Department of Engineering/Informatics, King's College London

Pattern Recognition, Neural Networks and Deep Learning (7CCSMPNN): Coursework 2: SVMs

Alexander Smerdon
23031306

March 2024

Q1:

First 7 digits of student ID: 2303130

Q2:

2 + 3 + 0 + 3 + 1 + 3 + 0 = 12
The remainder of $\frac{12}{4} = 0$ (12 % 4 = 0)
Method: One against one.

Q3:

Own dataset:

Sample of Class 1	Sample of Class 2	Sample of Class 3
(1,12)	(4,4)	(9,11)
(1,14)	(6,4)	(10,10)
(2,11)	(7,5)	(10,9)
(3,13)	(7,6)	(11,15)
(4,8)	(8,5)	(12,11)
(5,10)	(8,1)	(12,15)
(5,12)	(8,6)	(13,8)
(6,10)	(9,5)	(13,14)
(5,12)	(10,4)	(14,7)
(5,15)	(11,6)	(14,15)

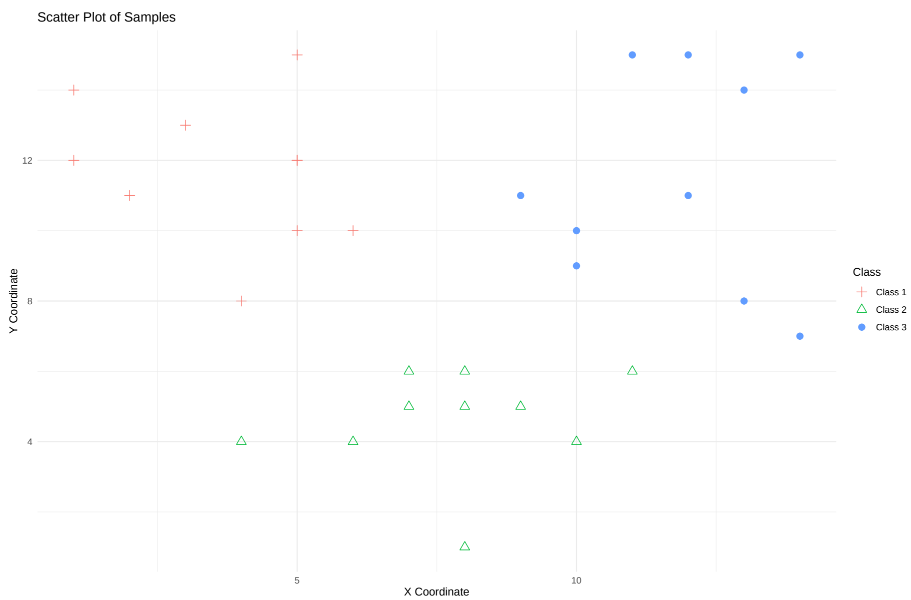

Figure 1: Own data plotted.

Q4:

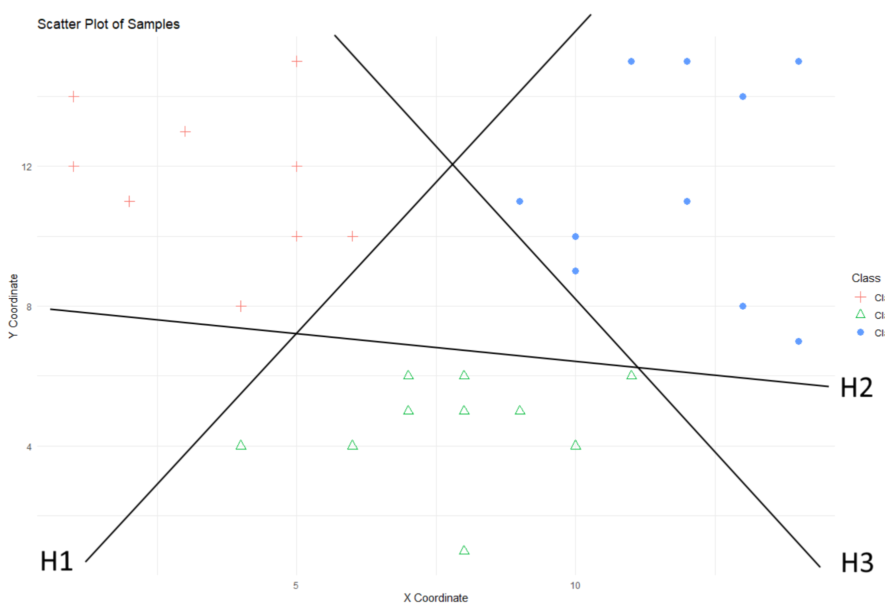

Figure 2: Plot of samples with drawn hyper-planes.

