



Τεχνητή Νοημοσύνη και Μηχανική Μάθηση

Κωνσταντίνος Καραμανής

The University of Texas at Austin & Archimedes/Athena RC

constantine@utexas.edu

<https://caramanis.github.io/>



Νευρωνικά Δίκτυα

Μία απείρως πλούσια
οικογένεια
αλγορίθμων που εάν
σχεδιαστούν σωστά,
ταιριάζουν με πολλές
εφαρμογές

Νευρωνικά Δίκτυα

- Που χρησιμοποιούνται τα νευρωνικά δίκτυα
- Παντού! Μηχανική όραση (computer vision), επεξεργασία φυσικής γλώσσας (natural language processing), γενετική τεχνητή νοημοσύνη (generative AI), ενισχυτική μάθηση (reinforcement learning)
- Alexnet, Resnet, ChatGPT, Claude, Dall-E, Gemini, Bard, Llama, CLIP, AlphaGO...

Νευρωνικά Δίκτυα

- Θα μάθουμε πως να τα χτίζουμε
- Πως να τα χρησιμοποιούμε
- Πως να τα εκπαιδεύουμε

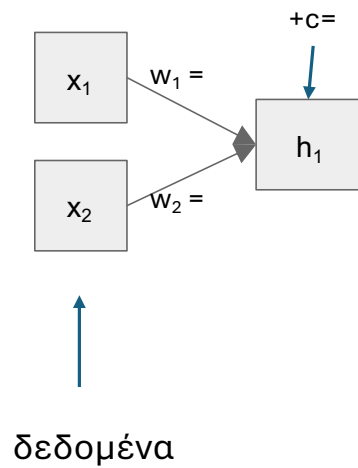
Νευρωνικά Δίκτυα

- Θα μάθουμε πως να τα χτίζουμε
- Πως να τα χρησιμοποιούμε
- Πως να τα εκπαιδεύουμε

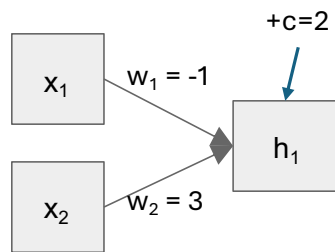
Σε αυτήν την διάλεξη:

