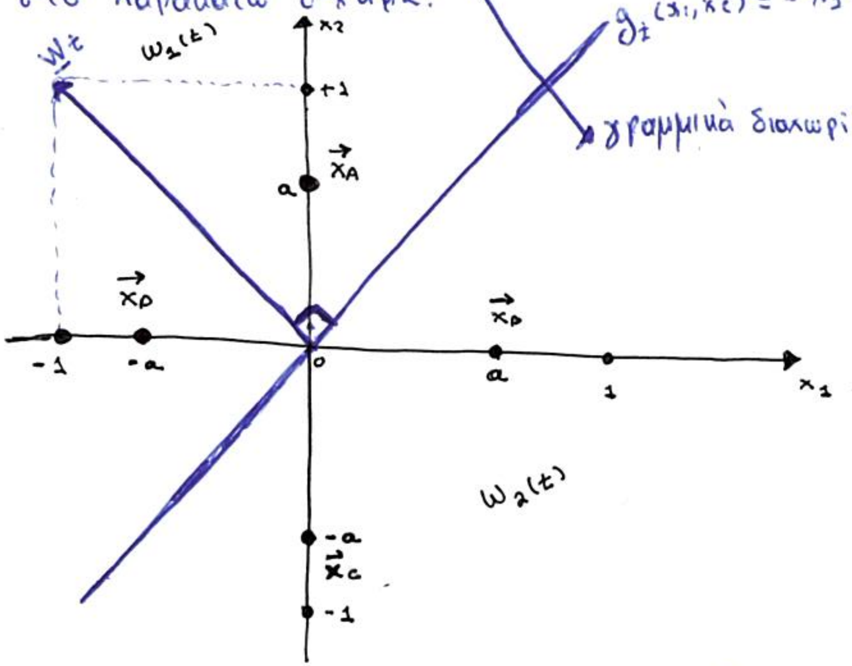
#1

Θεωρείστε το πρόβλημα δυαδικής ταξινόμησης μεταξύ των κλάσεων ω_1 και ω_2 που παρουσιάζεται στο παρακάτω σχήμα:



Έστω $X = \{\vec{x}_A, \vec{x}_B, \vec{x}_C, \vec{x}_D\}$ με $\vec{x}_j \in \mathbb{R}^2$ για $j \in \{A,B,C,D\}$ όπου:

$$\vec{x}_A = \begin{bmatrix} 0 \\ a \end{bmatrix}, \quad \vec{x}_B = \begin{bmatrix} a \\ 0 \end{bmatrix}, \quad \vec{x}_C = \begin{bmatrix} 0 \\ -a \end{bmatrix}, \quad \vec{x}_D = \begin{bmatrix} -a \\ 0 \end{bmatrix}$$

Με $\boxed{a > 1/2}$. Επίσης, έχουμε ότι:

$P(\omega_1) = P(\omega_2)$ and (★) $W_t = \begin{bmatrix} -1 \\ +1 \end{bmatrix}$.

(★) Αν κατά την t-οστή επανάληψη του αλγορίθμου Perceptron υπάρχουν 2 λάθος ταξινομημένα πρότυπα, $|Y_t| = 2$, τότε κατά την $(t+1)$-οστή επανάληψη η ενημέρωση της γραμμής διάκρισης οδηγεί στην ορθή ταξινόμηση όλων των προτύπων.

(I): Οι πιθανές επιλογές του συνόλου Y_t είναι οι παρακάτω:

(1): $Y_t = \{\vec{x}_A, \vec{x}_B\}$

(2): $Y_t = \{\vec{x}_A, \vec{x}_C\}$

(3): $Y_t = \{\vec{x}_A, \vec{x}_D\}$

(4): $Y_t = \{\vec{x}_B, \vec{x}_C\}$

(5): $Y_t = \{\vec{x}_B, \vec{x}_D\}$

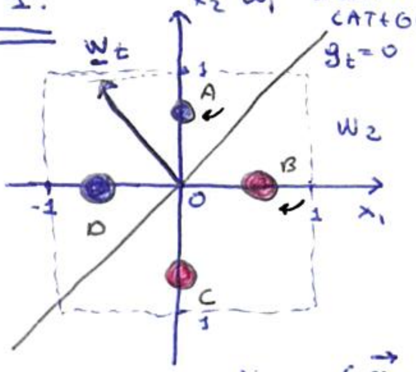
(6): $Y_t = \{\vec{x}_C, \vec{x}_D\}$

Let ● indicate patterns from ω_1

Let ● indicate patterns from ω_2.

