

Alfie Lamerton - Support Vector Machine Assignment

Metadata

Module: Pattern Recognition, Neural Networks and Deep Learning (7CCSMPNN)

Name: Alfie Lamerton

Student Number: K23080018

1: First 7 Digits of Student ID

$\{2, 3, 0, 8, 0, 0, 1\}$

2: Find R_1

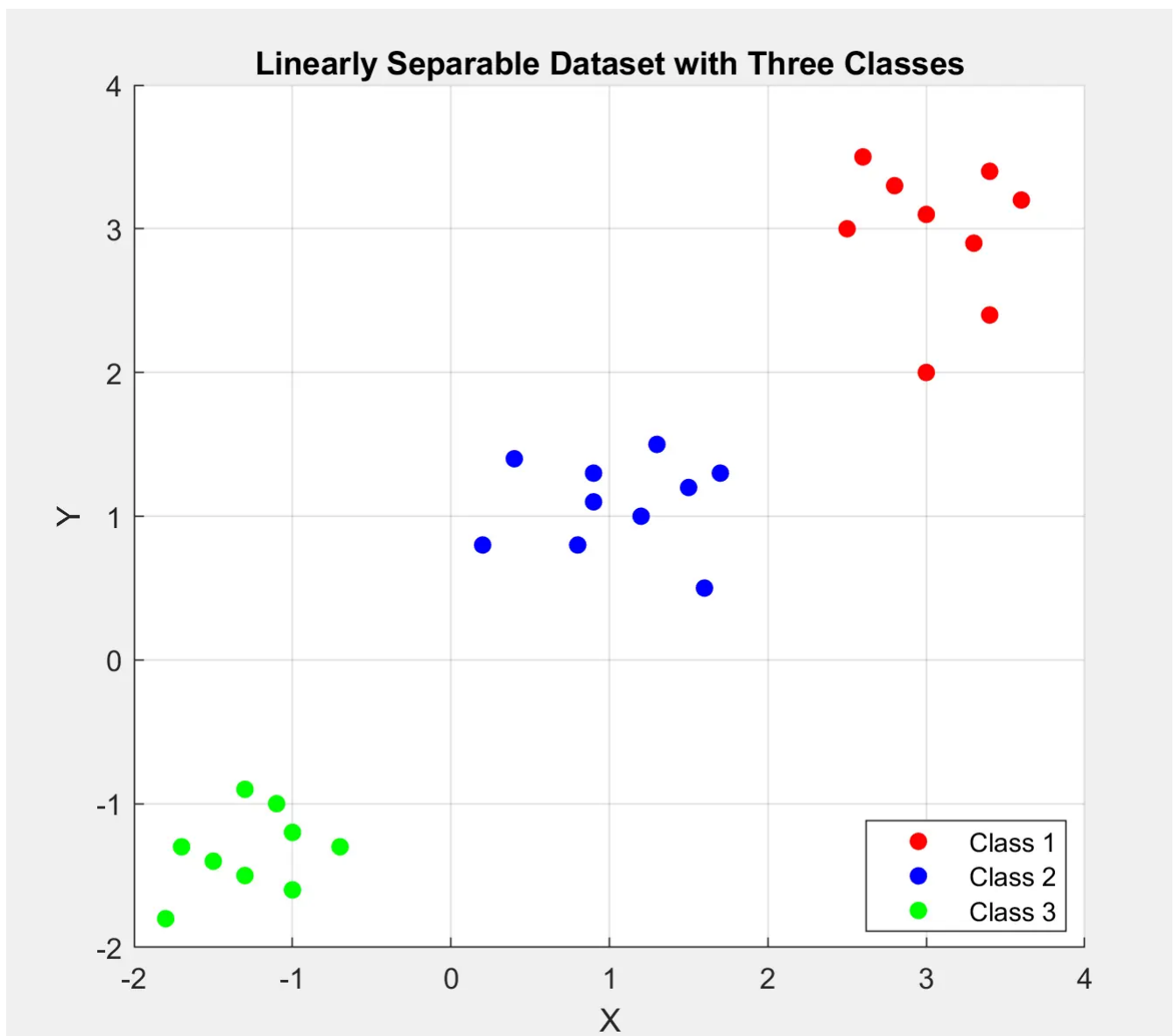
The remainder of $\frac{2+3+0+8+0+0+1}{4} = 2$, therefore the method for this coursework is **Binary Decision Tree**.