- Γραμμικά επίπεδα
- Softmax

Γραμμικά Επίπεδα: Fully Connected



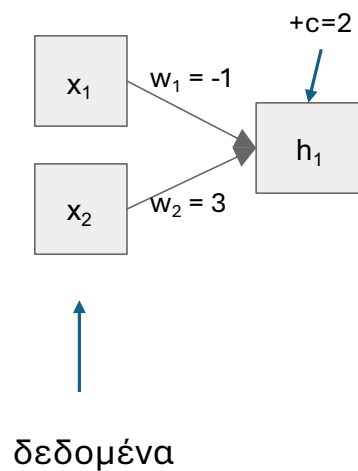
Γραμμικά Επίπεδα: Fully Connected



x_1	x_2	h_1
-1	2	
2	1	
1	0	

↑
δεδομένα

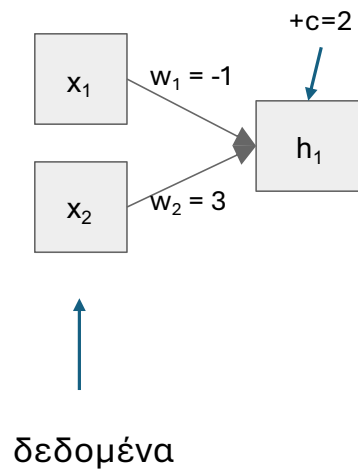
Γραμμικά Επίπεδα: Fully Connected



x_1	x_2	h_1
-1	2	
2	1	
1	0	

$$x_1 \cdot w_1 + x_2 \cdot w_2 + c$$

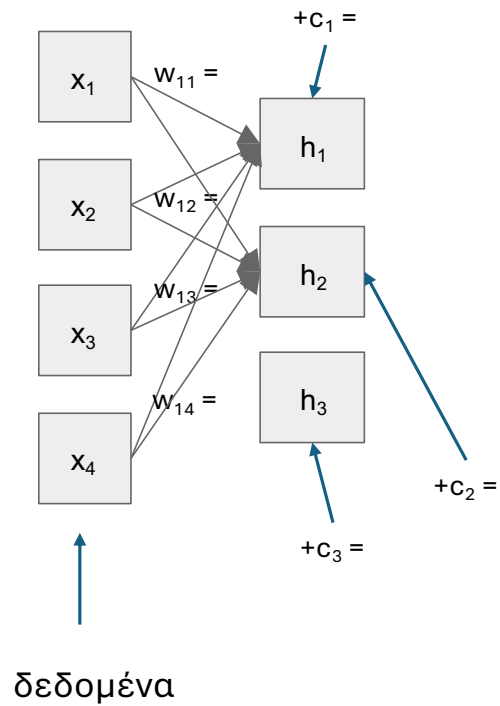
Γραμμικά Επίπεδα: Fully Connected



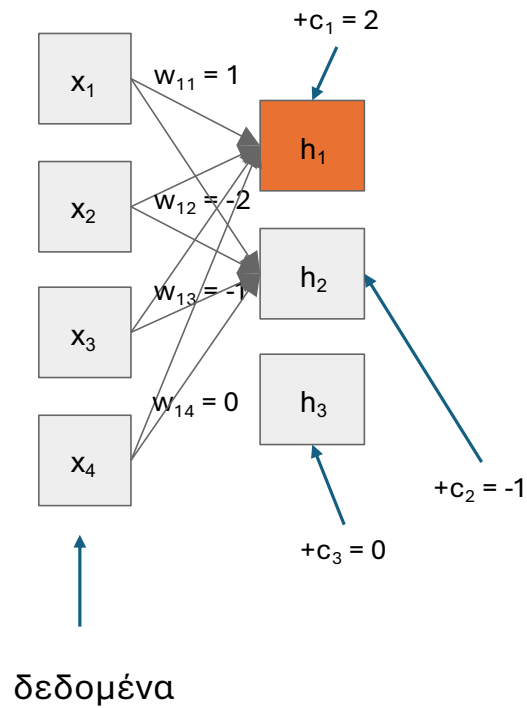
x_1	x_2	h_1
-1	2	9
2	1	3
1	0	1

$$x_1 \cdot w_1 + x_2 \cdot w_2 + c = -1 \cdot -1 + 2 \cdot 3 + 2 = 9$$

Γραμμικά Επίπεδα: Fully Connected

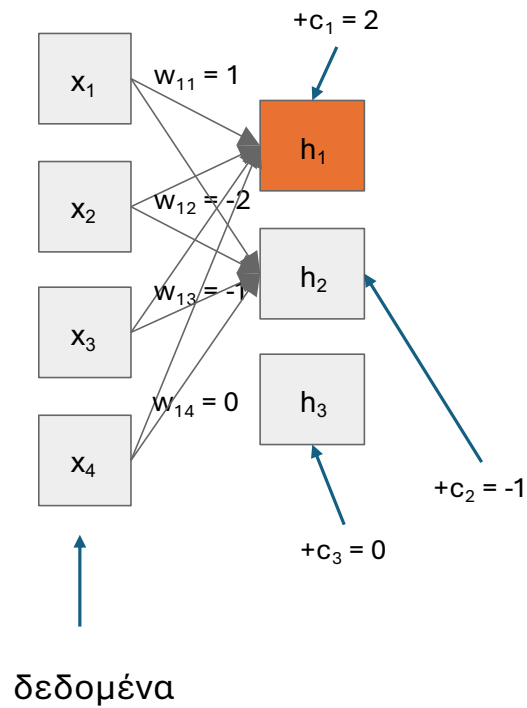


Γραμμικά Επίπεδα: Fully Connected



x_1	x_2	x_3	x_4	h_1
0	2	1	4	
-1	-1	2	-3	
0	3	0	2	

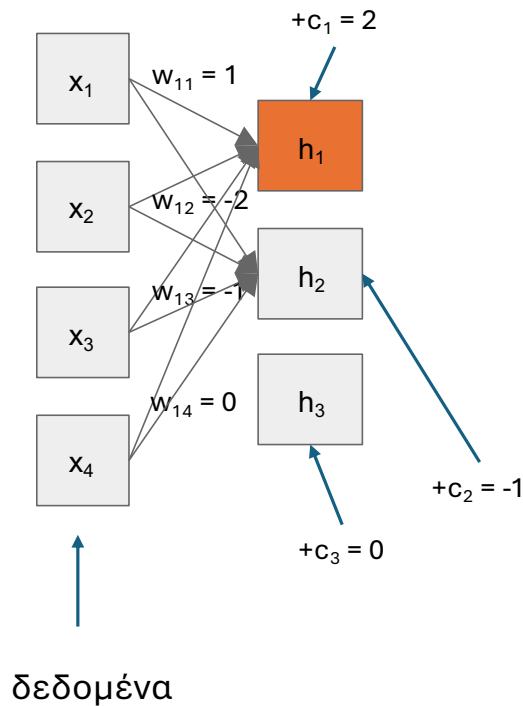
Γραμμικά Επίπεδα: Fully Connected



x_1	x_2	x_3	x_4	h_1
0	2	1	4	
-1	-1	2	-3	
0	3	0	2	

$$x_1 \cdot w_{11} + x_2 \cdot w_{12} + x_3 \cdot w_{13} + x_4 \cdot w_{14} + c_1$$

