Machine Learning CAP6610

Drew Gonsalves

HOMEWORK #4

**Question #1:**

**Varying epsilon:**

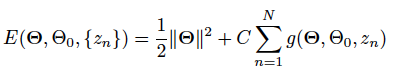
I varied epsilson from 1,0.1 and 0.00001 and saw no discernable difference in the misclassification errors. I did notice that the length of time to compute the linSVM increased as epsilon decreased.

**Varying C:**

I varied C from 1 to 20. I noticed that the linSVM took longer to converge to my convergence rate threshold as C increased.

**Varying threshold\_rate for finding minimum value of *z*:**

In order to find when I reached the MINIMUM of my MAJORIZED OBJECTIVE FUNCTION:



where:



I calculated the RATE OF CONVERGENCE of the MAJORIZED OBJECTIVE FUNCTION by taking the difference of the OLD CALCULATED MAJORIZED OBJECT FUNCTION VALUE and the NEW CALCULATED MAJORIZED OBJECTIVE FUNCTION VALUE. I compared this RATE OF CONVERGENCE against my threshold\_rate to determine if the minimum had been found.

As the difference of the Old and the New grew to be smaller than my threshold\_rate, I concluded the objective function had reached a MINIMUM. I noticed that as the threshold\_rate increased (ie “tighter bounds” or trying to get closer to the true “minimum”) I increased the LENGTH OF TIME OF CONVERGENCE.

**Increasing the size of the training set:**

As I increased the size of my training set from 0.1 to 0.5 I found the length of time to complete all trials and get the data increased. This is due to the increased number of training (N) and increased computation.

**Final parameters:**

Epsilon – 10^-6

Threshold\_Rate\_Convergence - 0.001

C-Varied but found about C=1.5 to be best

**RESULTS:**

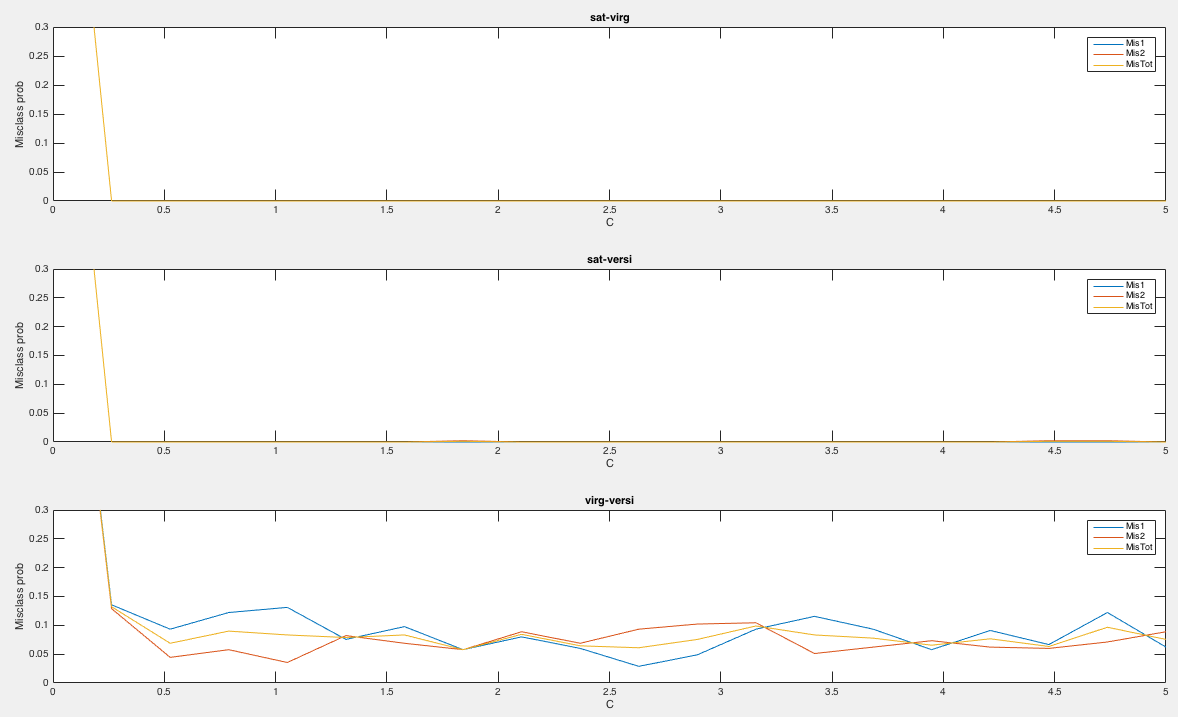
**FOR 10% TRAINING:**

**Best Parameters:**

Epsilon – 10^-6

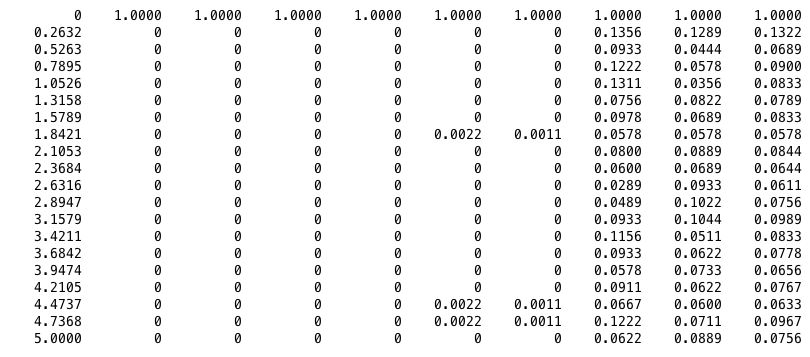
Threshold\_Rate\_Convergence - 0.001

C-~1.5 or between 1 and 5.