3: Dataset

Sample of Class 1	Sample of Class 2	Sample of Class 3
(2.8, 3.3)	(0.9, 1.3)	(-1.3, -0.9)
(3.3, 2.9)	(0.2, 0.8)	(-0.7, -1.3)
(3.0, 3.1)	(0.9, 1.1)	(-1.1, -1.0)
(3.4, 3.4)	(1.3, 1.5)	(-1.3, -1.5)
(2.5, 3.0)	(0.8, 0.8)	(-1.0, -1.2)
(3.0, 2.0)	(1.2, 1.0)	(-1.1, -1.0)
(3.4, 3.4)	(1.5, 1.2)	(-1.5, -1.4)
(3.4, 2.4)	(1.6, 0.5)	(-1.8, -1.8)
(3.6, 3.2)	(1.7, 1.3)	(-1.7, -1.3)
(2.6, 3.5)	(0.4, 1.4)	(-1.0, -1.6)

4: Plot

Linear separability is a property of two or more sets of data points, meaning the data clusters are spread in a way where members of each cluster don't mix at all. Two sets of data points are linearly separable if there is a linear function which classifies all of the data points with no overlap.



This plot shows that the dataset is linearly separable. This is shown by the distinct cluster each data point falls in. This dataset is linearly separable because each cluster of data points can be completely separated only using straight lines.

