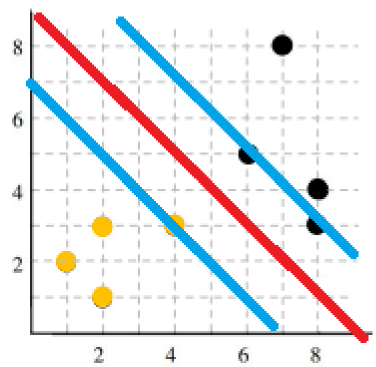
Bronco ID:	0 1 5 2 6 2 6 2 4
Last Name:	Nguyen
First Name:	Loc

1a. Decision boundary and parallel hyperplanes (hard margin)



Support vectors are Yellow(4,3) to satisfy w.x + b = -1 and Black (8,3) & (6,5) to satisfy w.x + b = 1

1b. Adding black circle at coordinate (7,5) won't affect previously learned decision boundary because the point lies on the upper parallel hyperplanes, classifying it as black circle while not violating the margin.

1c. Adding yellow circle at coordinate (4,2) won't affect previously learned decision boundary because the point lies below the lower parallel hyperplanes, classifying it as yellow circle while not violating the margin.

1d.The newly added black circle at coordinate (7,5) will be classified correctly because the point lies above the previously learned decision boundary that classify points as black circle.

1e. The newly added black circle at coordinate (6,4) will be classified correctly because the point lies above the previously learned decision boundary that classify points as black circle.

1f. The newly added yellow circle at coordinate (4,2) will be classified correctly because the point lies below the previously learned decision boundary that classify points as yellow circle

1g. The newly added yellow circle at coordinate (5,3) will be classified correctly because the point lies below the previously learned decision boundary that classify points as yellow circle

1h. The newly added black circle at coordinate (5,3) will be classified incorrectly because the point lies below the previously learned decision boundary that classify points yellow circle

1i. The newly added yellow circle at coordinate (6,4) will be classified incorrectly because the point lies above the previously learned decision boundary that classify points as black circle

1j. If a black circle is added as a training sample in the position (4,4) with $C = 1$ and considering soft margin formulation, the decision boundary will probably remain the same or with little changes since the penalty isn't significant. However, if $C = \infty$, the decision boundary will change because of the significant penalty due to the big penalizing parameter C since the previously learned decision boundary will incorrectly classify the new training instance of black circle at (4,4) to be yellow.

2a.

Positive labels are +1 and negative labels are -1

Support vectors are x = 0 and x = 1. From calculation in 2b, we can find the decision boundary.

$w.x + b = 0 \Rightarrow 2.x + (-1) = 0 \Rightarrow 2*x = 1 \Rightarrow x = ½$

Therefore, the decision boundary is at x = 0.5

2b.

To find b, we use x = 0. Thus, we derive $w.0 + b = -1 \Rightarrow b = -1$

To find w, we use (1, +) so x = 1 and y = 1. Thus, we derive $1(w.(1) + (-1)) = 1 \Rightarrow w – 1 = 1 \Rightarrow w = 2$

The width of the margin is $2 / ||w|| = 2 / 2 = 1$

2c.

To check the test sample of (-1.5, -), we use w.z + b < - 0.5 to classify

$\Rightarrow 2.(-1.5) + (-1) = 2*(-1.5) – 1 = -3 – 1 = -4 \Rightarrow -4 < 0$ so this instance is classified as negative

To check the test sample of (1.5, +), we use w.z + b > 0.5 to classify

$\Rightarrow 2.(1.5) + (-1) = 2*1.5 – 1 = 3 – 1 = 2 \Rightarrow 2 > 0$ so this instance is classified as positive

2d.

Since the point (0, -) remain, b = -1

To find w, we use (2, +) so x = 2 and y = 1. Thus, we derive $1(w.(2) + (-1)) = 1$

$\Rightarrow w*2 – 1 = 1 \Rightarrow 2w = 2 \Rightarrow w = 1$

The width of the margin is $2 / ||w|| = 2 / 1 = 2$

3a. $\Phi(A) = (1^2, 2^2, \sqrt{2}*1*2, \sqrt{2}*1, \sqrt{2}*2, 1) = (1, 4, 2\sqrt{2}, \sqrt{2}, 2\sqrt{2}, 1)$

3b. $\Phi(\mathrm{B}) = (2^2, 4^2, \sqrt{2}*2*4, \sqrt{2}*2, \sqrt{2}*4, 1) = (4, 16, 8\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}, 1)$

3c. $\Phi(A) \,.\, \Phi(\mathrm{B}) = 1*4 + 4*16 + 2\sqrt{2}*8\sqrt{2} + \sqrt{2}*2\sqrt{2} + 2\sqrt{2}*4\sqrt{2} + 1*1) = 4 + 64 + 32 + 4 + 16 +1 = 121$

