**Report with Detailed Steps**

My computer cannot open \*.data file, so I copied processed.cleveland.data to “libsvm/process-data/input-data.txt”

First of all the data from the website has missing values and the format of it was not right. I preprocessed data to make it right. I created a “fill-data.py” to fill all the missing data with the average of that attribute (cd ./process-data/). After filling all values, I wrote a script “format-data.py” match with SVM format. I also switched label to -1 and +1.

<label> <index1>:<value1> <index2>:<value2> … The file name is “good-data.txt”

-Then I use 9 fold cross validation by the command terminal “SVM-train -v 9 good.data.txt” and get the accuracy of 54.7%

I scaled the data by the command “svm-scale -l -1 -u 1 good-data.txt > good-data.txt.scale”

.Then I use a 9-fold validation for the scale file to test the accuracy by using command “svm-train -v 9 good-data.txt.scale” the result came out to 85.1%

I tried to find “C” and “Gamma” of the scale data to see if it comes up with the better result. I used “python grid.py -v 9 -log2c -5,5,1 -log2g -4,0,1 ../good-data.txt.scale” the “Gamma” and “C” came out to be 0.00625 and 2.0.

I trained data with “C” and “Gamma” by command “svm-train -c 2 -g 0.0625 good-data.txt.scale”. This will output model file “good-data.txt.scale.model”

I pick random 41 svms from the “good-data.txt” and put it into “test-data.txt” and scaled it with “svm-scale test-data.txt > test-data.txt.scale” this will output “test-data.txt.scale” file. I tested the accuracy again using command “svm-predict test-data.txt.scale good-data.txt.scale.model t.predict” The result came out as 87.8%. The result is pretty good to me.