CS 584-04: Machine Learning
Fall 2019: Assignment 5

Amitdeb Prasad Bhattacharya   A20402789

Question 1

a) **(5 points) What percent of the observations have SpectralCluster equals to 1?**

```
50.0 % percent of the observations have SpectralCluster equals to 1
```

b) **(15 points) You will search for the neural network that yields the lowest loss value and the lowest misclassification rate. You will use your answer in (a) as the threshold for classifying an observation into SpectralCluster = 1. Your search will be done over a grid that is formed by cross-combining the following attributes: (1) activation function: identity, logistic, relu, and tanh; (2) number of hidden layers: 1, 2, 3, 4, and 5; and (3) number of neurons: 1 to 10 by 1. List your optimal neural network for each activation function in a table. Your table will have four rows, one for each activation function. Your table will have five columns: (1) activation function, (2) number of layers, (3) number of neurons per layer, (4) number of iterations performed, (5) the loss value, and (6) the misclassification rate.**

```
   index      Loss activation  misclassification nHiddenNeuron nLayer niter  \
0      0  0.666290  identity                 0.5             1      1    10
1      0  0.000586  logistic                 0.0             9      3   353
2      0  0.000112      tanh                 0.0            10      4    71
3      0  0.000071      relu                 0.0             8      4    40

   output layer activation function
0                         logistic
1                         logistic
2                         logistic
3                         logistic
```

c) **(5 points) What is the activation function for the output layer?**

```
 Activation function of outlayer is Logistic
```

d) **(5 points) Which activation function, number of layers, and number of neurons per layer give the lowest loss and the lowest misclassification rate? What are the loss and the misclassification rate? How many iterations are performed?**

```
   index      Loss activation  misclassification nHiddenNeuron nLayer niter  \
3      0  0.000071      relu                 0.0             8      4    40

   output layer activation function
3                         logistic
```

e) (5 points) Please plot the y-coordinate against the x-coordinate in a scatterplot. Please color-code the points using the predicted SpectralCluster (0 = Red and 1 = Blue) from the optimal MLP in (d). Besides, plot the hyperplane as a dotted line to the graph. To obtain the full credits, you should properly label the axes, the legend, and the chart title. Also, grid lines should be added to the axes.

PredictedClass	id	x	y	SpectralCluster
0	52.5	-0.052476	-0.446791	0.0
1	48.5	0.018718	0.429089	1.0

