CMP-6002B - Machine Learning

Dr Gavin Cawley

Lecture 4 - Linear Classifiers

Perceptron Training Rule

Algorithm PerceptronTraining(DataSet X, t)
Returns LinearModel *w*

initialise $\boldsymbol{w}$ to random values
initialise learning rate η
do
    for i=1 to n
        $y_i = \psi(\boldsymbol{w}, \boldsymbol{x}_i)$
        for j=1 to m
            $\Delta w_j \leftarrow 0.5\eta(t_i - y_i)x_{ij}$
            $w_j \leftarrow w_j + \Delta w_j$
while (Stopping($\boldsymbol{t}, \boldsymbol{y}$) == **false**)

return $\boldsymbol{w}$

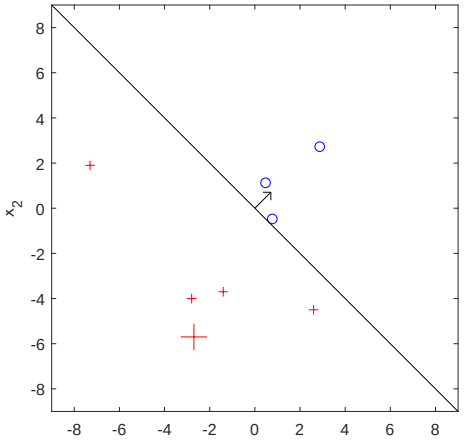
The Perceptron Rule

- Basic idea, cycle through training patterns, if a pattern is currently misclassified add/subtract the input vector to the weights.
  - This shifts the discriminant to make it more likely to classify that pattern correctly next time.
- η is the learning rate which moderates how much the weights are changed on each iteration.
  - The discriminant is unaffected by the magnitude of $\boldsymbol{w}^T\boldsymbol{x}$, so let $\eta = 1$.
- Stopping conditions vary, simplest is to stop when there is no change in $\boldsymbol{y}$ or error is zero.
  - A complete pass through the dataset is made without modifying the weights.

Perceptron Example - Step #1

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} +1.0 \\ +1.0 \end{bmatrix}$$

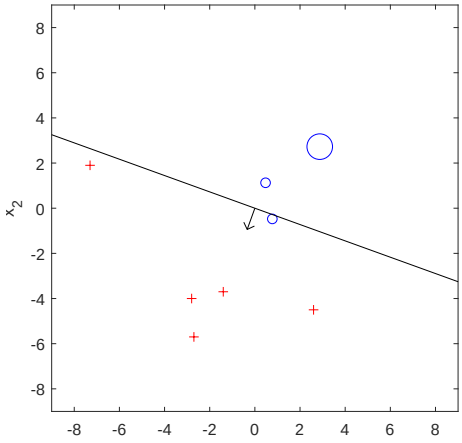


$$\begin{aligned} y_1 &= \psi(1.0 \times -2.7 + 1.0 \times -5.7) = \psi(-8.4) = -1 : \text{ WRONG!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [+1.0 + 1.0] + 0.5 \times 1 \times (+1.0 - -1.0)[-2.7 - 5.7] \\ &= [-1.7 - 4.7] \end{aligned}$$

Perceptron Example - Step #2

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -1.7 \\ -4.7 \end{bmatrix}$$



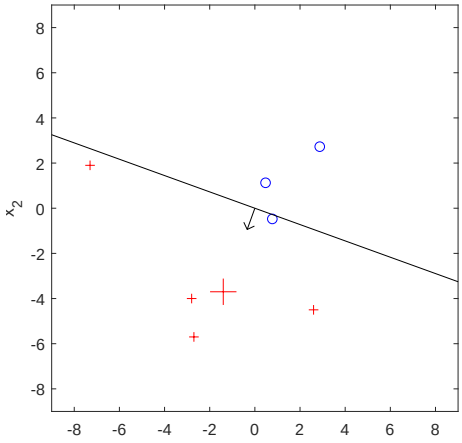
$$y_2 = \psi(-1.7 \times 2.9 + -4.7 \times 2.7) = \psi(-17.6) = -1: \text{ RIGHT!}$$

$$\begin{aligned} \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-1.7 - 4.7] + 0.5 \times 1 \times (-1.0 - -1.0)[+2.9 + 2.7] \\ &= [-1.7 - 4.7] \end{aligned}$$

Perceptron Example - Step #3

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -1.7 \\ -4.7 \end{bmatrix}$$