My dataset is linearly separable because all three classes can be separated from each other using straight lines. This can be seen in the plot (Figure 2). Hyperplane $H1$ separates class 1 from classes 2 and 3. Hyperplane $H2$ separates class 2 from classes 1 and 3. Finally, Hyperplane $H3$ separates class 3 from classes 1 and 2.

Q5:

The one against one approach to support vector machines works by first creating a binary SVM between each of the classes one by one. In my case, this means calculating the SVMs for class 1 vs class 2, class 1 vs class 3, and class 2 vs class 3. When a new sample needs to be classified, a classification is calculated for each of the binary SVMs. Based on the class it is most assigned to (majority voting), it is assigned that class. This can be seen in Figure 3.

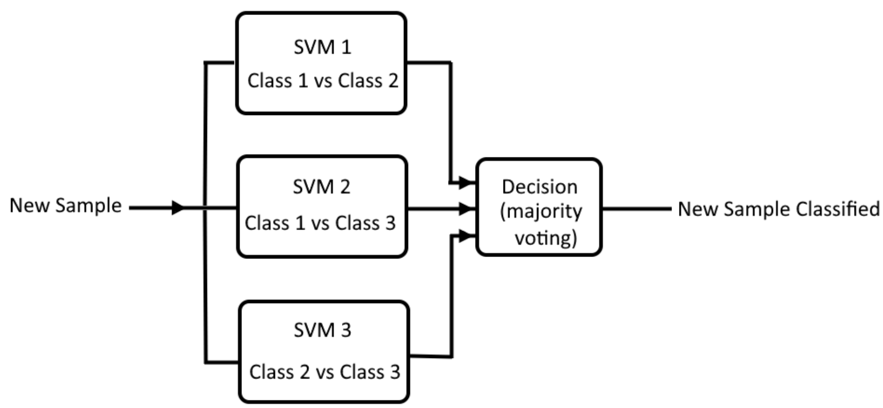

Figure 3: Block diagram of one against one SVM method.

Q6:

Support Vectors of linear SVMs by inspection:

SVM 1 can be seen on Figure 4. The support vectors are highlighted in purple and correspond to the following coordinates: (6,10) and (4,4).

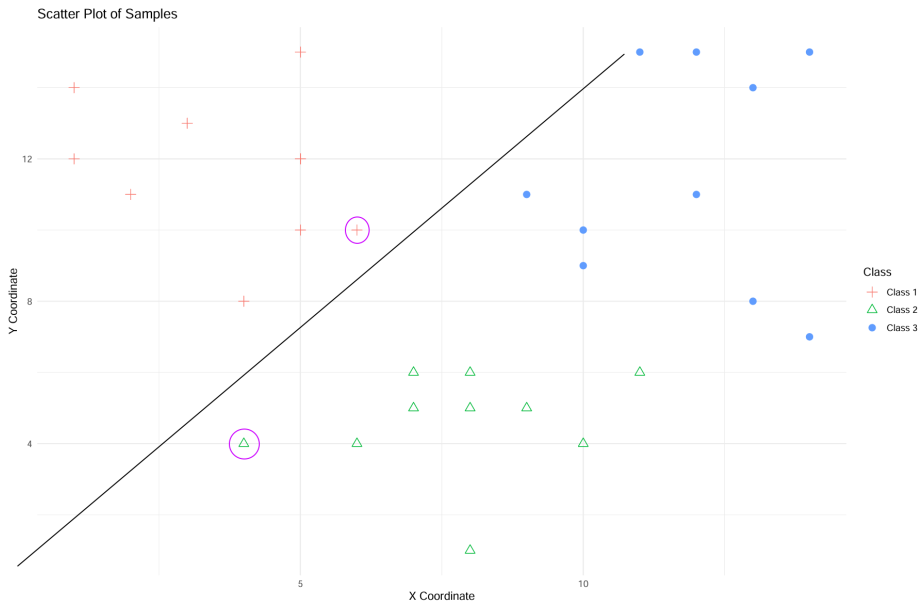

Figure 4: Support Vectors for class 1 vs class 2.