MLP (3 Layers, 7 Neurons)
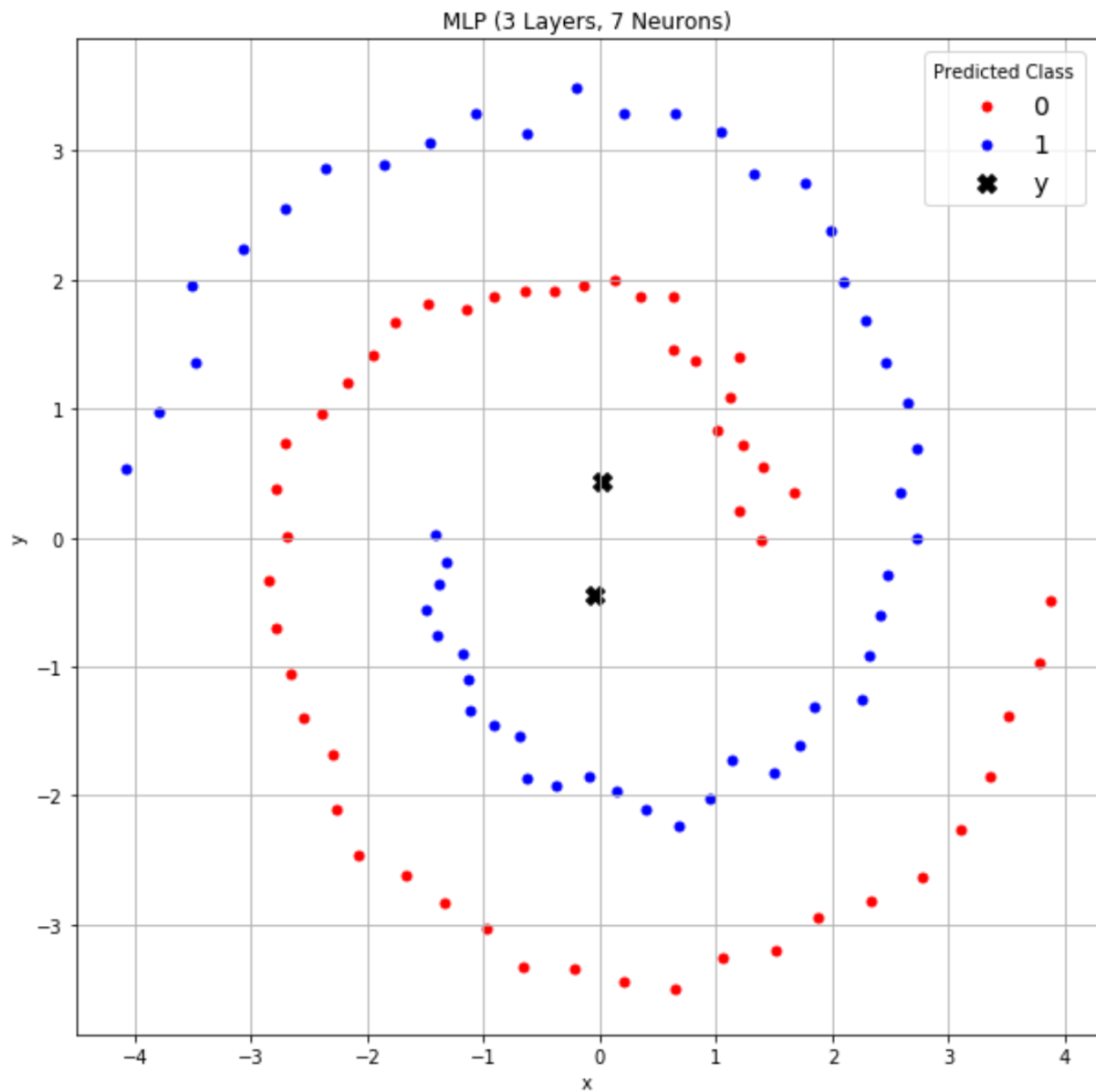
f) **(5 points) What is the count, the mean and the standard deviation of the predicted probability Prob(SpectralCluster = 1) from the optimal MLP in (d) by value of the SpectralCluster? Please give your answers up to the 10 decimal places.**

```
count:               50.00002237510553
mean:                0.5000002237510554
standard deviation:  0.4999958328598052
```

Question 2

a) **(5 points) What is the equation of the separating hyperplane? Please state the coefficients up to seven decimal places.**

```
                  id        x         y  SpectralCluster
_PredictedClass_
0              60.16 -0.110362 -1.672481              0.5
1              40.84  0.076604  1.654779              0.5

Intercept =  [0.003345]
Coefficients =  [[0.05333512 0.32868383]]
```