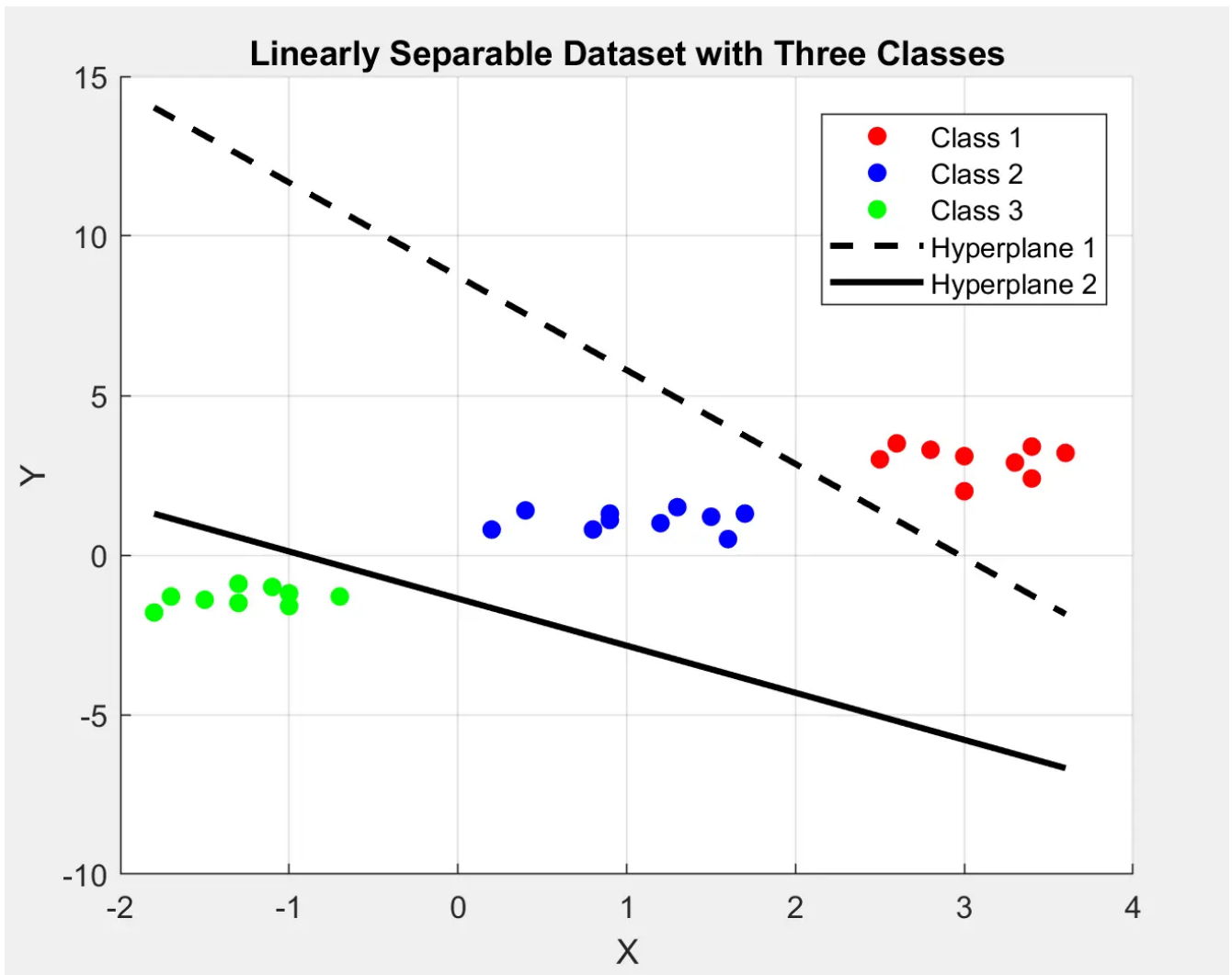


Figure: Plot of Dataset with Hand-Drawn Hyperplanes

In this image, hyperplane 1 serves the job of separating class 3 from classes 1 and 2, and hyperplane 2 serves the job of separating class 2 from classes 3 and 4. Therefore, this dataset is linearly separable because these straight hyperplanes can be used to separate the data.