SVM 2 can be seen on Figure 5. The support vectors are highlighted in purple and correspond to the following coordinates: (6,10) and (9,11).

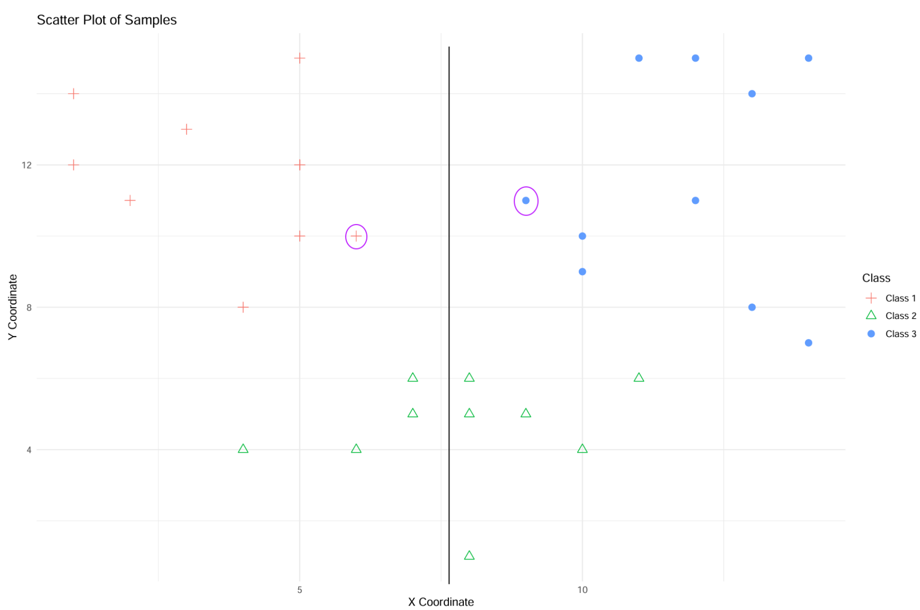

Figure 5: Support Vectors for class 1 vs class 3.

SVM 3 can be seen on Figure 6. The support vectors are highlighted in purple and correspond to the following coordinates: (11,6) and (10,9).

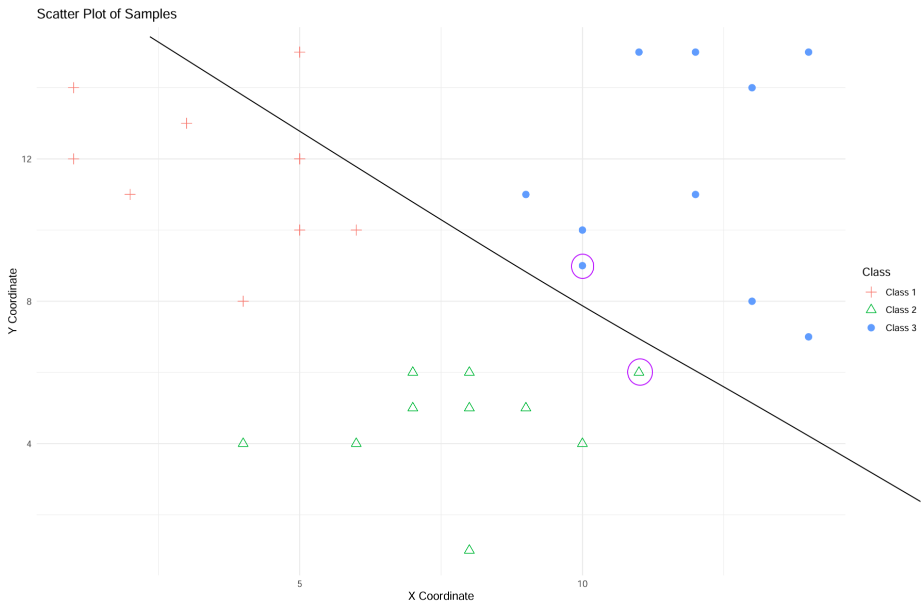

Figure 6: Support Vectors for class 2 vs class 3.

