Short Assignment 3

This is an individual assignment.

Due: Friday, November 4 @ 11:59PM

Problem 1

Consider the two classes represented by Gaussians distributions P1 and P2 in Figures 1 and 2. Calculate Fisher's univariate separation indices to answer the following questions.

In [1]:
```
from IPython.display import Image
Image('figures/two-gaussian-distributions.png',width=800)
```

Out[1]:

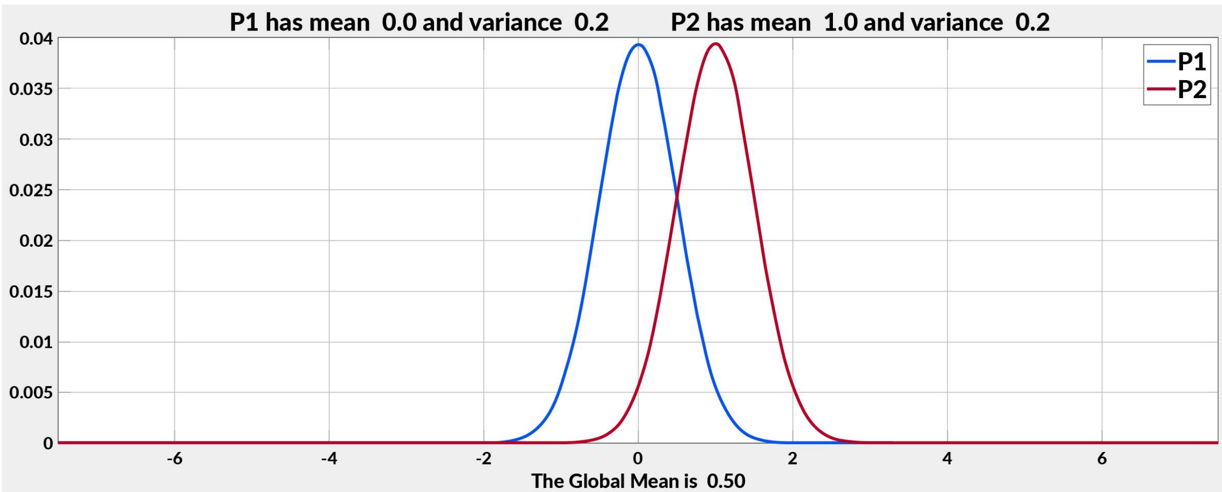


Figure 1: Two Gaussian Classes

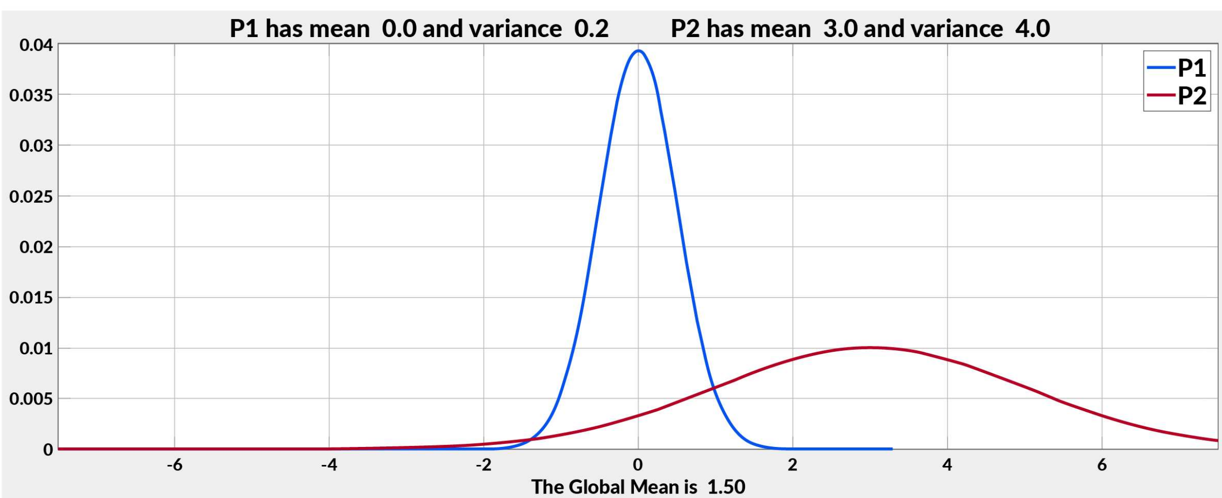


Figure 2: Two Gaussian Classes

1. What is the Within-Class Separation in Figure 1?

2. What is the Within-Class Separation in Figure 2?
3. What is the Between-Class Separation in Figure 1?
4. What is the Between-Class Separation in Figure 2?
5. Which Distributions are more separated: (A) P1,P2 in Figure 1 or (B) P1,P2 in Figure 2? Justify your answer.