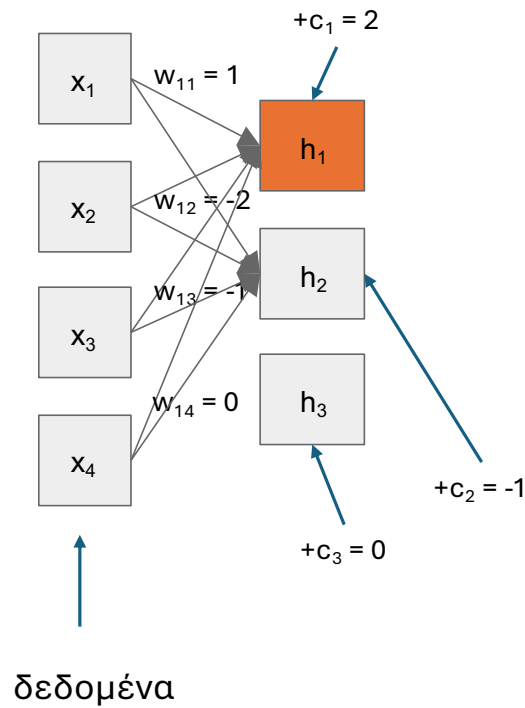
Γραμμικά Επίπεδα: Fully Connected



x_1	x_2	x_3	x_4	h_1
0	2	1	4	-3
-1	-1	2	-3	
0	3	0	2	

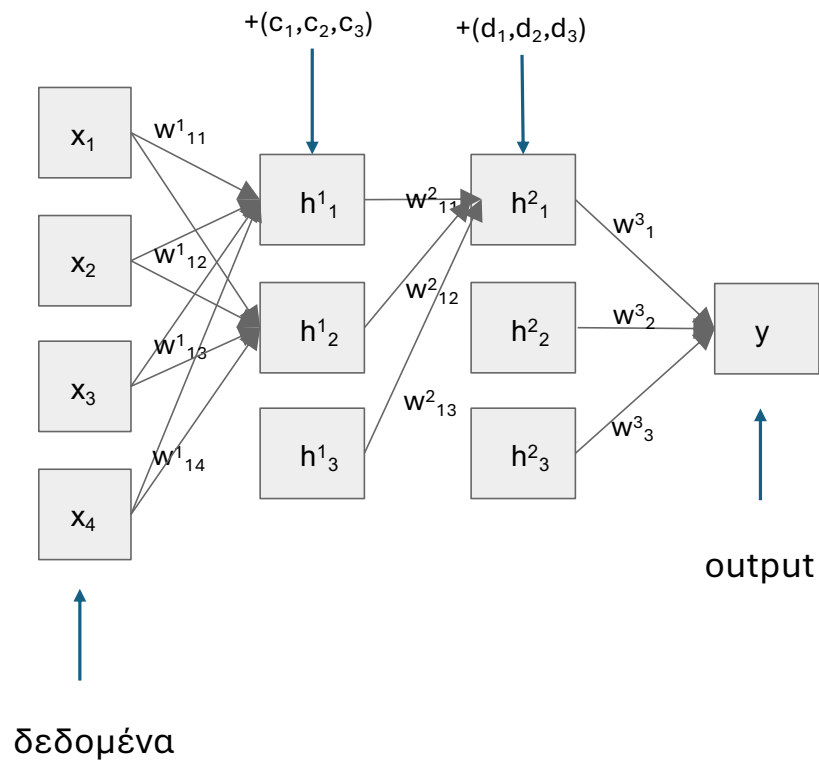
$$\begin{aligned} & x_1 \cdot w_{11} + x_2 \cdot w_{12} + x_3 \cdot w_{13} + x_4 \cdot w_{14} + c_1 \\ &= 0 \cdot 1 + 2 \cdot -2 + 1 \cdot -1 + 4 \cdot 0 + 2 \\ &= -3 \end{aligned}$$

Γραμμικά Επίπεδα: Fully Connected



x_1	x_2	x_3	x_4	h_1
0	2	1	4	-3
-1	-1	2	-3	1
0	3	0	2	-4

Γραμμικά Επίπεδα: Fully Connected

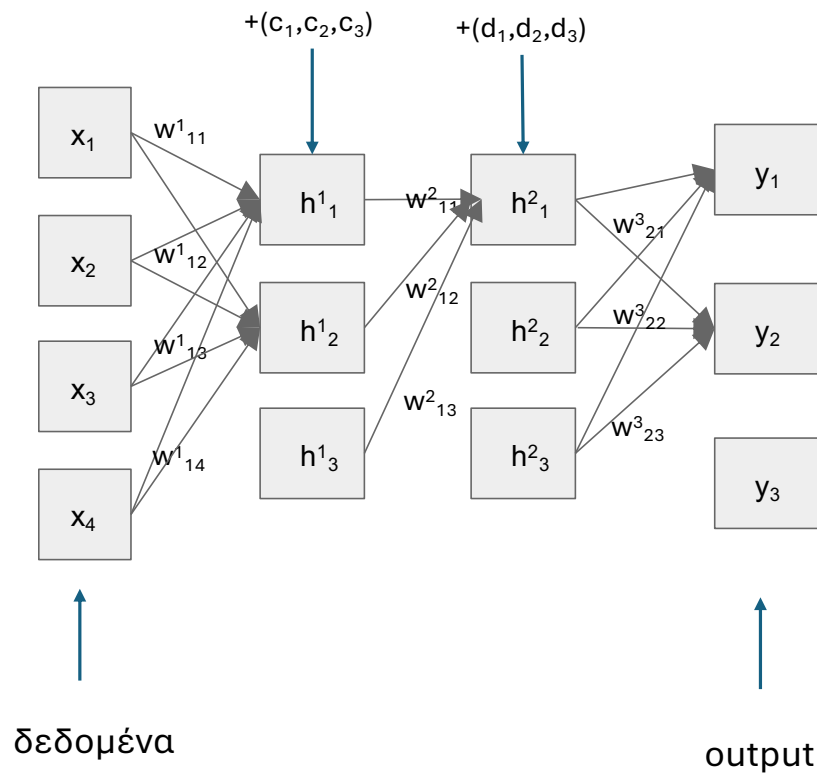


Για παλινδρόμηση/regression έχουμε ένα output

Παράδειγμα:

- Sentiment Detection
- Pose estimation

Γραμμικά Επίπεδα: Fully Connected



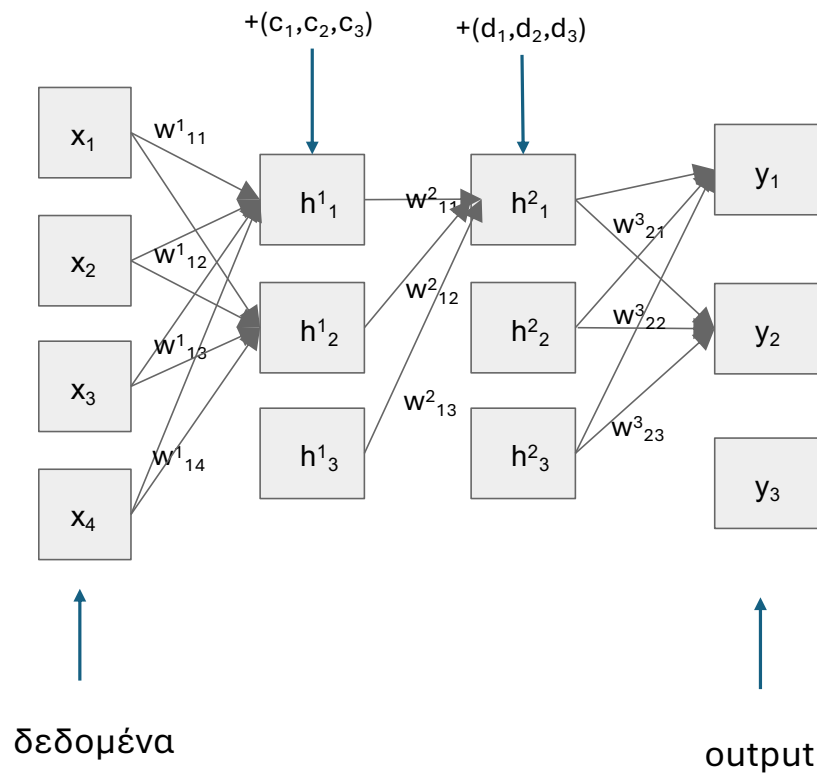
Για ταξινόμηση χρειαζόμαστε 2 ή παραπάνω output

Παράδειγμα:

- Ταξινόμηση καρκίνου μαστού: malignant / benign / normal



Γραμμικά Επίπεδα: Fully Connected



Για ταξινόμηση χρειαζόμαστε 2 ή παραπάνω output

Παράδειγμα:

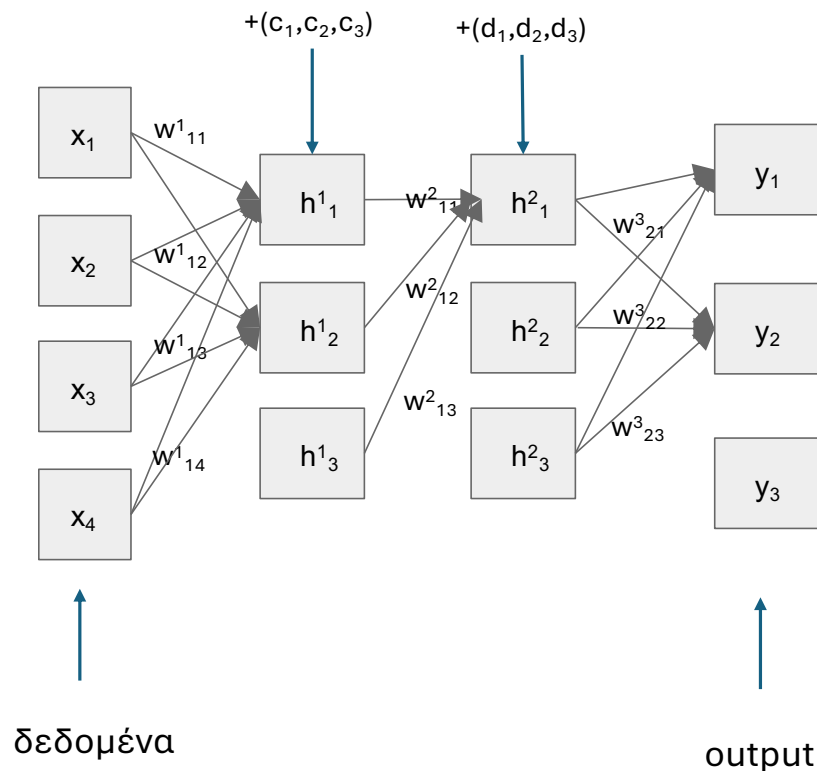
- Ταξινόμηση καρκίνου μαστού: malignant / benign / normal

$y_1 = 1, y_2 = 0, y_3 = 0$: το x ανήκει στην κατηγορία 1

$y_1 = 0, y_2 = 1, y_3 = 0$: το x ανήκει στην κατηγορία 2

$y_1 = 0, y_2 = 0, y_3 = 1$: το x ανήκει στην κατηγορία 3

Γραμμικά Επίπεδα: Fully Connected



Για ταξινόμηση χρειαζόμαστε 2 ή παραπάνω output

Παράδειγμα:

- Ταξινόμηση καρκίνου μαστού: malignant / benign / normal

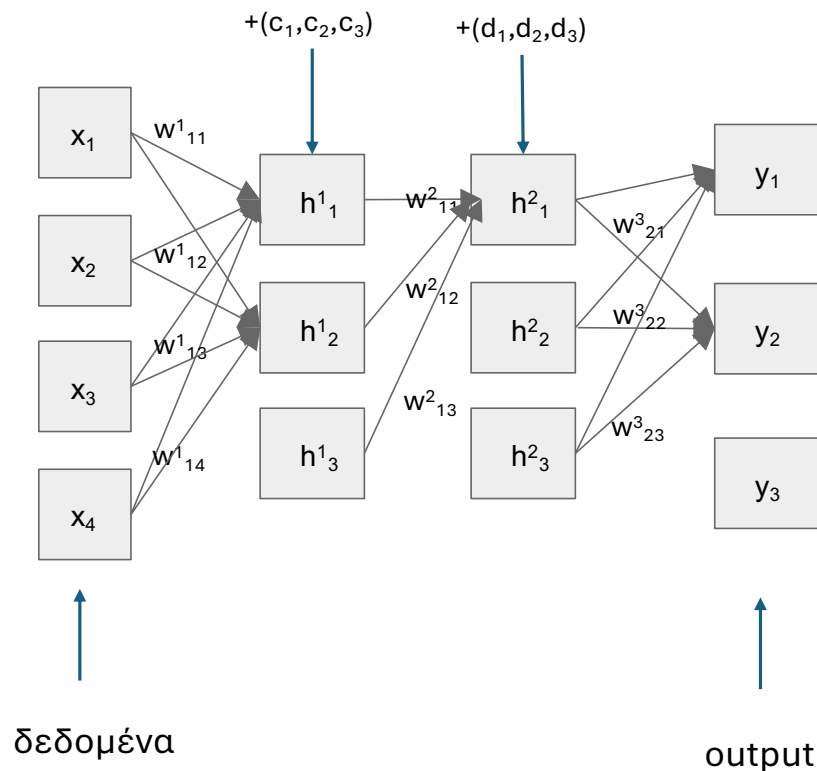
$y_1 = 1, y_2 = 0, y_3 = 0$: το x ανήκει στην κατηγορία **1**

$y_1 = 0, y_2 = 1, y_3 = 0$: το x ανήκει στην κατηγορία **2**

$y_1 = 0, y_2 = 0, y_3 = 1$: το x ανήκει στην κατηγορία **3**

$y_1 = 0,94, y_2 = 0,02, y_3 = 0,04$:

Γραμμικά Επίπεδα: Fully Connected



Για ταξινόμηση χρειαζόμαστε 2 ή παραπάνω output

Παράδειγμα:

- Ταξινόμηση καρκίνου μαστού: malignant / benign / normal

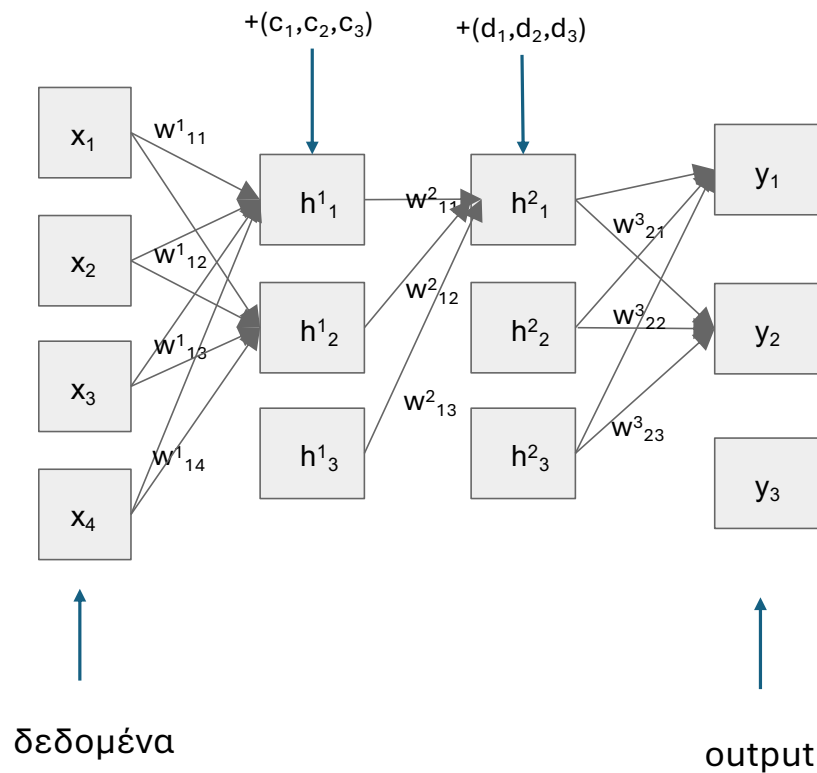
$y_1 = 1, y_2 = 0, y_3 = 0$: το x ανήκει στην κατηγορία 1 **1**

$y_1 = 0, y_2 = 1, y_3 = 0$: το x ανήκει στην κατηγορία 2 **2**

$y_1 = 0, y_2 = 0, y_3 = 1$: το x ανήκει στην κατηγορία 3 **3**

$y_1 = 0,94, y_2 = 0,02, y_3 = 0,04$: το x ανήκει στην κατηγορία 1
σχεδόν σίγουρα (94%)