Calculations for SVM1 by hand (class 1 vs class 2):

$$\textbf{Class 1 } (x_1)\textbf{:} \begin{pmatrix} 6 \\ 10 \end{pmatrix}. \textbf{ Class 2 } (x_2)\textbf{:} \begin{pmatrix} 4 \\ 4 \end{pmatrix}.$$

$$w = \lambda_1 \begin{pmatrix} 6 \\ 10 \end{pmatrix} - \lambda_2 \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

$$x = x_1 : [\lambda_1 \begin{pmatrix} 6 \\ 10 \end{pmatrix} - \lambda_2 \begin{pmatrix} 4 \\ 4 \end{pmatrix}]^T \begin{pmatrix} 6 \\ 10 \end{pmatrix} + w_0 = 1 \quad \Rightarrow \quad 136\lambda_1 - 64\lambda_2 + w_0 = 1$$

$$x = x_2 : [\lambda_1 \begin{pmatrix} 6 \\ 10 \end{pmatrix} - \lambda_2 \begin{pmatrix} 4 \\ 4 \end{pmatrix}]^T \begin{pmatrix} 4 \\ 4 \end{pmatrix} + w_0 = -1 \quad \Rightarrow \quad 64\lambda_1 - 32\lambda_2 + w_0 = -1$$

Putting this in matrix form and solving for λ_1, λ_2, and w_0:

$$\begin{pmatrix} 136 & -64 & 1 \\ 64 & -32 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ w_0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

Which gives:

$$\lambda_1 = 0.05,\ \lambda_2 = 0.05,\ w_0 = -2.6$$

To calculate matrix w:

$$0.05 \begin{pmatrix} 6 \\ 10 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.5 \end{pmatrix}$$

$$0.05 \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.2 \end{pmatrix}$$

$$w = \begin{pmatrix} 0.3 \\ 0.5 \end{pmatrix} - \begin{pmatrix} 0.2 \\ 0.2 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.3 \end{pmatrix}$$

Support vector for class 1:

$$w^T x_1 = \begin{pmatrix} 0.1 & 0.3 \end{pmatrix} \begin{pmatrix} 6 \\ 10 \end{pmatrix} = 0.6 + 3 = 3.6$$

$$3.6 - 2.6 = 1$$

Support vector for class 2:

$$w^T x_2 = \begin{pmatrix} 0.1 & 0.3 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 0.4 + 1.2 = 1.6$$

$$1.6 - 2.6 = -1$$

The calculations for SVM 2 (class 1 vs class 3) and SVM 3 (class 2 vs class 3) are much the same, and give these results:

$$\text{SVM 2: } \begin{pmatrix} -0.6 \\ -0.2 \end{pmatrix}^\top \mathbf{x} + 6.6 = 0$$

$$\text{SVM 3: } \begin{pmatrix} 0.2 \\ -0.6 \end{pmatrix}^\top \mathbf{x} + 2.4 = 0$$

Where values of $\mathbf{x}$ give the decision boundary.

Q7:

Sample of Class 1	Sample of Class 2	Sample of Class 3	Mean of Samples
(1,12)	(4,4)	(9,11)	(4.66, 9)
(1,14)	(6,4)	(10,10)	(5.6, 9.3)
(2,11)	(7,5)	(10,9)	(6.3, 8.3)
(3,13)	(7,6)	(11,15)	(7, 11.3)
(4,8)	(8,5)	(12,11)	(8, 8)
(5,10)	(8,1)	(12,15)	(8.3, 8.6)
(5,12)	(8,6)	(13,8)	(8.6, 8.6)
(6,10)	(9,5)	(13,14)	(9.3, 9.6)
(5,12)	(10,4)	(14,7)	(9.6, 7.6)
(5,15)	(11,6)	(14,15)	(10, 12)