*Plot of all three pairs and misclassification errors*

C-Value Sat-Virg (C1/C2/Total) Sat-Versi (C1/C2/Total) Virg-Versi (C1/C2/Total)

**

**Discussion:**

I looked at C from 0 to 30 with only 0 to 5 shown on the graphs above. I noticed above 5 a low rising trend up of the misclassification error. In looking at the graph, and keeping GENERALIZATION in mind by choosing a low C which results in a larger margin, I choose a C=~1.5 due to the large misclassification errors below C=1. **Average misclassification error was about 6%.**

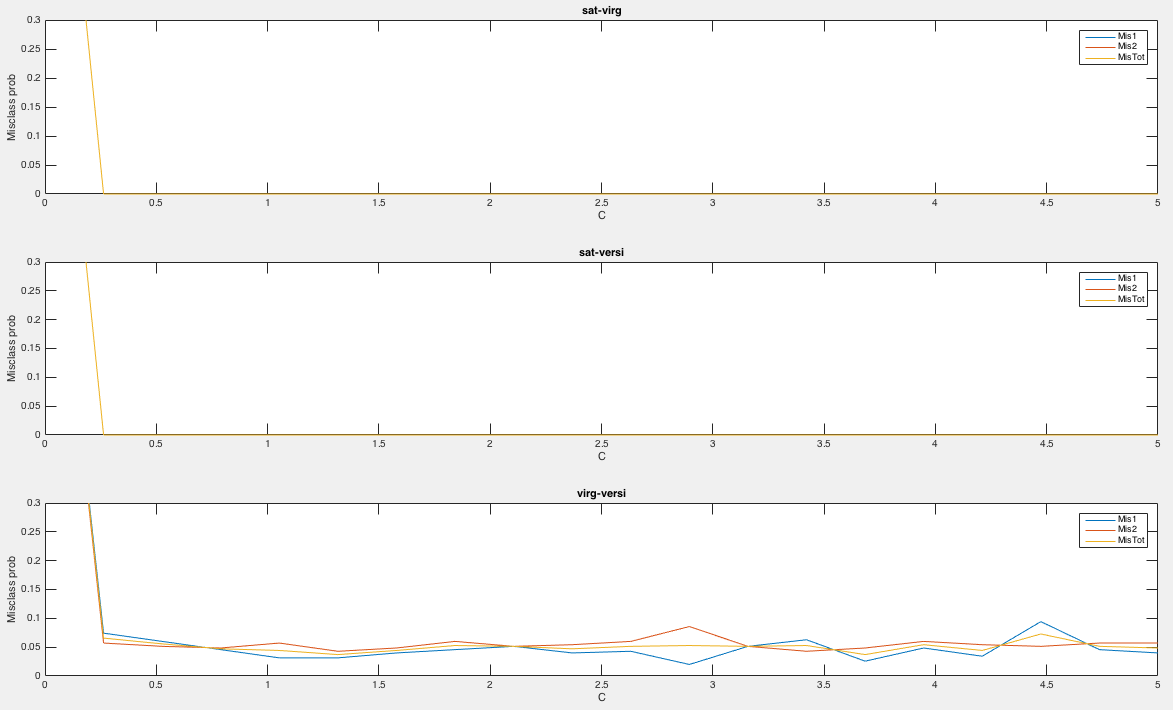
**FOR 30% TRAINING:**

**Best Parameters:**

Epsilon – 10^-6

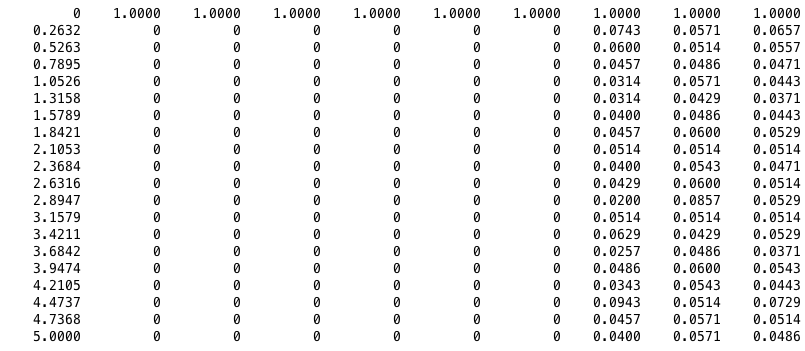
Threshold\_Rate\_Convergence - 0.001

C-~1.5 or between 1 and 5



*Plot of all three pairs and misclassification errors*

C-Value Sat-Virg (C1/C2/Total) Sat-Versi (C1/C2/Total) Virg-Versi (C1/C2/Total)

