ENGR 421 DASC 521
Homework 08: Modeling Credit Card Customer Behaviors

Deadline: January 3, 2020, 11:59 PM

In this homework, you will develop a machine learning solution in R, Matlab, or Python for six real-life classification problems from finance industry. Your machine learning algorithm needs to predict whether a credit card customer will take specific actions (six different scenarios named as TARGET_1, TARGET_2, TARGET_3, TARGET_4, TARGET_5, and TARGET_6) using the information given about each customer. Here are the steps you need to follow:

1. You are given three input data files, namely, hw08_training_data.csv, hw08_training_label.csv, and hw08_test_data.csv. The training and test sets contain 108,904 and 105,560 data instances, respectively, where each data instance has 225 features. You are also given a very simple solution strategy using a boosting classifier in the file named hw08_quick_and_dirty_solution.R.

2. Develop your own machine learning solution for these six problems. You are free to use any publicly available packages in R, Matlab, or Python. The predictive quality of your solutions will be evaluated in terms of AUROC (area under the receiver operating characteristics curve) values for these six problems on the test set.

3. Use the trained algorithms from the previous step to perform predictions for the test data set, which contains 105,560 customers for six problems. You are not given the correct labels for test instances. You need to predict the scores or posterior probabilities for positive class in each problem and to write these estimates into a file. For example, the strategy implemented in hw07_quick_and_dirty_solution.R file generates the estimates for the test set and writes these values into a file named as hw08_test_predictions.csv.

What to submit: You need to submit your source code in a single file (.R file if you are using R, .m file if you are using Matlab, or .py file if you are using Python), the estimated scores or posterior probabilities for positive class on the test set (hw08_test_predictions.csv), and a detailed report explaining your approach (.doc, .docx, or .pdf file). You will put these three files in a single zip file named as ***STUDENTID.zip***, where ***STUDENTID*** should be replaced with your 7-digit student number.

How to submit: Submit the zip file you created to Blackboard. Please follow the exact style mentioned and do not send a zip file named as ***STUDENTID.zip***. Submissions that do not follow these guidelines will not be graded.

Late submission policy: Late submissions will not be graded.

Cheating policy: Very similar submissions will not be graded.