Figure 7 shows the sample means plotted (in yellow) with the original data points:

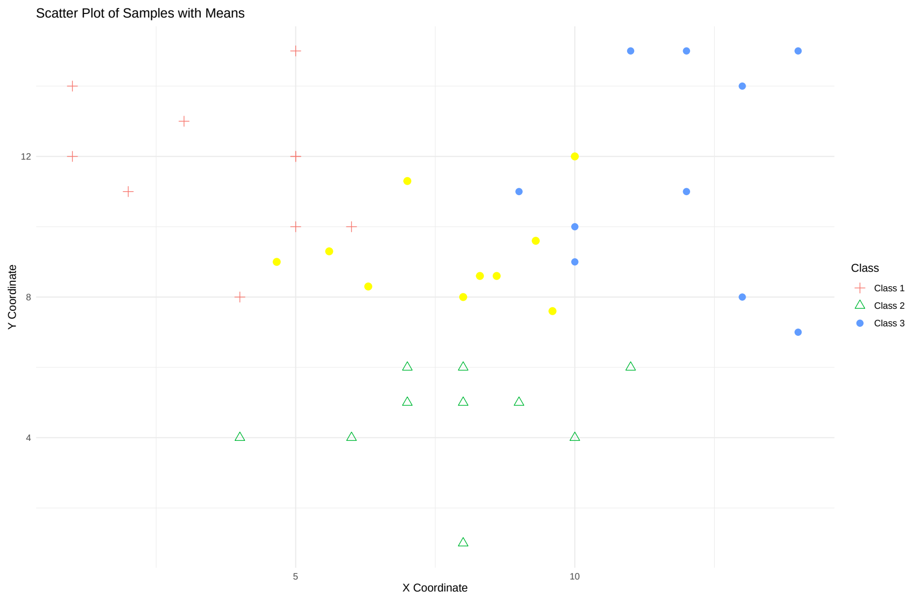

Figure 7: Plotted samples.

Here is a worked example of how to classify the first mean sample $\begin{pmatrix} 4.66 \\ 9 \end{pmatrix}$:

Calculating class using SVM 1 (class 1 vs class 2):

$$(0.1, 0.3)\begin{pmatrix} 4.66 \\ 9 \end{pmatrix} = 0.466 + 2.7 = 3.166.$$

$w_0 = -2.6$, so $3.166 - 2.6 = 0.566$, which is positive so corresponds to class 1.

Calculating class using SVM 2 (class 1 vs class 3):

$$(-0.6, -0.2)\begin{pmatrix} 4.66 \\ 9 \end{pmatrix} = -2.796 - 1.8 = -4.596.$$

$w_0 = 6.6$, so $-4.596 + 6.6 = 2.004$, which is positive so corresponds to class 1.

Calculating class using SVM 3 (class 2 vs class 3):

$$(0.2, -0.6)\begin{pmatrix} 4.66 \\ 9 \end{pmatrix} = 0.932 - 5.4 = -4.468.$$

$w_0 = 2.4$, so $-4.468 + 2.4 = -2.068$, which is negative so corresponds to class 3.

With the One against one method, the new classification is done through majority voting, so because we have two votes for class 1, and one vote for class 3, the sample $\begin{pmatrix} 4.66 \\ 9 \end{pmatrix}$ is classified as class 1.

Below is a table with a few new samples classified:

New Sample	Output of SVM 1	Output of SVM 2	Output of SVM 3	Classification
(4.66, 9)	0.566 (Class 1)	2.004 (Class 1)	-2.068 (Class 3)	Class 1
(8, 8)	0.6 (Class 1)	0.2 (Class 1)	-0.8 (Class 3)	Class 1
(10, 12)	2 (Class 1)	-1.8 (Class 3)	-2.8 (Class 3)	Class 3

Although my mean samples don't demonstrate it, the main flaw of the One against one method is that, if each SVM classifies the sample differently, a majority decision about the class cannot be made and so a different method is needed for classification.