# Homework 3
# CS 6375: Machine Learning

Spring 2012

Due date: Wednesday, March 21, midnight

## Perceptrons, Neural Networks and Support Vector Machines

In this homework, you will implement the Perceptron algorithm and compare it with WEKA implementations of Neural networks and Support Vector machines. The data sets that we will use for this homework are available on the class web page. As in homework 2, the classification task is spam/ham.

**25 points** Implement the perceptron algorithm (use the perceptron training rule and not the gradient descent rule). As discussed in class, if the data is not linearly separable, the algorithm will not converge. Therefore, your task here is to experiment with different values of number of iterations and the learning rate. Report the accuracy for 20 suitable combinations of number of iterations and the learning rate. Repeat your experiment by filtering out the stop words. Compare the accuracy of your perceptron implementation with that of Naive Bayes and Logistic Regression (implemented in Homework 2. If you did not implement Naive Bayes and/or Logistic regression in homework 2, use the implementations available in WEKA).

**50 points** Support Vector Machines and Neural networks in WEKA.

- Download WEKA `http://www.cs.waikato.ac.nz/ml/weka/`.
- Convert the spam/ham dataset into the ARFF format used by WEKA.
- Using the SVM implementation in WEKA, report the accuracy on the test set. Experiment with the following kernels: linear, polynomial and sigmoid.

– Using the Neural networks implementation in WEKA (called MultiLayered Perceptron in WEKA), report the accuracy of the test set. Experiment with different number of hidden layers and report your results.

**15 points** Problem 4.8 from Tom Mitchell's book.

**10 points** Problem 4.10 from Tom Mitchell's book.

# What to Turn in:

- Your code for the perceptron algorithm
- Your report describing the results of your experimental analysis. Make sure that you report your results on all the three data sets.
- Answers to Exercises 4.8 and 4.10 from Tom Mitchell's book. If you prefer, you can turn them in class on Wednesday, March 21.