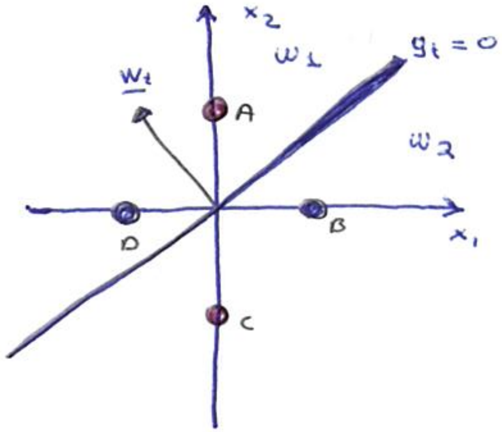
Case 1:

[CLASSIFIER-BASED CATEGORIZATION]



$Y_t = \{ \vec{x}_A , \vec{x}_B \}$

[TRUE CATEGORIZATION]

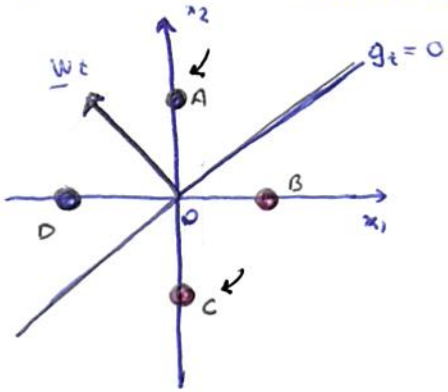


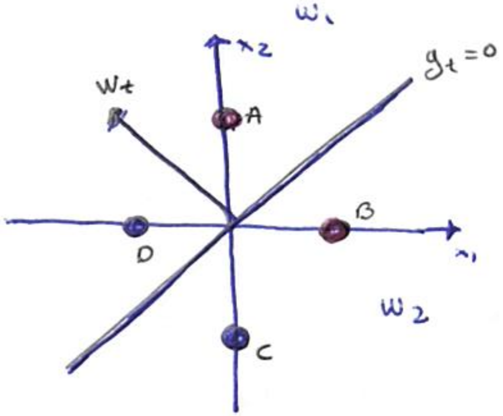
(*) According to the true categorization of patterns classes ω_1 and ω_2 are not linearly separable. Thus, this case is not possible.

[IMPOSSIBLE]

Case 2:

[CLASSIFIER-BASED CATEGORIZATION]





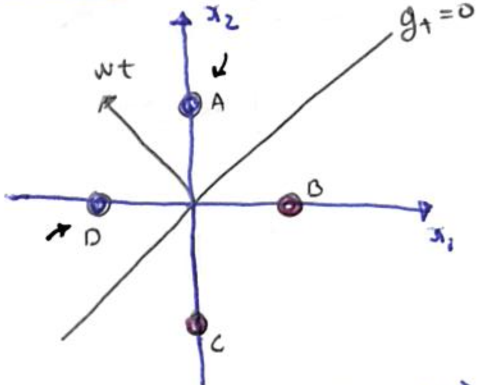
$Y_t = \{ \vec{x}_A , \vec{x}_C \}$

(*) This is a possible scenario, since according to the true categorization of patterns, they are linearly separable.

[POSSIBLE]

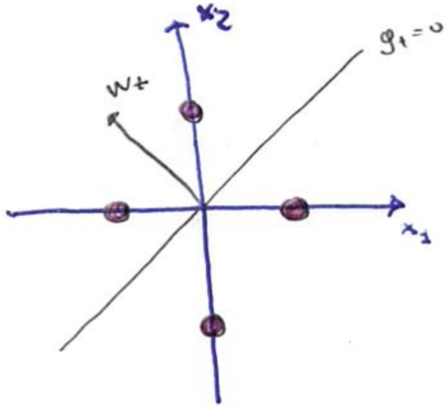
Case 3:

[CLASSIFIER-BASED CATEGORIZATION]



$Y_t = \{ \vec{x}_A \ \vec{x}_D \}$

[TRUE CATEGORIZATION]

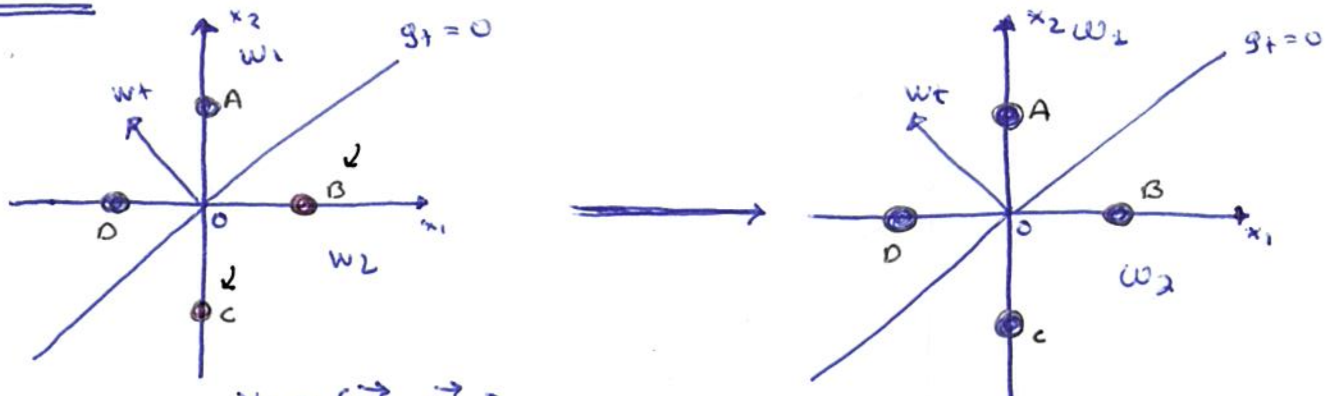


(*) This is not a possible scenario since according to the true categorization of patterns, all patterns would be assigned to same class. But, it is true that $\ell(\omega_1) = \ell(\omega_2)$.

[IMPOSSIBLE]

#3

Case 4:

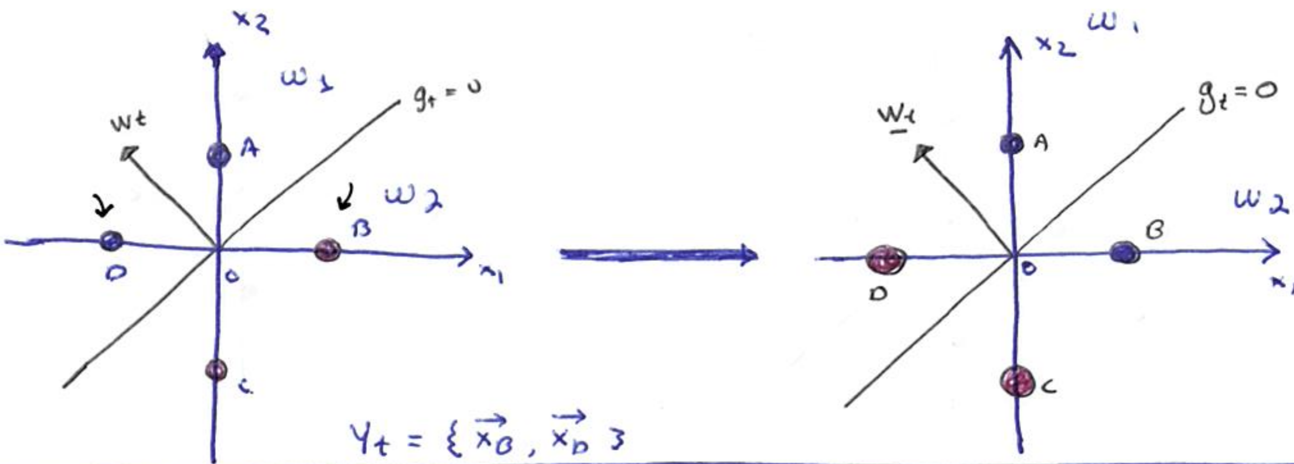


$Y_t = \{ \vec{x}_B, \vec{x}_C \}$

(A) This case is symmetric to case (3). It is not possible, since, it suggests that all patterns belong to ω_1. This is not true since $P(\omega_1) = P(\omega_2)$.

