UCLA Computer Science Department

CS 260 Machine Learning Research Project

You are asked to implement two machine learning algorithms using time series data. The first one is kNN (k should be an input, by default should be 1-NN), and the second one is a Neural Network (NN). The goal is to fully understand the underpinnings of these two algorithms. The input you receive for each subject/patient/participant will be two time series signals (example: blood pressure over 26 days and weight over 26 days to predict a healthy from an unhealthy patient, or x-axis and y-axis accelerometer signals for activity recognition).

To test your algorithms, you will perform leave one out cross validation. You will be asked to have each of the two algorithms (kNN and NN) output the following average measures of performance: accuracy, precision, recall, sensitivity, specificity, F-measure, ROC curves, confusion matrices.

You will be given a data set with N subjects, and each subject i is represented by two signals t1i and t2i.

Each data point will be associated with one of two labels (one will be a positive class and the other a negative class). For example: Cancer and No Cancer, or Walking and Running, or 1 and 0.

You will need to extract features (statistical features and other well-known features in time series analysis) from the time series data sets and attempt to distinguish between two class labels.

Uniqueness: You are then asked to add a unique contribution to the project. Either testing new and unique features on the time series data set (as discussed in class), or a unique machine learning algorithm. You may also test your new algorithm or features on a new data set to see how well it works (this is a bonus).

Submission: You will submit all your code and a clear explanation of how your code runs (this should be very clear). You will also submit a report that provides the output of the above two algorithms and the results of your unique

contribution. You will be tested on how well your algorithm works on a test set, your report as well as your contribution to this research project.

Honesty Measure: You are to work independently! You are not to use any sources from the internet (or copy code from somewhere). There will be no collaboration. The goal is for you to completely understand how these two machine learning algorithms using time series data.