1.

1. Within-class separation in Figure 1 = 0.08

1. Within-class separation in Figure 2 = 16.04

1. Between-class separation in Figure 1 = 1

1. Between-class separation in Figure 2 = 9

1. Despite having more difference between the mean's of classes in figure 2, the significant variance in distributions causes notable overlap along the projection vector. Therefore, distributions in figure 1 are more separated.

Short Assignment 3 -

Fisher's Criterion: $J(w) = \dfrac{\overbrace{(m_2 - m_1)^2}^{\text{between class variance}}}{\underbrace{s_1^2 + s_2^2}_{\text{within class variance}}}$

Fig1. (1) $s_1^2 + s_2^2 = 0.2^2 + 0.2^2 = 0.04 + 0.04 = 0.08$

Fig2. (2) $s_1^2 + s_2^2 = 0.2^2 + 4^2 = 0.04 + 16 = 16.04$

Fig1. (3) $(m_1 - m_2)^2 = (0-1)^2 = (-1)^2 = 1.$

Fig2. (4) $(m_1 - m_2)^2 = (0-3)^2 = 9$

(5) J(W) per fig 1.

$J(W) = \dfrac{1}{0.08} = 12.5$

J(W) per fig 2,

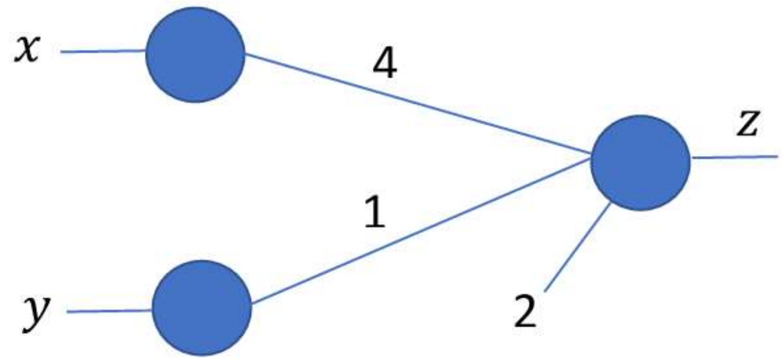
$J(W) = \dfrac{9}{16.04} = 0.5611$

Problem 2

Consider the following perceptron:

In [2]:
```
Image('figures/Perceptron.png', width=400)
```

Out[2]:



Recall that the perceptron uses the activation function:

$$\phi(x) = \begin{cases} -1 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

And the cost function is:

$$E_p(\mathbf{w}, b) = -\sum_{m \in \mathcal{M}} (\mathbf{w}^T \mathbf{x}_n + b)^T t_n$$

where is the set of all misclassified points. The update equations for the weights and bias term are:

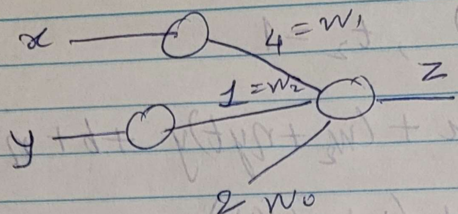
$$\begin{aligned} \mathbf{w}^{(t+1)} \leftarrow & \quad \mathbf{w}^{(t)} - \eta \frac{\partial E_p(\mathbf{w}, b)}{\partial \mathbf{w}} = \mathbf{w}^{(t)} + \eta \mathbf{x}_n t_n \\ b^{(t+1)} \leftarrow & \quad b^{(t)} - \eta \frac{\partial E_p(\mathbf{w}, b)}{\partial b} = b^{(t)} + \eta t_n \end{aligned}$$

Suppose you have the following 5 data samples (x, y) and their corresponding labels t:

$$\begin{gathered} (x_1, y_1) = (1, 0) \text{ with } t_1 = 1 \\ (x_2, y_2) = (4, 2) \text{ with } t_2 = 1 \\ (x_3, y_3) = (0, -1) \text{ with } t_3 = -1 \\ (x_4, y_4) = (-1, -1) \text{ with } t_4 = -1 \\ (x_5, y_5) = (-2, 1) \text{ with } t_5 = -1 \end{gathered}$$