Γραμμικά Επίπεδα: Fully Connected



Για ταξινόμηση χρειαζόμαστε 2 ή παραπάνω output

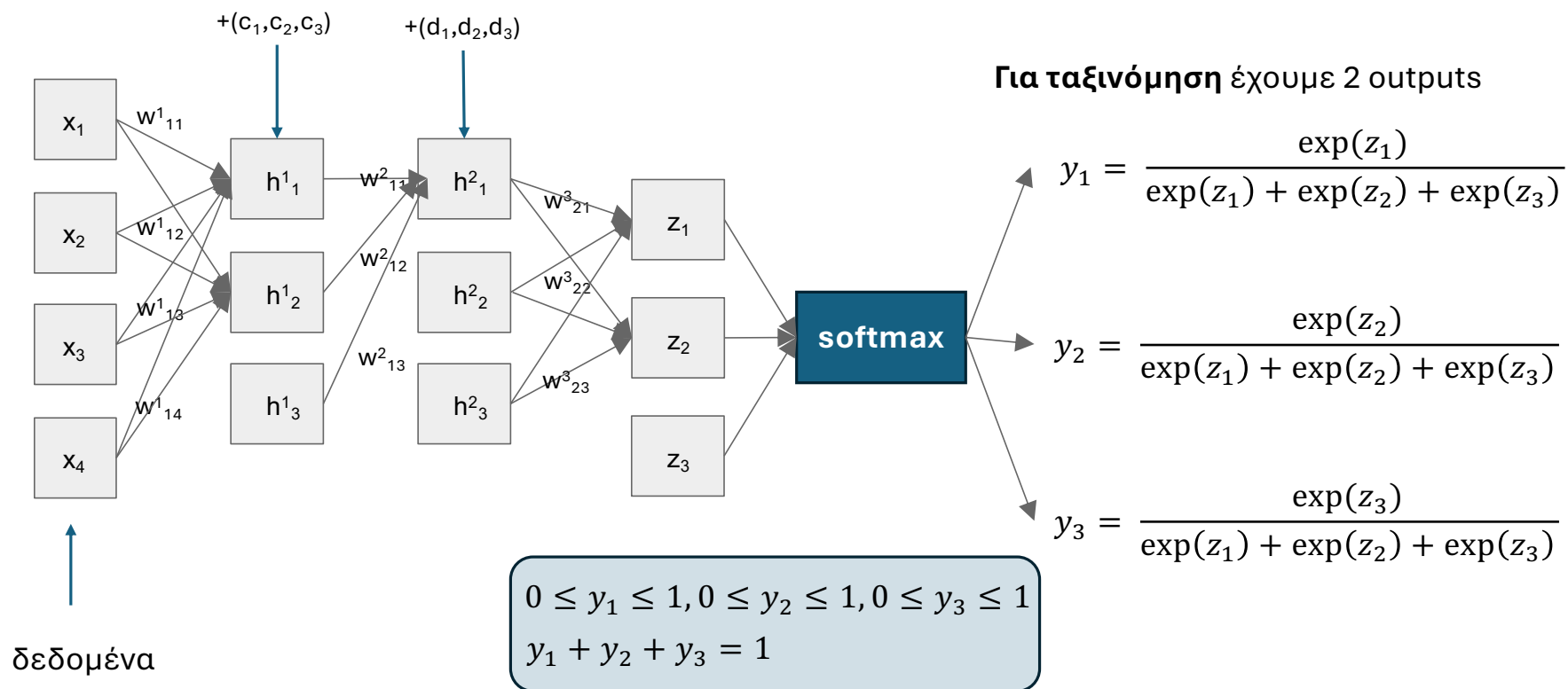
$y_1 = 0,94, y_2 = 0,02, y_3 = 0,04$: το x ανήκει στην κατηγορία 1 σχεδόν σίγουρα (94%)

Οπότε απαιτούμε:

$$0 \leq y_1 \leq 1, 0 \leq y_2 \leq 1, 0 \leq y_3 \leq 1$$
$$y_1 + y_2 + y_3 = 1$$

Πως θα το πετύχουμε αυτό;

