1. Data Preprocessing:

* Data format is txt file, so I convert its format into SCV file.
* Data has a duplicate percentage = 0.8 % of the data set so we will drop that to remove redundancy.

1. I will use the linear classifies kernel in each design (soft margin and hard margin)

* For hard margin:

The value of C = 100 (The C parameter tells the SVM optimization how much you want to avoid misclassifying each training example. For large values of C, the optimization will choose a smaller-margin hyperplane if that hyperplane does a better job of getting all the training points classified correctly)

* For soft margin:

The value of C will be as default (C =1).

1. For soft margin SVM:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| accuracy | TPR | FPR | specificity | precision |
| 1 | 0.99 | 0 | 1 | 1 |

For hard margin:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| accuracy | TPR | FPR | specificity | precision |
| 1 | 0.99 | 0 | 1 | 1 |

1. Chart

   Description automatically generated with medium confidence

Chart

Description automatically generated with medium confidence

Comment:

As we see that the accuracy has a good value which means that the model success to classify the data