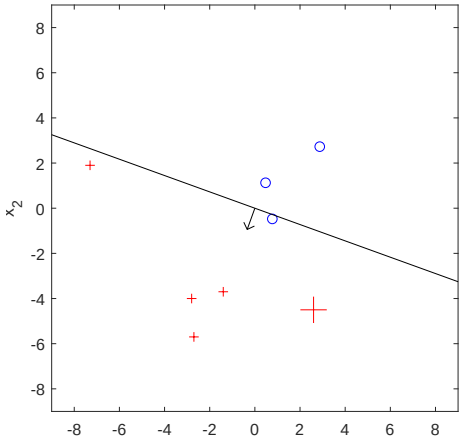
$$y_3 = \psi(-1.7 \times -1.4 + -4.7 \times -3.7) = \psi(+19.8) = 1: \text{ RIGHT!}$$

$$\begin{aligned} \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-1.7 - 4.7] + 0.5 \times 1 \times (+1.0 - +1.0)[-1.4 - 3.7] \\ &= [-1.7 - 4.7] \end{aligned}$$

Perceptron Example - Step #4

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -1.7 \\ -4.7 \end{bmatrix}$$



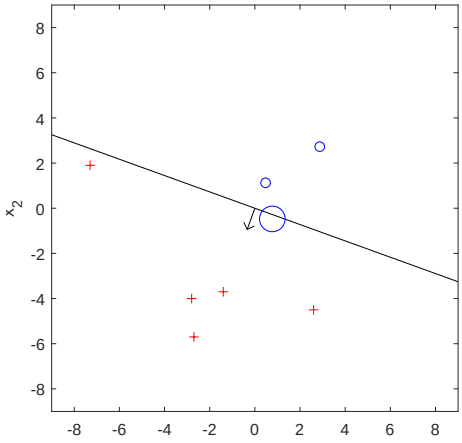
$$y_4 = \psi(-1.7 \times 2.6 + -4.7 \times -4.5) = \psi(+16.7) = 1: \text{ RIGHT!}$$

$$\begin{aligned} \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-1.7 - 4.7] + 0.5 \times 1 \times (+1.0 - +1.0)[+2.6 - 4.5] \\ &= [-1.7 - 4.7] \end{aligned}$$

Perceptron Example - Step #5

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -1.7 \\ -4.7 \end{bmatrix}$$



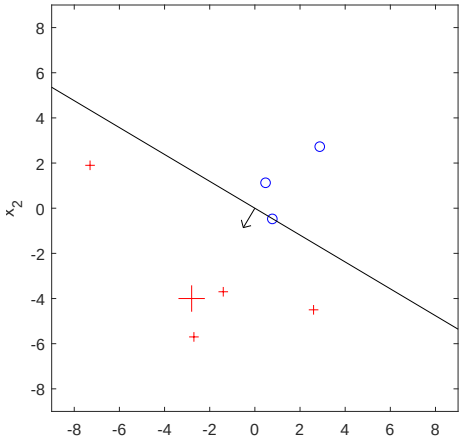
$$y_5 = \psi(-1.7 \times 0.8 + -4.7 \times -0.5) = \psi(+1.0) = 1: \text{ WRONG!}$$

$$\begin{aligned} \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-1.7 - 4.7] + 0.5 \times 1 \times (-1.0 - +1.0)[+0.8 - 0.5] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #6

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

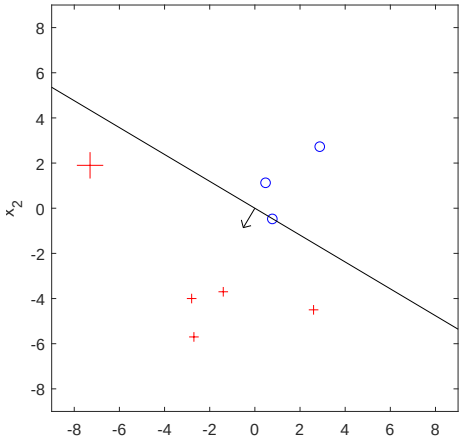


$$\begin{aligned} y_6 &= \psi(-2.5 \times -2.8 + -4.2 \times -4.0) = \psi(+23.8) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (+1.0 - +1.0)[-2.8 - 4.0] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #7

