare the data for model building.

Dummies were created for the categorical columns, scaled the features and split the data to train and test datasets, built the model using logistic regression, removed statistically unwanted variables, and evaluated the model performance.

The model evaluation metrics are as shown:

Metric/Data	Training Data	Test Data
Accuracy	0.91	0.9
Precision	0.87	0.86
Recall	0.91	0.9
F1 Score	0.89	0.88
Sensitivity	0.91	0.91
Specificity	0.91	0.9
ROC AUC	0.96	

Based on the above metrics, it is evident that the model is reliable, robust and accurate about the predictions.

These are the important features of the model:

- Total Time Spent on Website
- Lead Source\_Welingak Website
- Occupation\_Missing
- Tags\_Busy
- Tags\_Closed by Horizon
- Tags\_Lost to EINS
- Tags\_Ringing
- Tags\_Will revert after reading the email
- Tags\_switched off
- Last Notable Activity\_Olark Chat Conversation
- Last Notable Activity\_SMS Sent

## **Insights and Recommendation:**

As seen from the model's important features, it is necessary to assign a tag to each of the leads. The leads generated by opening the Welingak website and redirected to X Education are most likely to get converted.

Monitor people who spent more time on the X Education website and contact them as they are likely to get converted. We can increase the ads on the Welingak website to increase the number of redirections and also to increase the view time of the website. Also, monitor people who are not joining competitors as they would most likely get converted.

There is also a need to strategize whom to call for lead conversion, such as avoiding repetitive calls to people whose phones were switched off or keep on ringing during a call, and avoiding contacting people who haven't mentioned their occupation.