Neeraj Asthana
nasthan2
2/22/2016
CS 498 Homework 3 Report

Problem 1 Part 1 – SVM
I used the SVMlight software to produce these results. I tried lambda values of .001, .01, .1, 1, and 10. I subsetted 90% of the training data to train the SVM model on and used 10% as a validation set to test different values of lambda. After checking the validation accuracies, lambda of .01 produces the best results (which as summarized below).

Set	Accuracy
Full Training Set	.5173103
Validation Set 1	.4969748
Validation Set 2	.5015751
Validation Set 3	.4987749

Problem 1 Part 2 – Naive Bayes
I used the Naive Bayes classifier I wrote for Homework 1 to produce these results. The results of this classifier are summarized below.

Set	Accuracy
Full Training Set	.5086102
Validation Set 1	.4974249
Validation Set 2	.4933747
Validation Set 3	.4946747

Problem 1 Part 3 – Random Forest
I used the Random Forest function inside of the R package called “randomForest” to produce these results. I tried lambda values of .001, .01, .1, 1, and 10. I subsetted 90% of the training data to train the SVM model on and used 10% as a validation set to test different values of lambda. After checking the validation accuracies, lambda of .01 produces the best results (which as summarized below).

Set	Accuracy
Full Training Set	1
Validation Set 1	.8063613
Validation Set 2	.8079341
Validation Set 3	.8053613

Problem 3 – Neural Network

I used a Neural Network as a classifier for my choice method. I only trained on the first 5000 examples since it was taking too long to classify the full training set. I used 2 hidden layers with 20 and 10 hidden nodes each. I the neuralnet function inside of the neuralnet package to accomplish this task. I believe I could have gotten much higher accuracies if I had used more training examples but I did not have the time to accomplish that.

Set	Accuracy
Full Training Set	0.6937338747
Validation Set 1	0.5838291915
Validation Set 2	0.5749787489
Validation Set 3	0.5764288214