# Logistic regression classifier for a classification problem

The dataset for this assignment has data on whether customers purchased organic products during a campaign. The data dictionary is given below.

Table

Description automatically generated

To keep the assignment answers consistent, the training and test datasets have been created for you and are named “org.train.csv” and “org.test.csv” respectively. Both datasets can be downloaded from Canvas.

Your primary task is to build logistic regression model(s) to predict if a given customer will purchase organic products or not (TargetBuy is the response variable). Each of the tasks is to be executed in R-Studio and your answers should be written in a Word document (with the appropriate question/sub-question number).

After completing the assignment, save the R code and your Word document as “Logistic\_Regression\_**username**”, where **username** is **your Ninernet username** (e.g., Logistic\_Regression\_csubrama.R and Logistic\_Regression\_csubrama.docx). Upload both documents through Canvas assignment.

1. Read the training and test datasets into R-Studio.
2. If there are any categorical variables in the datasets, convert them to factors (use as.factor() function).
3. Fit a logistic regression to your training dataset with all the independent variables, except ID.
4. Use the estimated model to predict the classes for the test dataset. Use the threshold of 0.5 probability to determine if the customer is predicted buy organics products (i.e., predicted TargetBuy = 1).
5. Show the confusion matrix for the predictions on the test data.
6. Report the following values:
   * True Positive
   * True Negative
   * False Positive
   * False Negative

Note: To ensure you got the right values, calculate the count of actual classes in your org.test and cross-reference with the above 4 values.

1. Report the following values:
   * Recall (sensitivity)
   * Specificity
   * Precision
   * F1 Score