**BUSINESS SUMMARY REPORT**

* With the data we had, we imported to a notebook, handled outliers, imputed missing values, removed columns with missing data and redundancy so as to prepare the data for model building.
* We have created dummies for various categorical columns, scaled the features and split to train and test datasets, built the model using logistic regression, removed statistically unwanted variables and evaluated the model performance.
* With the accuracy score of 92%, Precision of 93%, Recall of 86% and F1 score of 89%, we can say that this model performs better on the test datasets.
* These are the features that are built by the model:
* Total Time Spent on Website
* Lead Source\_Welingak Website
* Occupation\_Missing
* Tags\_Busy
* Tags\_Closed by Horizzon
* 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

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

* Monitor people who spent more time on 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 no. of redirections and also to increase the view time of the website.
* There is also need to strategize whom to call for lead conversion, such as avoiding repetitive calls to people whose phones were switch off or keeps on ringing during a call, avoid contacting people that haven’t mentioned their occupation. Also monitor people who are not joining competitors as they would most likely to get converted.