5: Binary Decision Tree Block Diagram

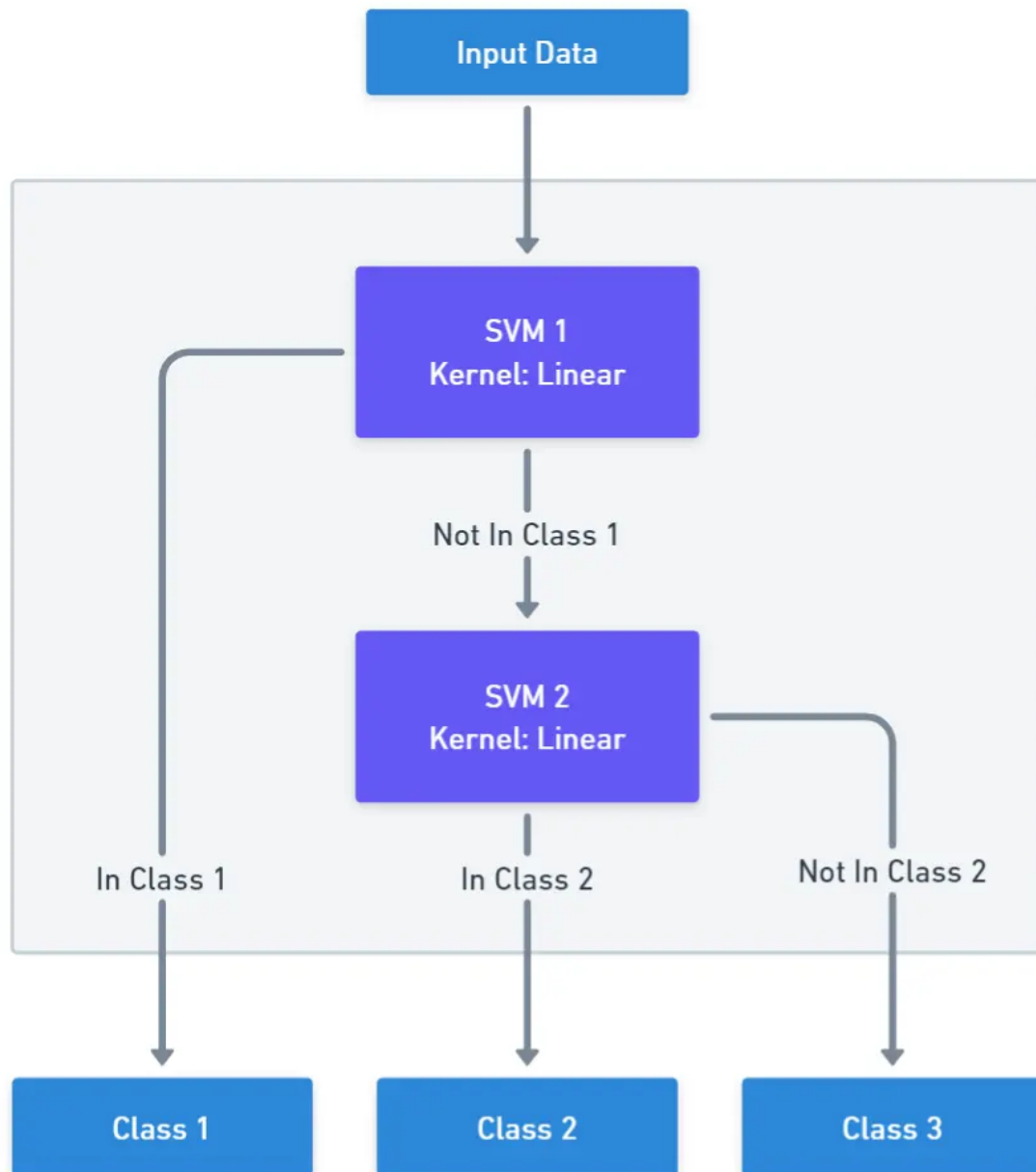


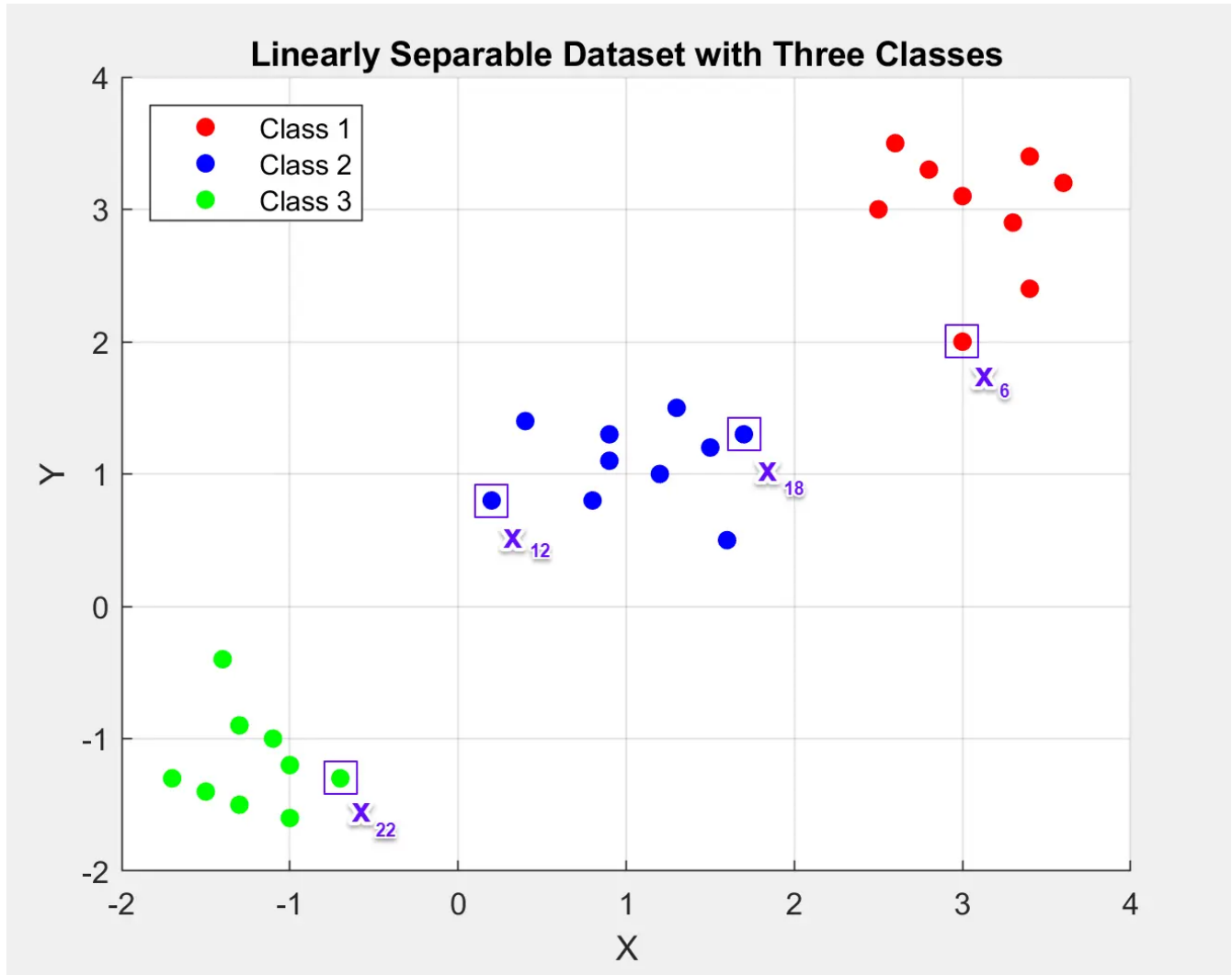
Figure: Binary decision tree separating 3 classes using 2 SVMs