[IMPOSSIBLE]

CASE 5:

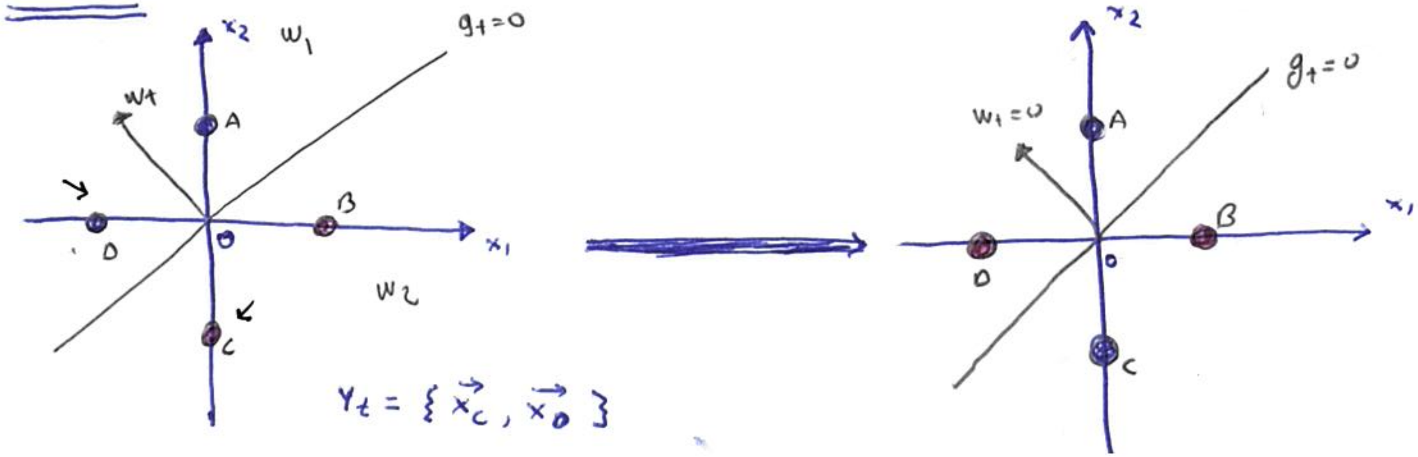


$Y_t = \{ \vec{x}_B, \vec{x}_D \}$

(A) This case is symmetric to case (2)

[POSSIBLE]

CASE 6:



$Y_t = \{ \vec{x}_C, \vec{x}_D \}$

(A) This case is symmetric to case (1). It is not possible since it suggests that the true categorization of patterns is not linear.

CASE 2: $Y_t = \{\vec{x}_A, \vec{x}_C\}$ where $\vec{x}_A \in \omega_2$ and $\vec{x}_c \in \omega_1$.

Thus, we have that:
$$\begin{cases} \delta(\vec{x}_A) = +1 \\ \delta(\vec{x}_c) = -1 \end{cases}$$

We also have that the extended vectors $\underline{x}_A'$ and $\underline{x}_c'$ are as follows:

$$\underline{x}_A' = \begin{bmatrix} 0 \\ a \\ 1 \end{bmatrix} \text{ and } \underline{x}_c' = \begin{bmatrix} 0 \\ -a \\ 1 \end{bmatrix} \text{ and } \underline{w}_t' = \begin{bmatrix} -1 \\ +1 \\ 0 \end{bmatrix}$$

The correction vector, for the $(t+1)$ step of the perceptron algorithm will be:

$$\underline{c}_{t+1} = \sum_{\underline{x} \in Y_t} \delta(\underline{x}) \circ x' \Rightarrow \underline{c}_{t+1} = \delta(\underline{x}_A) \circ \underline{x}_A' + \delta(\underline{x}_c) \circ \underline{x}_c' \Rightarrow$$

$$\underline{c}_{t+1} = \begin{bmatrix} 0 \\ a \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -a \\ 1 \end{bmatrix} \Rightarrow \underline{c}_{t+1} = \begin{bmatrix} 0 \\ 2a \\ 0 \end{bmatrix}$$