What is the smallest value for the learning rate η such that the updated network will result in zero misclassified points using only one iteration?

problem2.



$$W = \begin{bmatrix}2\\4\\1\end{bmatrix} \qquad x = \begin{bmatrix}x_0\\x_1\\x_2\end{bmatrix}$$

$$y = w_1 x_0 + w_2 y_0 + w_0$$
$$y = 4x_0 + y_0 + 2$$

$t_1 = 1$ $(x_1, y_1) = (1, 0) \Rightarrow y = 4 + 0 + 2 = 6$

$t_2 = 1$ $(x_2, y_2) = (4, 2) \Rightarrow y = 4(4) + 2 + 2 = 20$

$t_3 = -1$ $(x_3, y_3) = (0, -1) \Rightarrow y = 4(0) - 1 + 2 = 1$ ← misclassified

$t_4 = -1$ $(x_4, y_4) = (-1, -1) \Rightarrow y = 4(-1) - 1 + 2 = -3$

$t_5 = -1$ $(x_5, y_5) = (-2, 1) \Rightarrow y = 4(-2) + 1 + 2 = -5$

$(x_1, y_1) \to t_1 = 1$, $x_1 = 1$, $y_1 = 0$

$\therefore y(x_1, y_1) > 0$

$\therefore y = (w_1 + \eta x t)x + (w_2 + \eta y t)y + b + \eta t > 0$

$\therefore (4 + \eta(1)(1))1 + 0 + 2 + \eta(1) > 0$

$\therefore 4 + \eta + 2 + \eta > 0 \Rightarrow 2\eta > -6$

$\Rightarrow \underline{\underline{\eta_1 > -3}}$

$(x_2, y_2) = (4, 2)$, $t_2 = 1$

$\therefore (w_1 + \eta x t)x + (w_2 + \eta y t)y + b + \eta t > 0$

$\therefore (4 + \eta(4))4 + (1 + \eta(2)(1))2 + 2 + \eta > 0$

$\therefore 16 + 16\eta + 2 + 4\eta + 2 + \eta > 0$

$\therefore 20 + 21\eta > 0$

$\therefore \eta_2 > \frac{-20}{21} \Rightarrow \eta_2 > -0.9524$

$(x_3, y_3) = (0, -1)$, $t_3 = -1$

$\therefore (4 + \eta(0))0 + (1 + \eta(-1)(-1))(-1) + 2 + \eta(-1) < 0$

$\therefore 0 - 1 - \eta + 2 - \eta < 0$

$\therefore 1 - 2\eta < 0$

$\therefore -2\eta_3 < -1 \Rightarrow \eta_3 > \frac{1}{2} \Rightarrow \eta_3 > 0.5$

$(x_4, y_4) = (-1, -1)$, $t_4 = -1$

$\therefore (4 + \eta(-1)(-1))(-1) + (1 + \eta(-1)(-1))(-1) + 2 + \eta(-1) < 0$

$\therefore -4 - \eta + (-1) - \eta + 2 - \eta < 0$

$\therefore -3 - 3\eta < 0$

$\therefore -3\eta < 3$

$\therefore \eta_4 > -1$

$(x_5, y_5) = (-2, 1)$, $t_5 = -1$

$\therefore \; (4 + \eta(-2)(-1))(-2) + (1 + \eta(1)(-1)) + 2 + \eta(-1) < 0$

$\therefore \; -8 - 4\eta + 1 - \eta + 2 - \eta < 0$

$\therefore \; -5 - 5\eta < 0$

$\therefore \; -5\eta_5 < 5 \Rightarrow \underline{\underline{\eta_5 > -1}}$

$\eta_1 > -3$

$\eta_2 > -0.9524$

$\eta_3 > 0.5$

$\eta_4 > -1$

$\eta_5 > -1$

$\therefore$ $\eta > 0.5$ is the condition that satisfies all the points.

$$\boxed{\eta > 0.5}$$

Submit Your Solution

Confirm that you've successfully completed the assignment.

Along with the Notebook, include a PDF of the notebook with your solutions.

`add` and `commit` the final version of your work, and `push` your code to your GitHub repository.

Submit the URL of your GitHub Repository as your assignment submission on Canvas.