This support vector machine (SVM) has 30 inputs, corresponding to each data point in the dataset. The SVM outputs a class for each data point, there are 3 classes and therefore 3 outputs. There are two SVMs. SVM 1 separates data points in Class 1 from those in Classes 2 and 3. SVM 2 separates data points in Class 2 from those in Class 3.

6. Support Vectors and Hyperplane

Identifying Support Vectors

The support vectors and hyperplanes can be identified in this figure by inspection.



The support vector for Class 1 is $x_6 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. The support vectors for Class 2 are $x_{12} = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$ (for separating it from Class 3), and $x_{18} = \begin{pmatrix} 1.7 \\ 1.3 \end{pmatrix}$ (for separating it from Class 1). The support vector for Class 3 is $x_{22} = \begin{pmatrix} -0.7 \\ -1.3 \end{pmatrix}$.

Classes 1 and 2 can be separated with a hyperplane in the region in between $1.7 \leq x \leq 3$ and $1.3 \leq y \leq 2$, and Classes 2 and 3 can be separated with a hyperplane in the region in between $-0.7 \leq x \leq 0.2$ and $-1.3 \leq y \leq 0.8$.

Designing Hyperplanes

Calculations

Hyperplanes which optimally separate the classes can be designed based on these support vectors. The equation for calculating a hyperplane is $w^T x + b = 0$, where w is the normal vector to the hyperplane and b is the bias term. Two hyperplanes must be defined for separating the classes, one separating Class 1 from Class 2, and one separating Class 2 from Class 3.

Separating Class 1 from Class 2

The support vectors separating Class 1 from Class 2 are $x_6 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ and $x_{18} = \begin{pmatrix} 1.7 \\ 1.3 \end{pmatrix}$ respectively.

The class labels for these vectors can be defined as $+1$ or -1 depending on their membership to

classes 1 and 2. x_6 belongs to Class 1, so its label y_6 is set to +1. x_{18} belongs to Class 2, so its label y_{18} is set to -1.

The equation for calculating a hyperplane is:

$$\mathbf{w}^T \mathbf{x} + w_0 = 0.$$

The values of the vectors and their labels can be substituted into this equation:

$$\mathbf{w} = \lambda_1 y_1 x_1 + \lambda_2 y_2 x_2 = \lambda_1 \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1.7 \\ 1.3 \end{pmatrix}$$

This can then be decomposed and solved by substitution:

$$\left[\lambda_1 \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \lambda_2 \begin{pmatrix} 1.7 \\ 1.3 \end{pmatrix} \right]^T \begin{pmatrix} 3 \\ 2 \end{pmatrix} + b = 1 \implies 13\lambda_1 - 7.7\lambda_2 + b = 1$$

$$\left[\lambda_1 \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \lambda_2 \begin{pmatrix} 1.7 \\ 1.3 \end{pmatrix} \right]^T \begin{pmatrix} 1.7 \\ 1.3 \end{pmatrix} + b = -1 \implies 7.7\lambda_1 - 4.58\lambda_2 + b = -1$$

$$w_1 + w_2 + 0 \times b = 0$$

$$\begin{pmatrix} 13 & 7.7 & 1 \\ 7.7 & -4.58 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

These equations can be rearranged and solved to produce the values of the unknown variables: $\lambda_1 = 0.9174$, $\lambda_2 = 0.9174$, and $w_0 = -3.8624$. These values are now substituted into the equation from above to find w_1 and w_2 :

$$\mathbf{w} = \lambda_1 y_1 x_1 + \lambda_2 y_2 x_2 = 0.9174 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} - 0.9174 \times \begin{pmatrix} 1.7 \\ 1.3 \end{pmatrix} = \begin{pmatrix} 1.1926 \\ 0.6422 \end{pmatrix}$$