Γραμμικά Επίπεδα: Fully Connected



Softmax: Παραδείγματα

$$(z_1, z_2) = (-1, 3) \quad (y_1, y_2) = \text{softmax}(z_1, z_2):$$

$$y_1 = \frac{\exp(z_1)}{\exp(z_1) + \exp(z_2)} = 0,018$$

$$y_2 = \frac{\exp(z_2)}{\exp(z_1) + \exp(z_2)} = 0,982$$

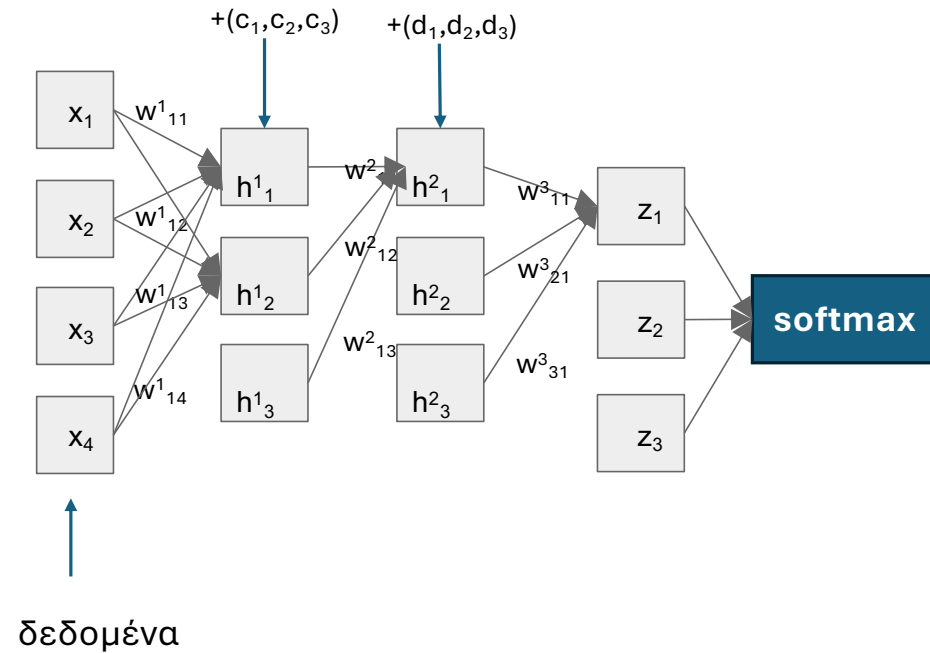
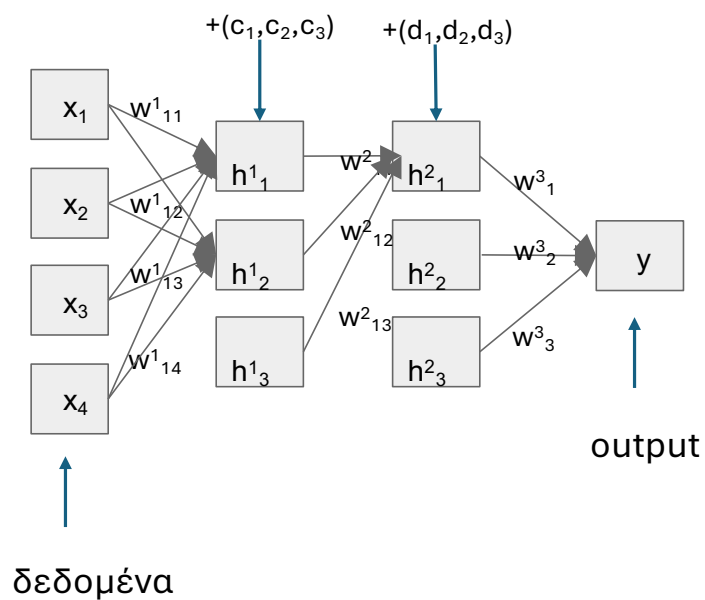
$$(z_1, z_2, z_3) = (3, 3, 3, -2) \quad (y_1, y_2, y_3) = \text{softmax}(z_1, z_2, z_3):$$

$$y_1 = \frac{\exp(z_1)}{\exp(z_1) + \exp(z_2) + \exp(z_3)} = 0,424$$

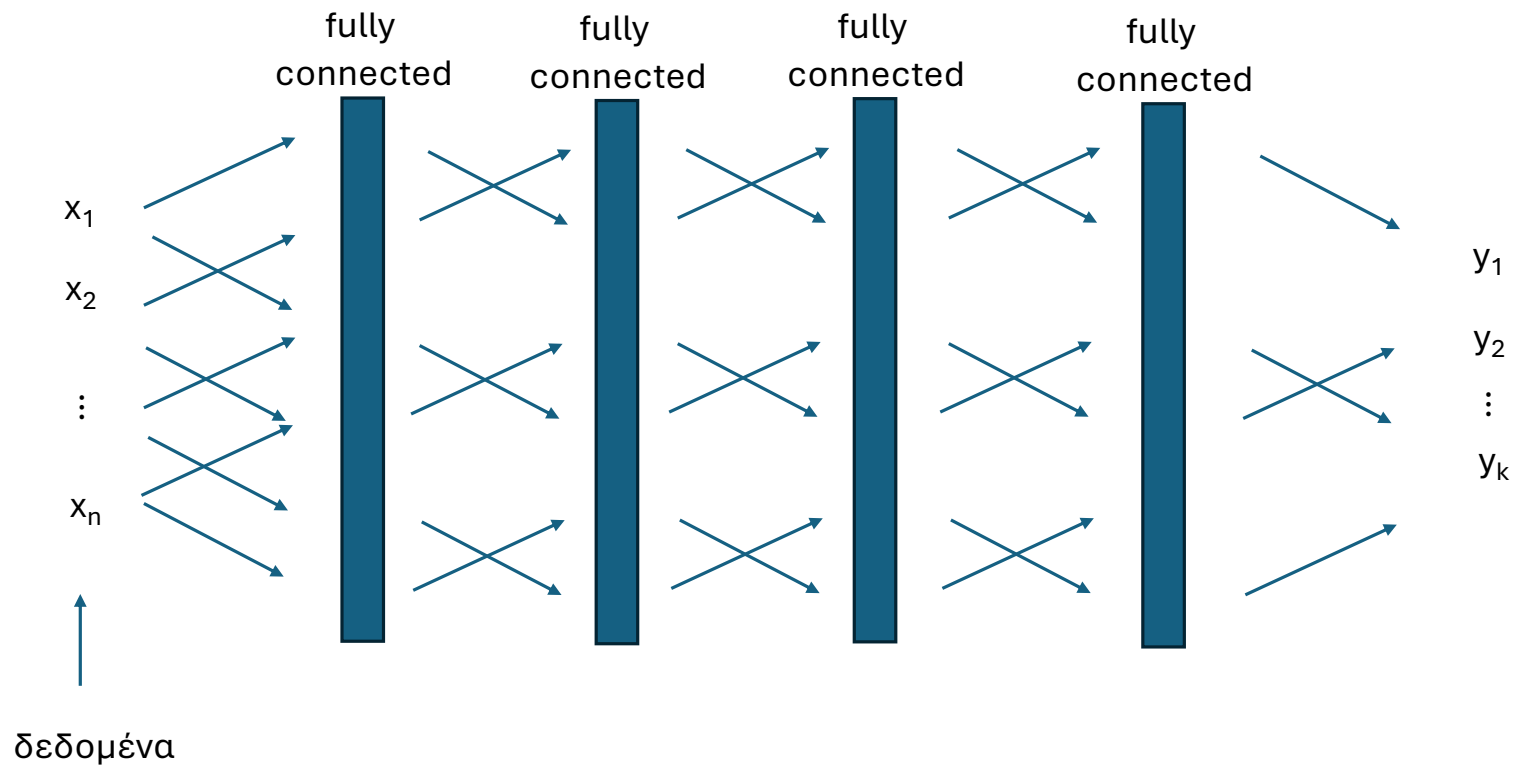
$$y_2 = \frac{\exp(z_2)}{\exp(z_1) + \exp(z_2) + \exp(z_3)} = 0,573$$

