ECE 4850/7650 Applied Computational Intelligence

Project 4

SVM Investigation

Fall 2018

Due Date: Thursday 15 November

1. Do Exercise 7 from Andrew Ng's course, but with the following changes:
http://openclassroom.stanford.edu/MainFolder/DocumentPage.php?course=MachineLearning&doc=exercises/ex7/ex7.html

   a. Remove the outlier from the positive training set.
   b. Remove the effect of regularization in the LIBSVM Matlab Interface, and apply the basic linear SVM to the resulting training set. Note: our Matlab SVM library on the machines in the lab is different than the SVM library Andrew uses, and if you choose to use the lab's version, then you will need to determine how to change the input data to become compatible with it.
   c. Plot the resulting decision boundary, and report the values of the optimal w and b.
   d. Use logistic regression on the same training set, and compare the resulting decision boundary with that obtained with SVM. Compare the training times also. Which algorithm is better?
   e. Find the center of gravity of the positive training examples and the center of gravity of the negative training examples. Find the midpoint of the line that joins them. Is this point on the decision boundary line you found in Step c?
   f. Suggest an algorithm for finding the slope and intercept given the midpoint of the line that joins the centers of gravity as found in Step e?
   g. Add 1000 positive training examples as follows: choosing the point clusterCenter = (40.0, 50.0) and generate 1000 neighbors by randomly permuting clusterCenter by +-2.0.
   h. Plot your new training set to verify you have created a cluster of positive training examples far away for the original cluster.
   i. Perform SVM on this new training set, and compare the resulting decision boundary with that obtained by Step c. Are they the same? Explain why they are the same or why they are not the same.

2. Prove mathematically or by an argument that the SVM algorithm will produce at least one positive class support vector and one negative class support vector.

SUBMISSION: Prepare and submit a report to the TA for marking.