Making $w_1 = 1.1926$, and $w_2 = 0.6422$. These values can then be substituted into equation for the hyperplane separating Class 1 and Class 2:

$$\mathbf{w}^T \mathbf{x} + w_0 = 1.1926x_1 + 0.6422x_2 - 3.8624$$

Separating Class 2 from Class 3

The support vectors separating Class 2 from Class 3 are $x_{12} = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$ and $x_{22} = \begin{pmatrix} -0.7 \\ -1.3 \end{pmatrix}$ respectively. The same labelling method is used to label x_{12} and x_{22} with respect to their membership in Class 2 and Class 3. x_{12} belongs to Class 2, so its label y_{12} is set to +1. x_{22} belongs to Class 3, so its label y_{22} is set to -1. The values of the vectors and their labels can be substituted into the same equation:

$$\mathbf{w} = \lambda_1 y_1 x_1 + \lambda_2 y_2 x_2 = \lambda_1 \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} + \lambda_2 \begin{pmatrix} -0.7 \\ -1.3 \end{pmatrix}$$

$$\left[\lambda_1 \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} - \lambda_2 \begin{pmatrix} -0.7 \\ -1.3 \end{pmatrix} \right]^T \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} + b = 1 \implies 0.68\lambda_1 + 1.18\lambda_2 + b = 1$$

Which can then be decomposed and solved by substitution:

$$\left[\lambda_1 \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} - \lambda_2 \begin{pmatrix} -0.7 \\ -1.3 \end{pmatrix} \right]^T \begin{pmatrix} -0.7 \\ -1.3 \end{pmatrix} + b = -1 \implies -1.18\lambda_1 - 2.18\lambda_2 + b = -1$$

$$w_1 + w_2 + 0 \times b = 0$$

$$\begin{pmatrix} 0.68 & 1.18 & 1 \\ -1.18 & -2.18 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

These equations can be rearranged and solved to produce the values of the unknown variables:

$\lambda_1 = 0.3831$, $\lambda_2 = 0.3831$, and $w_0 = 0.2874$. These values are now substituted into the equation from above:

$$\mathbf{w} = \lambda_1 y_1 x_1 + \lambda_2 y_2 x_2 = 0.3831 \times \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} - 0.3831 \times \begin{pmatrix} -0.7 \\ -1.3 \end{pmatrix} = \begin{pmatrix} 0.3448 \\ 0.8045 \end{pmatrix}$$

Making $w_1 = 0.3448$, and $w_2 = 0.8045$. These values can then be substituted into equation for the hyperplane separating Class 1 and Class 2:

$$\mathbf{w}^T \mathbf{x} + w_0 = 0.3448x_1 + 0.8045x_2 + 0.2874$$

This is the equation for the hyperplane separating Class 2 from Class 3.