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

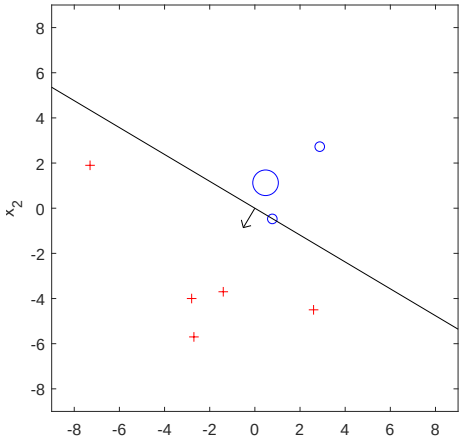


$$\begin{aligned} y_7 &= \psi(-2.5 \times -7.3 + -4.2 \times 1.9) = \psi(+10.3) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (+1.0 - +1.0)[-7.3 + 1.9] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #8

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

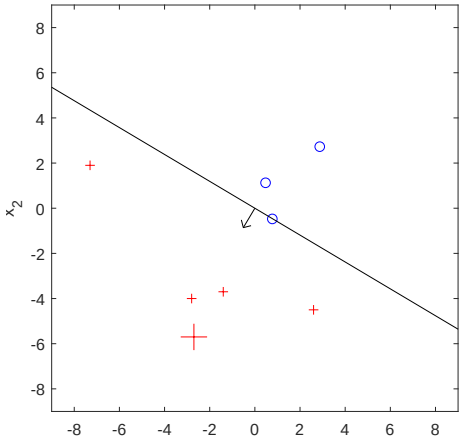


$$\begin{aligned} y_8 &= \psi(-2.5 \times 0.5 + -4.2 \times 1.1) = \psi(-5.9) = -1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (-1.0 - -1.0)[+0.5 + 1.1] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #9

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

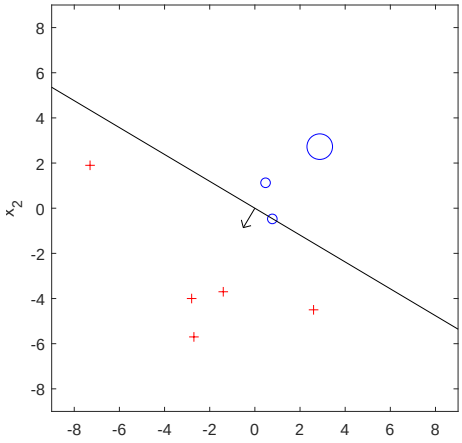


$$\begin{aligned} y_1 &= \psi(-2.5 \times -2.7 + -4.2 \times -5.7) = \psi(+30.7) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (+1.0 - +1.0)[-2.7 - 5.7] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #10

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

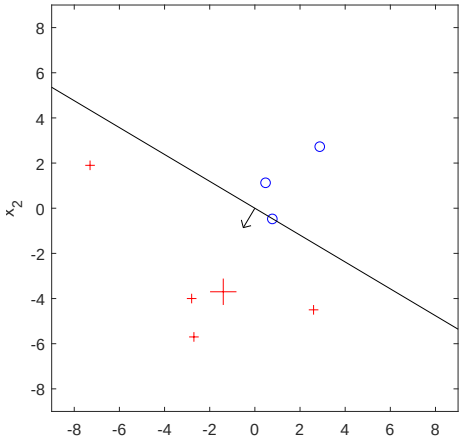


$$\begin{aligned} y_2 &= \psi(-2.5 \times 2.9 + -4.2 \times 2.7) = \psi(-18.6) = -1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (-1.0 - -1.0)[+2.9 + 2.7] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #11

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

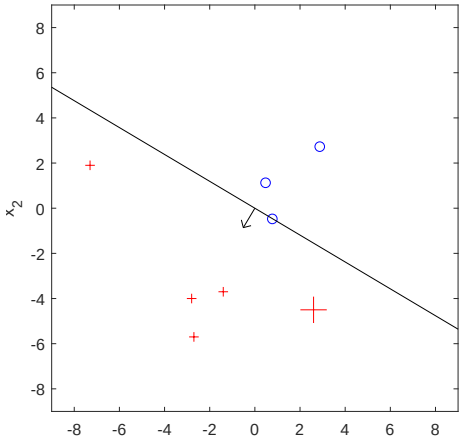


$$\begin{aligned} y_3 &= \psi(-2.5 \times -1.4 + -4.2 \times -3.7) = \psi(+19.0) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (+1.0 - +1.0)[-1.4 - 3.7] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #12