**b)** **(5 points) What is the misclassification rate?**

```
Misclassification : 0.5
```

c) **(5 points) Please plot the y-coordinate against the x-coordinate in a scatterplot. Please color-code the points using the predicted SpectralCluster (0 = Red and 1 = Blue). Besides, plot the hyperplane as a dotted line to the graph. To obtain the full credits, you should properly label the axes, the legend, and the chart title. Also, grid lines should be added to the axes.**
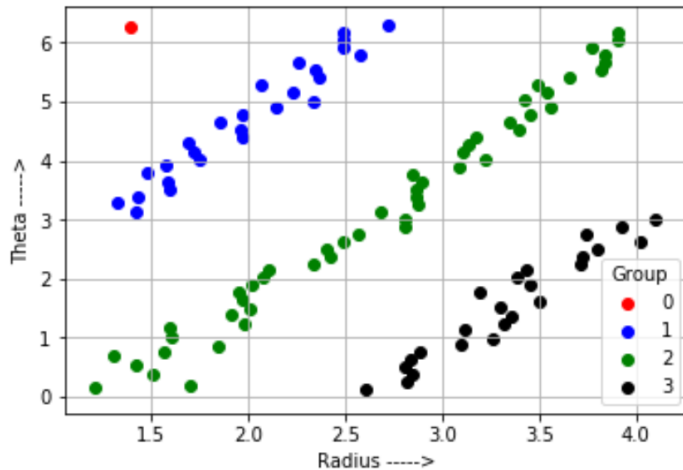
Support Vector Machines
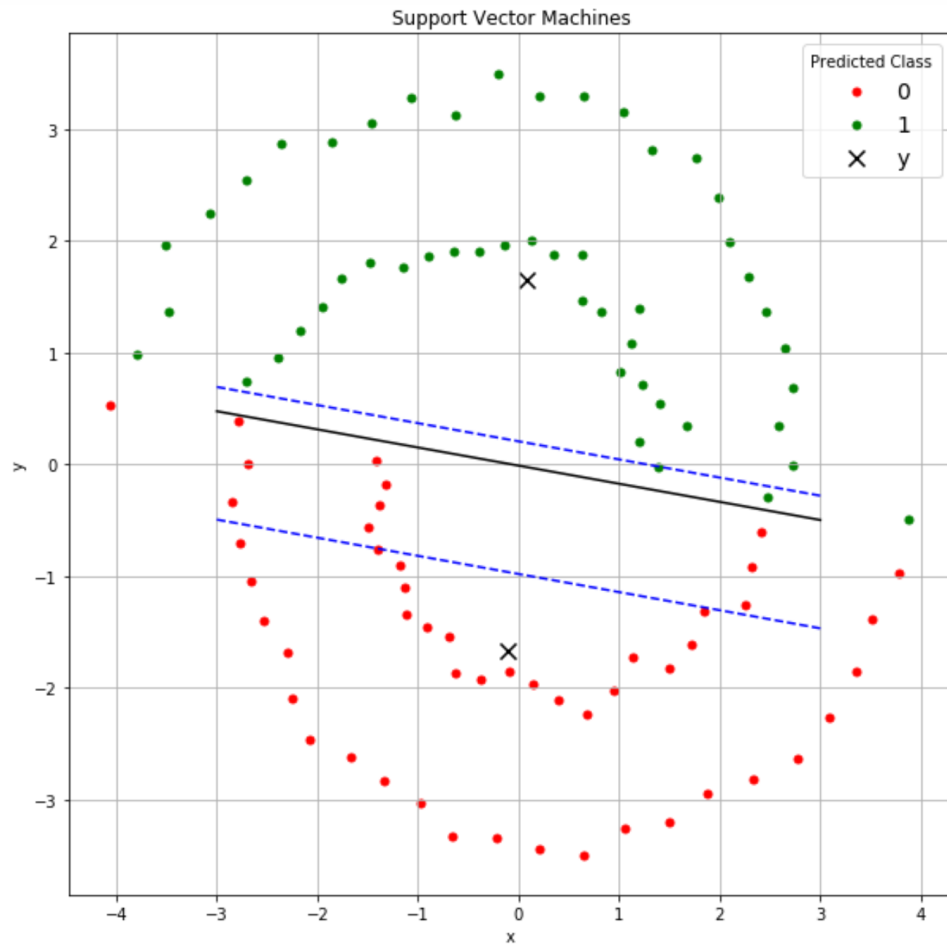
d) **(10 points) Please express the data as polar coordinates. Please plot the theta-coordinate against the radius-coordinate in a scatterplot. Please color-code the points using the SpectralCluster variable (0 = Red and 1 = Blue). To obtain the full credits, you should properly label the axes, the legend, and the chart title. Also, grid lines should be added to the axes.**

```
radius    0
theta     0
dtype: int64
Mean Accuracy =  0.55
Intercept =  [1.7610532]
Coefficients =  [[-0.57239859 -0.05567796]]
```

