ECE 7650 Applied Computational Intelligence

Project 3  
SVM Investigation

Fall 2024

Due Date: Nov 18, 1:30 PM

* 1. Given: training set DataSetv1.txt: Each training example (row) has two features and a class label (+1 or -1)
  2. Using Matlab’s SVM library ([fitcsvm](https://www.mathworks.com/help/stats/fitcsvm.html)), train an SVM (with no regularization) using the given (DataSetv1.txt) training data.  
     (Tutorial on Matlab SVM: <https://www.mathworks.com/discovery/support-vector-machine.html>)
  3. Plot the decision boundary, support vectors, margin, and all training examples. Your plot should look similar (not exactly the same) to that shown below:

A diagram of a line with text

Description automatically generated with medium confidence

* 1. Find the center of gravity (average) 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?
  2. Suggest a method for finding the slope and intercept, given the midpoint of the line that joins the centers of gravity, as found in Step d?
  3. 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.
  4. Plot your new training set to verify you have created a cluster of positive training examples far away for the original cluster.
  5. Perform SVM (with no regularization) 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.
  6. Do Step h again: this time perform SVM with a significant amount of regularization. Explain.

SUBMISSION: Submit your zipped solution folder to umlearn under Project 3.