$$y_3 = \frac{\exp(z_3)}{\exp(z_1) + \exp(z_2) + \exp(z_3)} = 0,0028$$

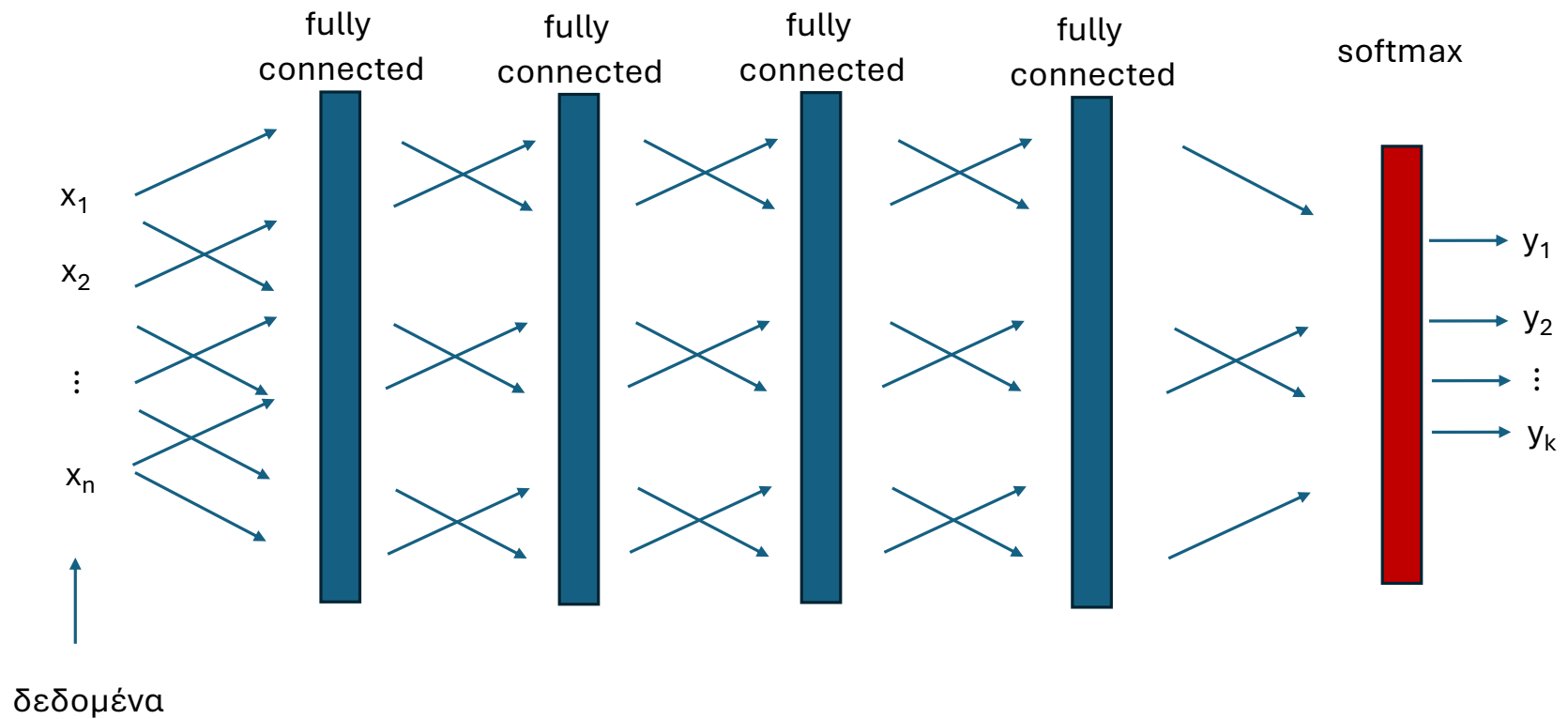
Fully Connected + Softmax



Fully Connected



Fully Connected + SoftMax



Αρκούν τα Γραμμικά Επίπεδα;

