CISC 520 Assignment 2

Late Summer 2017 Session
Due Date: September 12, 2017 by 11:55 P.M. EST

1 Description

In this assignment you will be implementing K-Nearest Neighbors (KNN) algorithm. Please review the video lectures and the lecture slides for the associated topic. Please implement your own algorithm and not use any available libraries.

You will be using KNN to predict whether the stock exchange index goes up or down. The training data consist of two columns determining how the stock exchange performed on a particular day between 2001 and 2004. The labels data provides the direction of the stock exchange for each day of the training data. The test data consists of two columns for the year 2005. You are supposed to obtain the labels of the test data and store them to a file with each line containing a label for the observation corresponding to the test file.

There are two labels for the market movement, 'Up' and 'Down'. Test your program with different number of neighbors, k. For each value of nearest neighbors (k=1,3,5,10) save your test results.

All the data files, train.csv, trainDirection.csv and test.csv are provided for you to download.

1.1 Grading Criteria

This assignment carries 10 points of your assignment grade. You will be graded on logic, comments, whether the code runs or not and the output. You are expected to do this work on your own. Please implement your code in one of the following languages, Java, C, C++, Python or R and turn in your source code files only and not the entire project, along with the test results.

1.2 Submission

You have the option to use 2 extra days to finish this assignment unless you have already used two extensions.

Upload all your source code files (.py, or .java, or .c or .c++) and the test results to Moodle. Please add a README file to include any instructions on how to successfully run your code if there are any specific steps.

Note: Please do not upload all the project files, just the source code. Submitting wrong files or in the wrong format will not be accepted nor will any re-submission be allowed for any such mistake.