Support Vector Machines on Three Segments
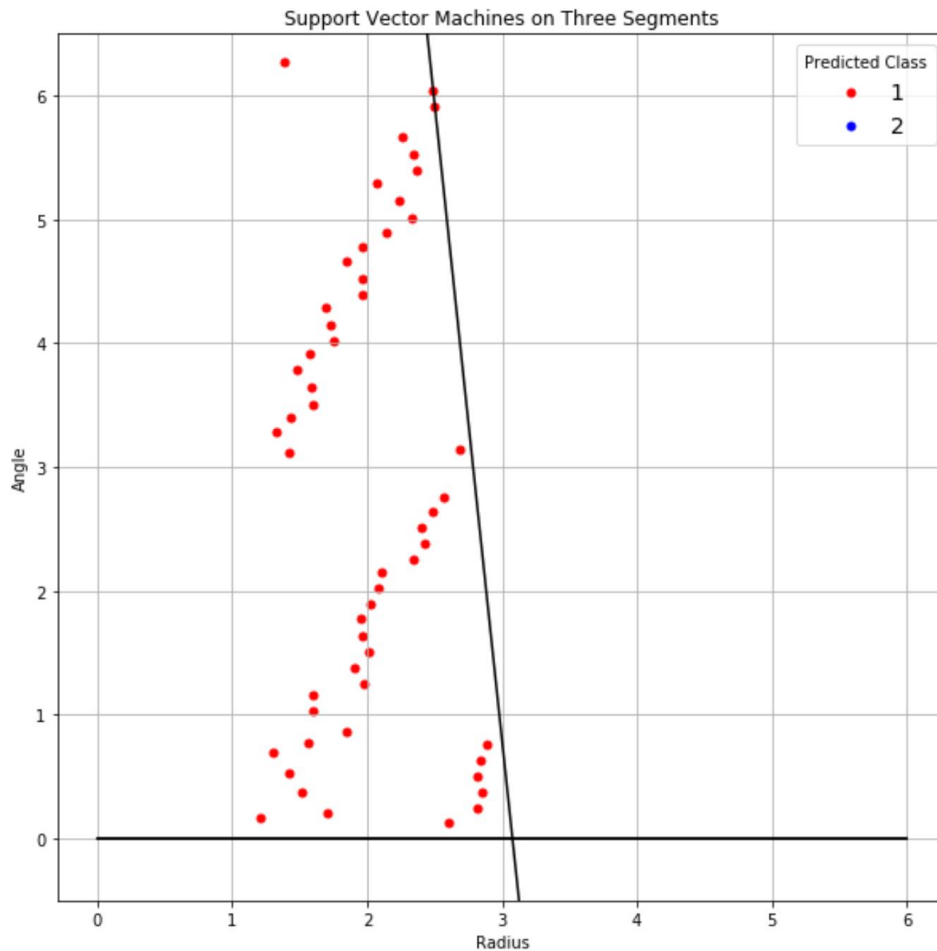
e) **(10 points) You should expect to see three distinct strips of points and a lone point. Since the SpectralCluster variable has two values, you will create another variable, named Group, and use it as the new target variable. The Group variable will have four values. Value 0 for the lone point on the upper left corner of the chart in (d), values 1, 2,and 3 for the next three strips of points. Please plot the theta-coordinate against the radius-coordinate in a scatterplot. Please color-code the points using the new Group target variable (0 = Red, 1 = Blue, 2 = Green, 3 = Black). To obtain the full credits, you should properly label the axes, the legend, and the chart title. Also, grid lines should be added to the axes.**