3d. $\mathrm{K(A,B)} = ((1*2) + (2*4) + 1)^2 = (11)^2 = 121$

4. https://github.com/Skyhorizon2021/CS_4210/blob/main/Assignment3/svm.py

5a. Activation function is Heaviside(z), learning rate = 0.4, initial weight = 1

First Iteration:

$X_1 = 0$, $X_2 = 0$, $t = 0$, $X_0 = 1$, $w_0 = w_1 = w_2 = 1$

$Z = (1*0) + (1*0) + (1*1) = 1$

Since $y = 1$ and $t = 0$ and $\Delta w_0 = \eta(t-y)\, x_i$, we can derive

$\Delta w_0 = 0.4 * (-1) * 1 = -0.4$, $\Delta w_1 = 0.4 * (-1) * 0 = 0$, $\Delta w_2 = 0.4 * (-1) * 0 = 0$

➔ $w_0 = 1 - 0.4 = 0.6$, $w_1 = 1$, $w_2 = 1$

Second Iteration:

$X_1 = 0$, $X_2 = 1$, $t = 0$, $X_0 = 1$, $w_0 = 0.6$, $w_1 = w_2 = 1$

$Z = (1*0) + (1*0) + (0.6*1) = 0.6$

Since $y = 1$ and $t = 0$, we can derive

$\Delta w_0 = 0.4 * (-1) * 1 = -0.4$, $\Delta w_1 = 0.4 * (-1) * 0 = 0$, $\Delta w_2 = 0.4 * (-1) * 1 = -0.4$

➔ $w_0 = 0.6 - 0.4 = 0.2$, $w_1 = 1$, $w_2 = 1 - 0.4 = 0.6$

Third Iteration:

$X_1 = 1$, $X_2 = 0$, $t = 0$, $X_0 = 1$, $w_0 = 0.2$, $w_1 = 1$, $w_2 = 0.6$

$Z = (1*1) + (0.6*0) + (0.2*1) = 1 + 0.2 = 1.2$

Since $y = 1$ and $t = 0$, we can derive

$\Delta w_0 = 0.4 * (-1) * 1 = -0.4$, $\Delta w_1 = 0.4 * (-1) * 1 = -0.4$, $\Delta w_2 = 0.4 * (-1) * 0 = 0$

➔ $w_0 = 0.2 - 0.4 = -0.2$, $w_1 = 1 - 0.4 = 0.6$, $w_2 = 0.6$

Fourth Iteration:

$X_1 = 1$, $X_2 = 1$, $t = 1$, $X_0 = 1$, $w_0 = -0.2$, $w_1 = 0.6$, $w_2 = 0.6$

$Z = (0.6*1) + (0.6*1) + (-0.2*1) = 0.6 + 0.6 - 0.2 = 1$

Since $y = 1$ and $t = 1$, we can derive

$\Delta w_0 = 0.4 * (0) * 1 = 0$, $\Delta w_1 = 0.4 * (0) * 1 = 0$, $\Delta w_2 = 0.4 * (0) * 1 = 0$

➔ $w_0 = -0.2$, $w_1 = 0.6$, $w_2 = 0.6$

Solution Table

	x_0	x_1	x_2	w_0	w_1	w_2	t	z(net)	y	t-y	Δw_0	Δw_1	Δw_2
1	1	0	0	1	1	1	0	1	1	-1	-0.4	0	0
2	1	0	1	0.6	1	1	0	1.6	1	-1	-0.4	0	-0.4
3	1	1	0	0.2	1	0.6	0	1.2	1	-1	-0.4	-0.4	0
4	1	1	1	-0.2	0.6	0.6	1	1	1	0	0	0	0
Final				-0.2	0.6	0.6							

5b. Activation function is Heaviside(z), learning rate = 0.1, initial weight = 0

First Iteration:

$X_1 = 0$, $t = 1$, $X_0 = 1$, $w_0 = w_1 = 0$

$Z = (0*0) + (0*1) = 0$

Since $y = 0$ and $t = 1$ and $\Delta w_0 = \eta(t-y)\, x_i$, we can derive

$\Delta w_0 = 0.1 * (1) * 1 = 0.1$, $\Delta w_1 = 0.1 * (1) * 0 = 0$

➔ $w_0 = 0 + 0.1 = 0.1$, $w_1 = 0$

Second Iteration:

$X_1 = 1$, $t = 0$, $X_0 = 1$, $w_0 = 0.1$, $w_1 = 0$

$Z = (0*1) + (0.1*1) = 0 + 0.1 = 0.1$

Since $y = 1$ and $t = 0$, we can derive

$\Delta w_0 = 0.1 * (-1) * 1 = -0.1$, $\Delta w_1 = 0.1 * (-1) * 1 = -0.1$

➔ $w_0 = 0.1 - 0.1 = 0$, $w_1 = 0 – 0.1 = -0.1$

Solution Table

	x_0	x_1	w_0	w_1	t	z(net)	y	t-y	Δw_0	Δw_1
1	1	0	0	0	1	0	0	1	0.1	0
2	1	1	0.1	0	0	0.1	1	-1	-0.1	-0.1
Final			0	-0.1						