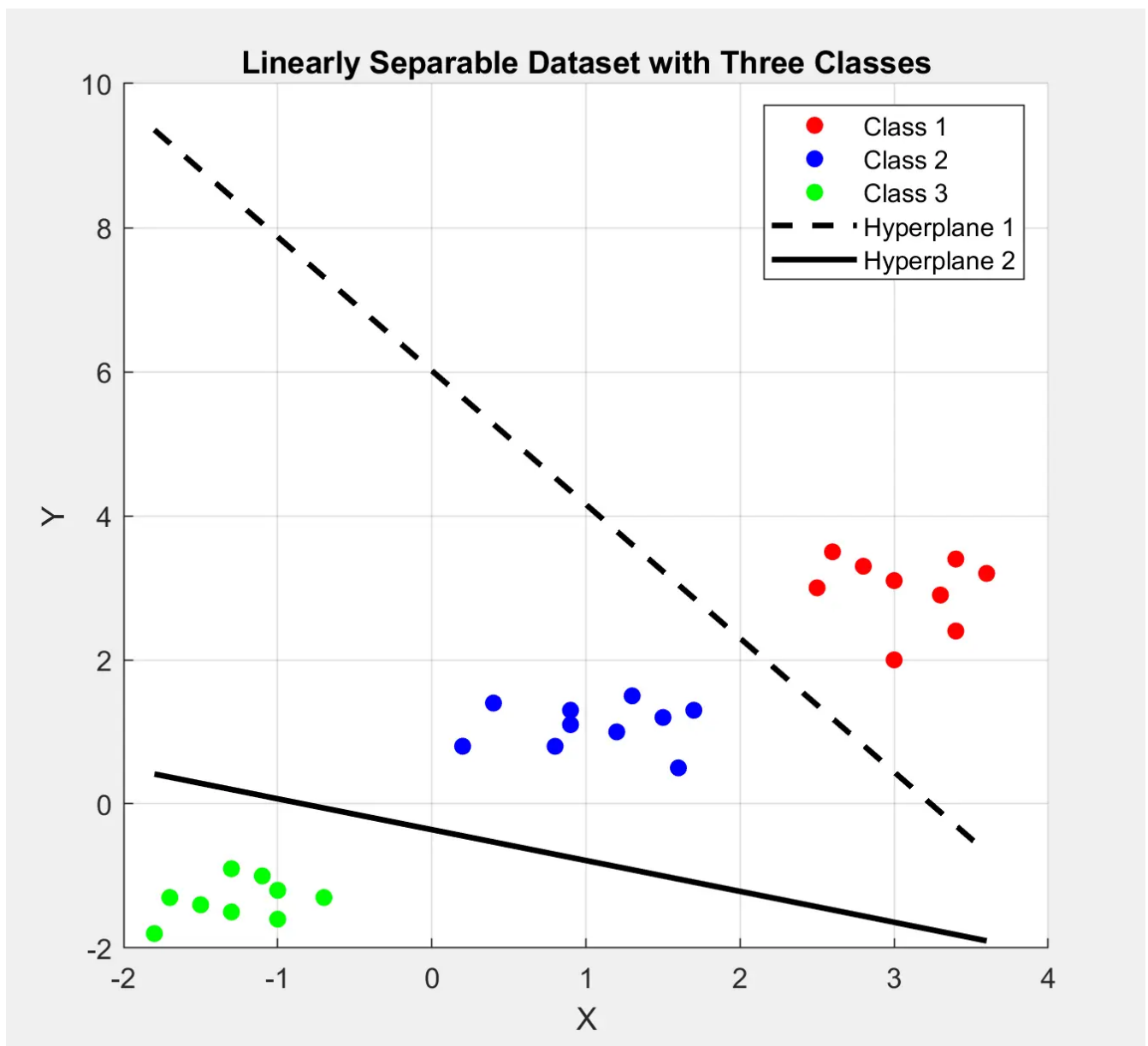


Figure: Plot Showing the Data Points Being Separated by the SVMs

7. Test Dataset

The following table is a summary of the application of the two support vector machines (SVMs) to the average values of the samples in the original dataset.

Test Sample	Output of SVM 2	Output of SVM 3	Classification
(0.8, 1.23)	Negative	Positive	Class 2
(0.93, 0.8)	Negative	Positive	Class 2
(0.93, 1.07)	Negative	Positive	Class 2
(1.13, 1.13)	Negative	Positive	Class 2
(0.77, 0.87)	Negative	Positive	Class 2
(1.03, 0.67)	Negative	Positive	Class 2
(1.13, 1.07)	Negative	Positive	Class 2
(1.07, 0.37)	Negative	Positive	Class 2
(1.2, 1.07)	Negative	Positive	Class 2

Test Sample	Output of SVM 2	Output of SVM 3	Classification
(0.67, 1.1)	Negative	Positive	Class 2

The classification of the first test sample, $\begin{pmatrix} 0.8 \\ 1.23 \end{pmatrix}$, is calculated by first passing the sample into SVM 1:

$$w^T x + w_0 = 1.1926x_1 + 0.6422x_2 - 3.8624$$

Which becomes:

$$(1.1926 \times 0.8) + (0.6422 \times 1.23) - 3.8624 = -2.1184$$

Resulting in a classification of **negative**, as $-2.1184 < -1$. SVM 1 therefore passes the sample to SVM 2:

$$w^T x + w_0 = 0.3448x_1 + 0.8045x_2 + 0.2874$$

Which becomes:

$$(0.3448 \times 0.8) + (0.6422 \times 1.23) + 0.2874 = 1.5528$$

Resulting in a classification of **positive**, as $1.5528 > 1$. SVM 2 therefore classifies the sample as a member of Class 2.