

# SUMMARY

Analysis done for X Education to find ways to get maximum professionals to join their course. The data gave us a lot of information about how the potential customer visits the site, time spent on the site and how they reached the site and the conversion rate.

We used the Following Steps:

1. Data Preparation and Cleaning: Some rows having high percentage of Null values were dropped and the 'select' option had to be replaced with a null value since it did not give us any information. Few of the Null Values were changed to 'not provided' so as to not lose much of the data.
2. EDA: It was found that a lot of elements in the categorical variables were irrelevant. Numeric values seemed good and no outliers were found.
3. Dummy Variables: Dummy variables were created and for Numeric Variables we used MinMax Scaler.
4. Train – Test Split: We split the data in 70-30 Ratio.
5. Model Building: We used RFE to select the top 15 Variables. Then we manually removed the variables based on p values and VIF scores.
6. Model Evaluation: A confusion matrix was made and then we used the optimum cutoff Value(ROC Curve) to find the accuracy, sensitivity and specificity which was around 80%
7. Prediction: Prediction was done on the Test Data frame and with an optimum cutoff as 0.35 with accuracy, sensitivity and specificity of 80%
8. Precision –Recall: This method was used to recheck and a cut off of 0.41 was found with Precision around 73% and recall around 75% on the Test set

It was found that the variables that mattered the most in the potential buyers are (In descending order) :

1. The total time spend on the Website.
2. Total number of visits.
3. When the lead source was:
  - a. Google
  - b. Direct traffic
  - c. Organic search
  - d. Welingak website
4. When the last activity was:
  - a. SMS
  - b. Olark chat conversation
5. When the lead origin is Lead add format.
6. When their current occupation is as a working professional.  
Keeping these in mind the X Education can flourish as they have a very high chance to get almost all the potential buyers to change their mind and buy their courses.