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

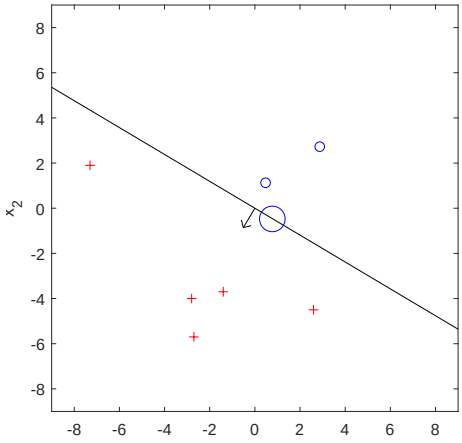


$$\begin{aligned} y_4 &= \psi(-2.5 \times 2.6 + -4.2 \times -4.5) = \psi(+12.4) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (+1.0 - +1.0)[+2.6 - 4.5] \\ &= [-2.5 - 4.2] \end{aligned}$$

Perceptron Example - Step #13

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -2.5 \\ -4.2 \end{bmatrix}$$

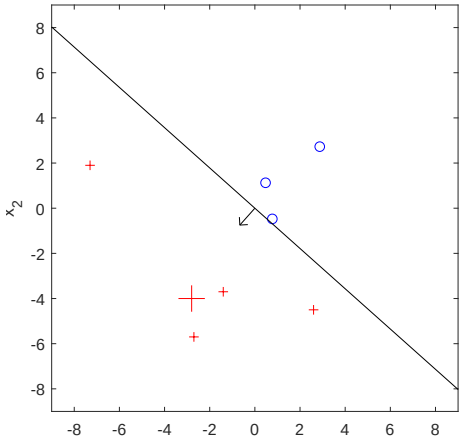


$$\begin{aligned} y_5 &= \psi(-2.5 \times 0.8 + -4.2 \times -0.5) = \psi(+0.1) = 1: \text{ WRONG!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-2.5 - 4.2] + 0.5 \times 1 \times (-1.0 - +1.0)[+0.8 - 0.5] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #14

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$

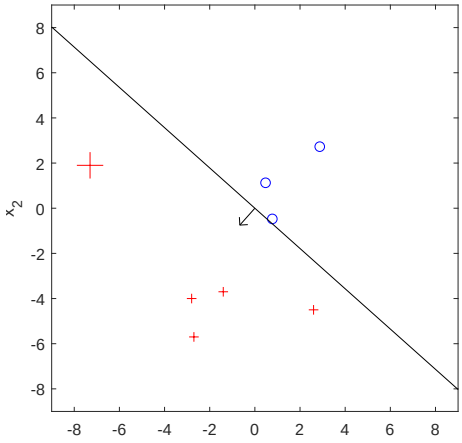


$$\begin{aligned} y_6 &= \psi(-3.3 \times -2.8 + -3.7 \times -4.0) = \psi(+24.0) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (+1.0 - +1.0)[-2.8 - 4.0] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #15

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$

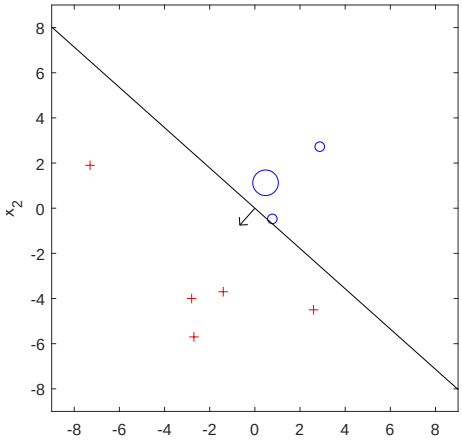


$$\begin{aligned} y_7 &= \psi(-3.3 \times -7.3 + -3.7 \times 1.9) = \psi(+17.1) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (+1.0 - +1.0)[-7.3 + 1.9] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #16

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$

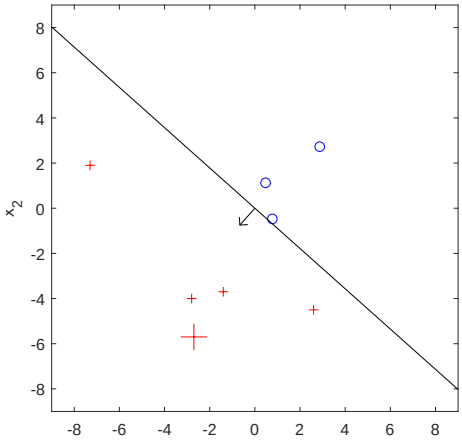


$$\begin{aligned} y_8 &= \psi(-3.3 \times 0.5 + -3.7 \times 1.1) = \psi(-5.7) = -1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (-1.0 - -1.0)[+0.5 + 1.1] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #17