Support Vector Machines on Four Segments

f) (10 points) Since the graph in (e) has four clearly separable and neighboring segments, we will apply the Support Vector Machine algorithm in a different way. Instead of applying SVM once on a multi-class target variable, you will SVM three times, each on a binary target variable.
SVM 0: Group 0 versus Group 1
SVM 1: Group 1 versus Group 2
SVM 2: Group 2 versus Group 3
Please give the equations of the three hyperplanes.

```
Equation of the separating hyperplane for SVM 0:
w_0+w_1*x_1+w_2*x_2=0
(1.4691251) + (0.9337841*x_1) + (-0.4538025*x_2) = 0

Equation of the separating hyperplane for SVM 1:
w_0+w_1*x_1+w_2*x_2=0
(0.8840632) + (-1.8867496*x_1) + (0.8914745*x_2) = 0

Equation of the separating hyperplane for SVM 2:
w_0+w_1*x_1+w_2*x_2=0
(-4.1328449) + (2.0125835*x_1) + (-0.8375616*x_2) = 0
```
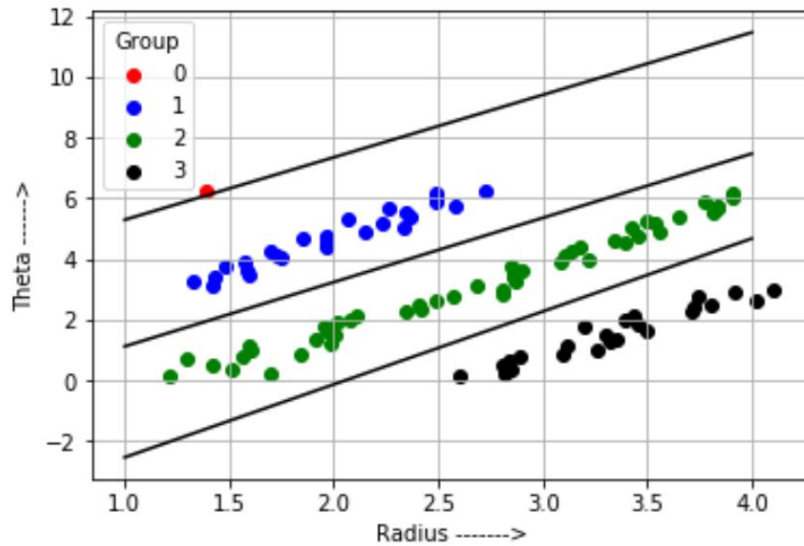
g) (5 points) Please plot the theta-coordinate against the radius-coordinate in a scatterplot. Please color-code the points using the new Group target variable (0 = Red, 1 = Blue, 2 = Green, 3 = Black). Please add the hyperplanes to the graph. To obtain the full credits, you should properly label the axes, the legend, and the chart title. Also, grid lines should be added to the axes.

Support Vector Machines on Four Segments

h) **(10 points) Convert the observations along with the hyperplanes from the polar coordinates back to the Cartesian coordinates. Please plot the y-coordinate against the x-coordinate in a scatterplot. Please color-code the points using the SpectralCluster (0 = Red and 1 = Blue). Besides, plot the hyper-curves as dotted lines to the graph. To obtain the full credits, you should properly label the axes, the legend, and the chart title. Also, grid lines should be added to the axes.**

Based on your graph, which hypercurve do you think is not needed?
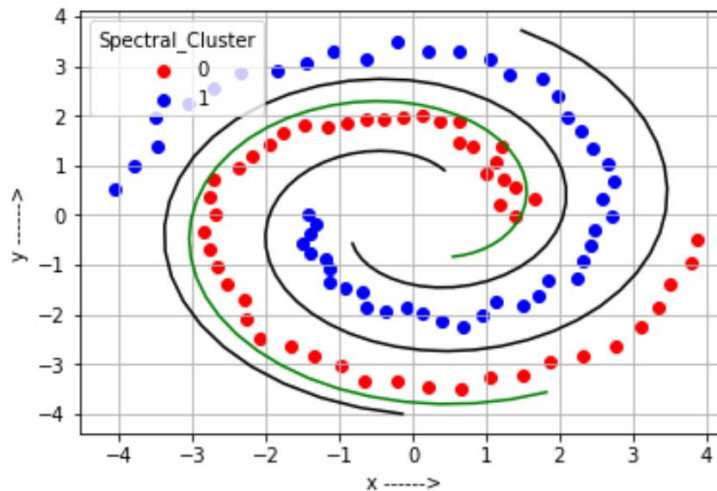
Support Vector Machines on Two Segments

Green colored hypercurve is not required.