Thus, yields that: $\underline{W}_{t+1} = \underline{W}_t - \underline{c}_{t+1} \Rightarrow$

$$\underline{W}_{t+1} = \begin{bmatrix} -1 \\ +1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 2a \\ 0 \end{bmatrix} \Rightarrow \boxed{\underline{W}_{t+1} = \begin{bmatrix} -1 \\ 1-2a \\ 0 \end{bmatrix}}$$

Thus, we have that $\boxed{g_{t+1}(x_1, x_2) = -x_1 + (1-2a)x_2 = 0}$

Let's see how this new line classifies all patterns.

$g_{t+1}(\vec{x}_A) = g_{t+1}(0, a) = a \circ (1-2a) < 0$ (correct) $[\vec{x}_A \in \omega_2]$

$g_{t+1}(\vec{x}_B) = g_{t+1}(a, 0) = -a < 0$ (correct) $[\vec{x}_B \in \omega_2]$

$g_{t+1}(\vec{x}_C) = g_{t+1}(0, -a) = -a(1-2a) > 0$ (correct) $[\vec{x}_c \in \omega_1]$

$g_{t+1}(\vec{x}_D) = g_{t+1}(-a, 0) = a > 0$ (correct) $[\vec{x}_D \in \omega_1]$

Case (5): $Y_t = \{ \vec{x}_B , \vec{x}_D \}$ where $\vec{x}_B \in \omega_1$ and $\vec{x}_D \in \omega_2$.

Thus, we have that: $\begin{cases} \delta(\vec{x}_B) = -1 \\ \delta(\vec{x}_D) = +1 \end{cases}$

We also have that the extended vectors $\underline{x}_B'$ and $\underline{x}_D'$ are as follows:

$$\underline{x}_B' = \begin{bmatrix} a \\ 0 \\ 1 \end{bmatrix} \text{ and } \underline{x}_D' = \begin{bmatrix} -a \\ 0 \\ 1 \end{bmatrix} \text{ with } \underline{W}_t' = \begin{bmatrix} -1 \\ +1 \\ 0 \end{bmatrix}.$$

The correction vector for the (t+1) step of the perceptron algorithm will be:

$$\underline{C}_{t+1} = \sum_{x \in Y_t} \delta(x) \cdot x' \Rightarrow \underline{C}_{t+1} = \delta(\underline{x}_B) \cdot \underline{x}_B' + \delta(\underline{x}_D) \cdot \underline{x}_D' \Rightarrow$$

$$\underline{C}_{t+1} = - \begin{bmatrix} a \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -a \\ 0 \\ 1 \end{bmatrix} \Rightarrow \underline{C}_{t+1} = \begin{bmatrix} -2a \\ 0 \\ 0 \end{bmatrix}$$

This, yields that: $\underline{W}_{t+1} = \underline{W}_t - \underline{C}_{t+1} \Rightarrow$

$$\underline{W}_{t+1} = \begin{bmatrix} -1 \\ +1 \\ 0 \end{bmatrix} - \begin{bmatrix} -2a \\ 0 \\ 0 \end{bmatrix} \Rightarrow \boxed{\underline{W}_{t+1} = \begin{bmatrix} -1+2a \\ 1 \\ 0 \end{bmatrix}}$$

Thus, we have that: $\boxed{g_{t+1}(x_1, x_2) = (2a-1)x_1 + x_2 = 0}$

Let's see how this line classifies all patterns:

$g_{t+1}(\vec{x}_A) = g_{t+1}(0, a) = a > 0$ [correct] $(\vec{x}_A \in \omega_1)$

$g_{t+1}(\vec{x}_B) = g_{t+1}(a, 0) = a(2a-1) > 0$ [correct] $(\vec{x}_B \in \omega_1)$

$g_{t+1}(\vec{x}_C) = g_{t+1}(0, -a) = -a < 0$ [correct] $(\vec{x}_C \in \omega_2)$

$g_{t+1}(\vec{x}_D) = g_{t+1}(-a, 0) = -a(2a-1) < 0$ [correct] $(\vec{x}_D \in \omega_2)$