**

**Discussion:**

The same as for 10% training: I looked at C from 0 to 30 with only 0 to 5 shown on the graphs above. I noticed above 5 a low rising trend up of the misclassification error. In looking at the graph, and keeping GENERALIZATION in mind by choosing a low C which results in a larger margin, I choose a C=~1.5 due to the large misclassification errors below C=1. **Average misclassification error was about 4%.**

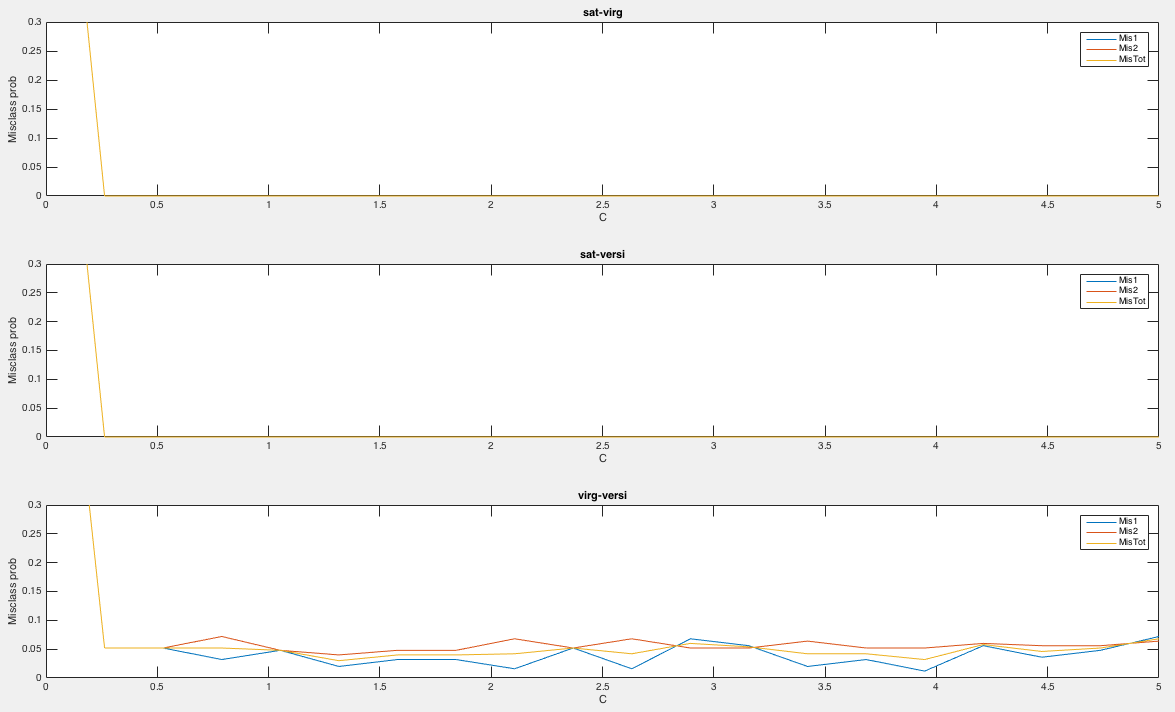
**FOR 50% TRAINING:**

**Best Parameters:**

Epsilon – 10^-6

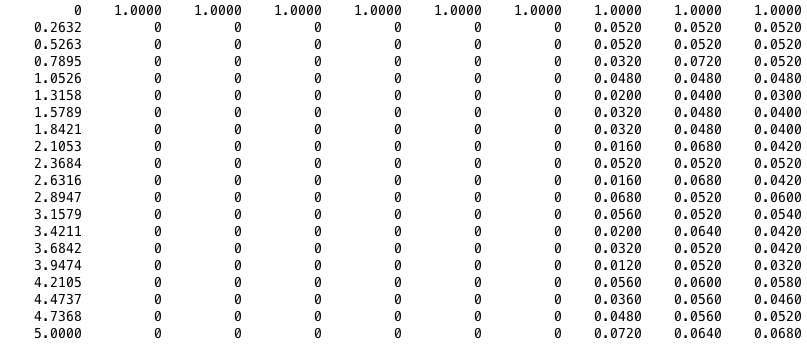
Threshold\_Rate\_Convergence - 0.001

C-~1.5 or between 1 and 5



*Plot of all three pairs and misclassification errors*

C-Value Sat-Virg (C1/C2/Total) Sat-Versi (C1/C2/Total) Virg-Versi (C1/C2/Total)

**

**Discussion:**

Same as for 10%/30%: I looked at C from 0 to 30 with only 0 to 5 shown on the graphs above. I noticed above 5 a low rising trend up of the misclassification error. In looking at the graph, and keeping GENERALIZATION in mind by choosing a low C which results in a larger margin, I choose a C=~1.5 due to the large misclassification errors below C=1. **Average misclassification error was about 6%.**