Summary

This analysis is to find ways to get more industry professionals to
join their courses of X Education. Basic data about how the potential customers visit the site,
and time they spend, how they reached the site and the conversion rate provided.

The following are the steps used:

1. Cleaning data:

The data was partially clean except for a few null values and the option select had to
be replaced with a null value since it did not give us much information. Few of the
null values were changed to 'not provided' to not lose much data.

2. EDA:

A quick EDA was done to check the condition of our data. It was found that a lot of
elements in the categorical variables were irrelevant. The numeric values seem good
and no outliers were found.

3. Dummy Variables:

The dummy variables were created and later the dummies with 'not provided 'elements were
removed. For numeric values we used the MinMaxScaler.

4. Train-Test split:

The split was done at 70% and 30% for train and test data respectively.

5. Model Building:

Firstly, RFE was done to attain the top 15 relevant variables. Later the rest of the
variables were removed manually depending on the VIF values and p-value.

6. Model Evaluation:

A confusion matrix was made. Later, the optimum cut off value (using ROC curve)
was used to find the accuracy, sensitivity and specificity which came to be around
70% each.

7. Prediction:

Prediction was done on the test data frame and with an optimum cut off as 0.3 with
accuracy, sensitivity, and specificity of 70%.

8. Precision – Recall:

This method was also used to recheck and a cut off 0.3 was found with Precision
around 66% and recall around 73% on the test data frame.

It was found that the variables that mattered the most in the potential buyers are.

- The total time spends on the Website.
- Total number of visits.
- When the lead source was:
  - Google
  - Direct traffic
  - Organic search
  - Welingak website
- When the last activity was:
  - SMS
  - Olark chat conversation
- When the lead origin is Lead add format.
- When their current occupation is as a working professional.