The goal of this problem is to run a linear perceptron algorithm. Assume that you have three training samples:

1. Sample $\mathbf{x}_1 = [1, 3]^T$ from Class 1 ($y_1 = 1$)
2. Sample $\mathbf{x}_2 = [3, 2]^T$ from Class 2 ($y_2 = -1$)
3. Sample $\mathbf{x}_3 = [4, 1]^T$ from Class 2 ($y_3 = -1$)

The linear perceptron is initialized with a line with corresponding weight $\mathbf{w}(0) = [-\frac{1}{3}, 1]^T$. In the following, for the sake of simplicity, you will assume that all lines of the perceptron intersect point $(0,0)$, therefore you do not have to include any intercept $w_0$ or $x_0$ in the following calculations.

(1) Plot $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ in the given 2D space. Find and plot the line corresponding to weight $\mathbf{w}(0)$.

For $\mathbf{w}(0) = [-\frac{1}{3}, 1]$, we have $-\frac{1}{3}x + y = 0 \Rightarrow y = \frac{1}{3}x$, therefore the line is plot in the above figure. Note that the direction of the vector $\mathbf{w}_0$ on the line is the same as the direction of $\mathbf{w}(0)$ starting from $(0,0)$.

(2) Using the rule $\text{sign}(\mathbf{w}(t)^T \mathbf{x}_n)$, please indicate in which class are samples $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ classified using the weight $\mathbf{w}(0)$. Which samples are not correctly classified based on this rule? **Note:** You have to compute the inner product $\mathbf{w}(0)^T \mathbf{x}_n$, $n = 1, 2, 3$, and see if it is greater or less than 0.

$\mathbf{w}(0)^T \mathbf{x}_1 = -\frac{1}{3} + 3 = \frac{8}{3} > 0$, therefore sample $\mathbf{x}_1$ is correctly classified using $\mathbf{w}(0)$.
\( w(0)^T x_2 = -1 + 2 = 1 > 0, \) therefore sample \( x_2 \) is incorrectly classified using \( w(0) \).
\( w(0)^T x_3 = -\frac{4}{3} + 1 = -\frac{1}{3} < 0, \) therefore sample \( x_3 \) is correctly classified using \( w(0) \).

(3) Using the weight update rule from the linear perceptron algorithm, please find the value of the new weight \( w(1) \). Find and plot the new line corresponding to weight \( w(1) \) in the 2D space.

\( w(1) = w(0) - x_2 = \left[ -\frac{1}{3}, 1 \right]^T - [3, 2]^T = \left[ -\frac{10}{3}, -1 \right]^T. \)

Note the “-” sign in the above equation, which is because the misclassified sample from \( w(0) \), \( x_2 \), belongs to Class 2 (\( y_2 = -1 \)).

The corresponding line is \(-\frac{10}{3} x - y = 0 \Rightarrow y = -\frac{10}{3} x \Rightarrow y = -3.33x\).

(4) Using the rule \( \text{sign}(w(t)^T x_n) \), please indicate in which class are samples \( x_1 \), \( x_2 \), and \( x_3 \) classified using the weight \( w(1) \). Which samples are not correctly classified based on this rule?

Note: You have to compute the inner product \( w(1)^T x_n \), \( n = 1, 2, 3 \), and see if it is greater or less than 0.

\( w(1)^T x_1 = -\frac{10}{3} - 3 = -\frac{19}{3} < 0, \) therefore sample \( x_1 \) is incorrectly classified using \( w(1) \).
\( w(1)^T x_2 = -10 - 2 = -12 < 0, \) therefore sample \( x_2 \) is correctly classified using \( w(1) \).
\( w(1)^T x_3 = -\frac{40}{3} - 1 = -\frac{43}{3} < 0, \) therefore sample \( x_3 \) is correctly classified using \( w(1) \).

(5) Using the weight update rule from the linear perceptron algorithm, please find the value of the new weight \( w(2) \). Find and plot the new line corresponding to weight \( w(2) \) in the 2D space. How many samples are correctly classified now?

Note: The update rule is \( w(t + 1) = w(t) + y_s x_s \), where \( x_s \) and \( y_s \in \{-1, 1\} \) is the feature and class label of misclassified sample \( s \).

\( w(2) = w(1) + x_1 = \left[ -\frac{10}{3}, -1 \right]^T + [1, 3]^T = \left[ -\frac{7}{3}, 2 \right]^T. \)

Note the “+” sign in the above equation, which is because the misclassified sample from \( w(1) \), \( x_1 \), belongs to Class 1 (\( y_1 = 1 \)).

The corresponding line is \(-\frac{7}{3} x + 2 y = 0 \Rightarrow y = \frac{7}{6} x \Rightarrow y = 1.16x\).

\( w(2)^T x_1 = -\frac{7}{3} + 6 = \frac{14}{3} > 0, \) therefore sample \( x_1 \) is correctly classified using \( w(2) \).
\( w(2)^T x_2 = -7 + 4 = -3 < 0, \) therefore sample \( x_2 \) is correctly classified using \( w(2) \).
\( w(2)^T x_3 = -\frac{28}{3} + 2 = -\frac{22}{3} < 0, \) therefore sample \( x_3 \) is correctly classified using \( w(2) \).

All samples are classified correctly based on \( w(2) \).