

Lead score assignment

Problem statement

- An education company having online courses to industry professionals. This education company having lots of leads, but those leads are not getting converted .
- Education company wants to factors , and each customer conversation rate so they can keep more efforts on those customers and the conversation process to more feasible for the company and time constrain also reduces.

Step-2 and step-3

- Inspecting the Data frame.
- Knowing about the data contains 9240 rows and 37 columns.
- Checking the missing values count and handling the missing values.
- There is a 'select' value in some columns considering as missing column changing all the values to null.
- Deleting the columns having more than 40% of missing values and dropping the unwanted columns.
- Handling missing values using different techniques and creating dummy values for the categorical variables.

Step-4 and step-5

- Test train split
- Y variable is having target variable converted.
- Using minmaxscaler converting numeric variables to standard values.
- Checking the variables.
- Next step is model building


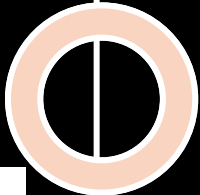
Step-6

- Model building
- Using logistic regression, we are building our model
- Binary classification, keeping 10 variables for model building pvalues should be less than 0.05
- Deleting occupation_Housewife having higher p values deleting and re running the model
- Deleting Last Notable Activity_Had a Phone Conversation having higher p values deleting and re running the model


Step-7

- Feature Selection Using RFE
- REF values is important feature while selecting the which variable are important to keep in the model.
- Values should be less than 5.
- Our model is having all the variables having less than value of 5.

Accuracy, Recall, F1 Score, confusion matrix



0.7941851568477429				
	precision	recall	f1-score	support
0	0.78	0.83	0.81	674
1	0.81	0.75	0.78	633
accuracy			0.79	1307
macro avg	0.80	0.79	0.79	1307
weighted avg	0.80	0.79	0.79	1307



```
[[561 113]
 [156 477]]
```