**M5 Lab: Building a SVM Model in LibSVM**

### **Part I: Installing the Libraries**

Steps:

1. Download LIBSVM that is suitable for your machine from this link: [https://www.csie.ntu.edu.tw/~cjlin/libsvm/ (Links to an external site.)](https://www.csie.ntu.edu.tw/~cjlin/libsvm/)
2. [(Links to an external site.)](https://www.csie.ntu.edu.tw/~cjlin/libsvm/)Install the library on your machine.

Graphical user interface, text, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

* 1. For windows, just download the zip file, unzip it to a known directory. The unzipped directory contains a folder named windows. This is where all the svm training & testing programs are located. They are called svm-train.exe & svm-predict.ext
  2. For MAC, download the zip or the tar.gz file and unzip it. Open a terminal, change directory to go to the unzipped folder. In the unzipped folder, there's a Makefile. You need to compile this file by simply typing   
     make  
     in the terminal.  
     Once the compilation process is over, you are now ready to run the SVM model training & testing programs!

Graphical user interface, application, table

Description automatically generated

1. Every time you want to run the library you will need to open a terminal and go to the directory that contains the programs.

A picture containing application

Description automatically generated

### **Part II: Training an SVM Model with a linear kernel**

Open a terminal window, change directory to go to where the library was installed, and run the svm-train & svm-predict programs by following these instructions:

1. First, copy the dataset files into the same directory as the svm programs.

A screenshot of a computer

Description automatically generated

Graphical user interface, application, table

Description automatically generated

1. Second, run the training program to create an SVM classifier with a linear decision boundary:
   1. For windows, type this command:  
      **svm-train.exe -t 0 colon\_cancer\_train\_1.libsvm linear\_model.txt**
   2. For MAC, type this command:  
      **./svm-train -t 0 colon\_cancer\_train\_1.libsvm linear\_model.txt**

Text

Description automatically generated

1. Third, run the testing program to test your model in a set of new data points and classify them.
   1. For windows, type this command:  
      **svm-predict.exe colon\_cancer\_test\_1.libsvm linear\_model.txt out.txt**
   2. For MAC, type this command:  
      **./svm-predict colon\_cancer\_test\_1.libsvm linear\_model.txt out.txt**

Text

Description automatically generated

1. What is the obtained classification error (or accuracy)?

Accuracy =91.3043%

1. Take a screen shot of the string that you used to run the code, and obtained results.

Done.

### **Part III: Playing with other kernels**

Check the libsvm page for how to train SVM to learn non-linear decision boundaries such as a Polynomial or an RBF curve.

1. Build a non-linear SVM model with a polynomial kernel with degree 2 & test it on the test set.
2. Repeat with a polynomial kernel with degree 3

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Repeat with an RBF kernel with gamma = 0.01

Repeat with an RBF kernel with gamma = 0.1

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, email

Description automatically generated

What are the obtained testing errors?

1 - 0.74 = 0.26

Which SVM model gives the lowest testing error?

Linear model has the lowest test Errors (0.09).

How to do you explain that?

As we go to the higher order the performance is reducing, accuracy of the model is decreasing. This data set is suitable for the linear separation.

### **Part IV: Building an SVM Classification Model for the US Census Dataset**

Text

Description automatically generated

1. What is the obtained classification error (or accuracy)?

Accuracy =83.8157% === 84%.

Text

Description automatically generated

Text

Description automatically generated

What are the obtained testing errors?

1 - 0.83 = 0.17

Which SVM model gives the lowest testing error?

radial basis function with gamma value 0.1 has the lowest test Errors (0.16).

How to do you explain that?

As we go to the higher order the performance is increasing, accuracy of the model is increasing. This data set is suitable for the radial basis function with gamma value 0.1 separation.