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$

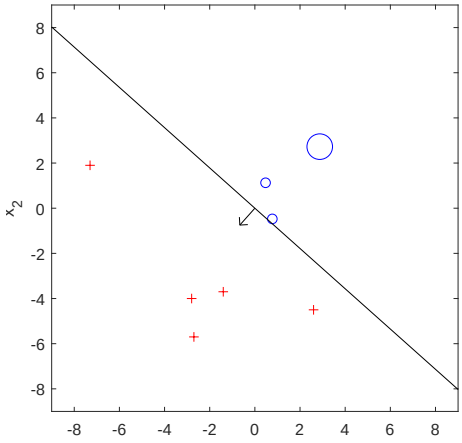


$$\begin{aligned} y_1 &= \psi(-3.3 \times -2.7 + -3.7 \times -5.7) = \psi(+30.0) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (+1.0 - +1.0)[-2.7 - 5.7] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #18

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$

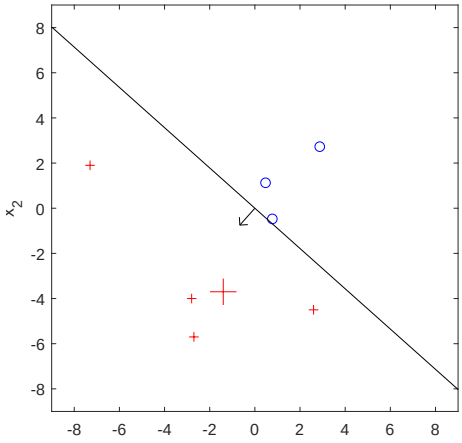


$$\begin{aligned} y_2 &= \psi(-3.3 \times 2.9 + -3.7 \times 2.7) = \psi(-19.6) = -1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (-1.0 - -1.0)[+2.9 + 2.7] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #19

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$

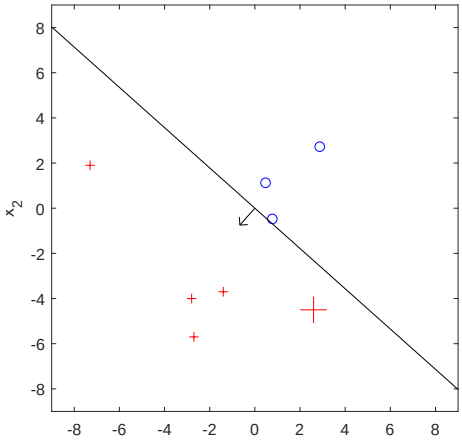


$$\begin{aligned} y_3 &= \psi(-3.3 \times -1.4 + -3.7 \times -3.7) = \psi(+18.3) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (+1.0 - +1.0)[-1.4 - 3.7] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #20

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$

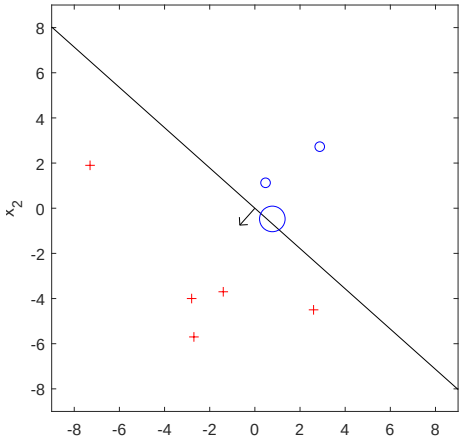


$$\begin{aligned} y_4 &= \psi(-3.3 \times 2.6 + -3.7 \times -4.5) = \psi(+8.1) = 1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (+1.0 - +1.0)[+2.6 - 4.5] \\ &= [-3.3 - 3.7] \end{aligned}$$

Perceptron Example - Step #21

x_1	x_2	t
-2.7	-5.7	+1.0
+2.9	+2.7	-1.0
-1.4	-3.7	+1.0
+2.6	-4.5	+1.0
+0.8	-0.5	-1.0
-2.8	-4.0	+1.0
-7.3	+1.9	+1.0
+0.5	+1.1	-1.0

$$\boldsymbol{w} = \begin{bmatrix} -3.3 \\ -3.7 \end{bmatrix}$$



$$\begin{aligned} y_5 &= \psi(-3.3 \times 0.8 + -3.7 \times -0.5) = \psi(-0.8) = -1: \text{ RIGHT!} \\ \boldsymbol{w} &\leftarrow \boldsymbol{w} + 0.5\eta(t_i - y_i)\boldsymbol{x}_j \\ &= [-3.3 - 3.7] + 0.5 \times 1 \times (-1.0 - -1.0)[+0.8 - 0.5] \\ &= [-3.3 - 3.7] \end{aligned}$$