**BUSINESS PROBLEM:** Client wants to predict successful all “unique Ids” from unique Id column who have high chance of getting 1 as “Target” column value.

# Input Two Datasets:

1. Training – train.csv
2. Testing – test.csv

# Train/Test Data Schema (Data Description):

1. 'Target' Column in the training set is Class Label. “-1” signifies missing value in the columns.
2. TOP: Time on Page (in seconds)
3. Exits: Number of times unique Id has exited the page
4. Binary\_Var: contains binary values
5. Metric\_Var: contains continuous values
6. Unique Id: The primary key/unique identifier
7. Page1\_Visited: Whether unique id has visited the page or not

# Submissions:

1. Detailed code including Complete Preprocessing & Feature Engineering Steps
2. Output file in .csv format which consists of 2 columns “Unique Id” & “Predict\_Target”
3. You need to share the following evaluation metrics
   1. Confusion Matrix
   2. F1 Score
   3. Accuracy Score
   4. AUC\_ROC\_Score
4. The code must be clearly commented. Visualizations are optional