



Παραγωγική Τεχνητή Νοημοσύνη: Generative AI

Κωνσταντίνος Καραμανής

The University of Texas at Austin & Archimedes/Athena RC

constantine@utexas.edu

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





Ας θυμηθούμε τα
προηγούμενα...



Εσωτερικό Γινόμενο

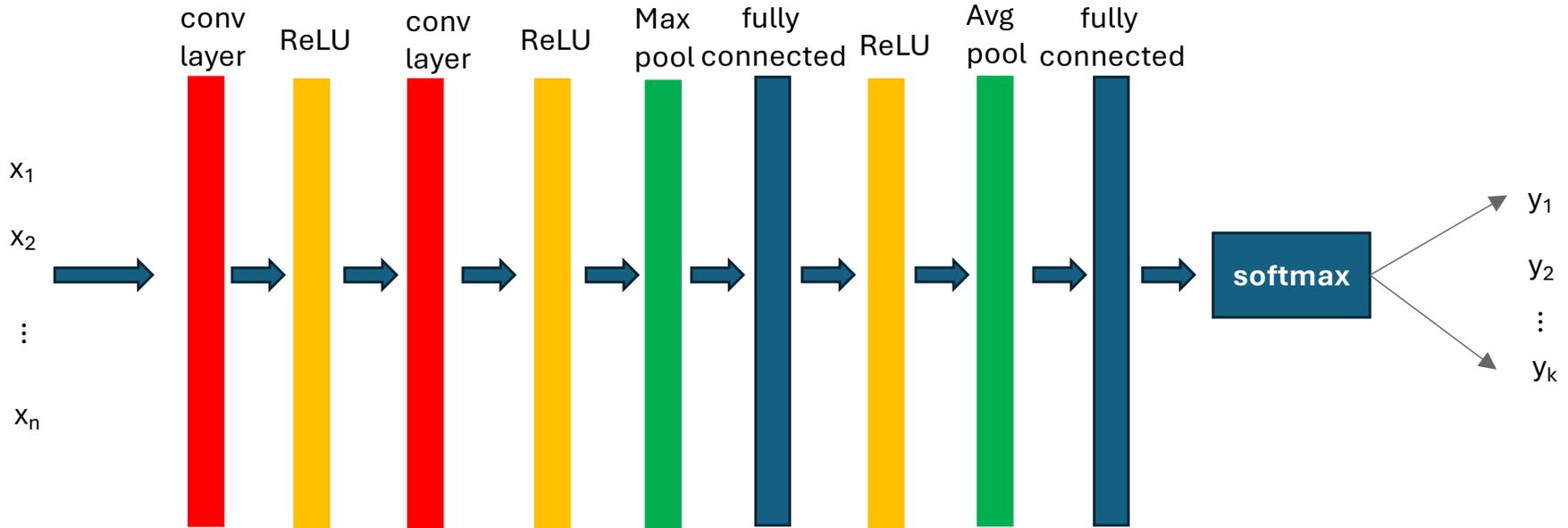
Το Εσωτερικό Γινόμενο $\langle W, X \rangle$ είναι μεγάλο όταν το W και το X :

- «Μοιάζουν»
- Είναι κοντά στον χώρο
- Σχηματίζουν μικρή γωνία $\angle WX$

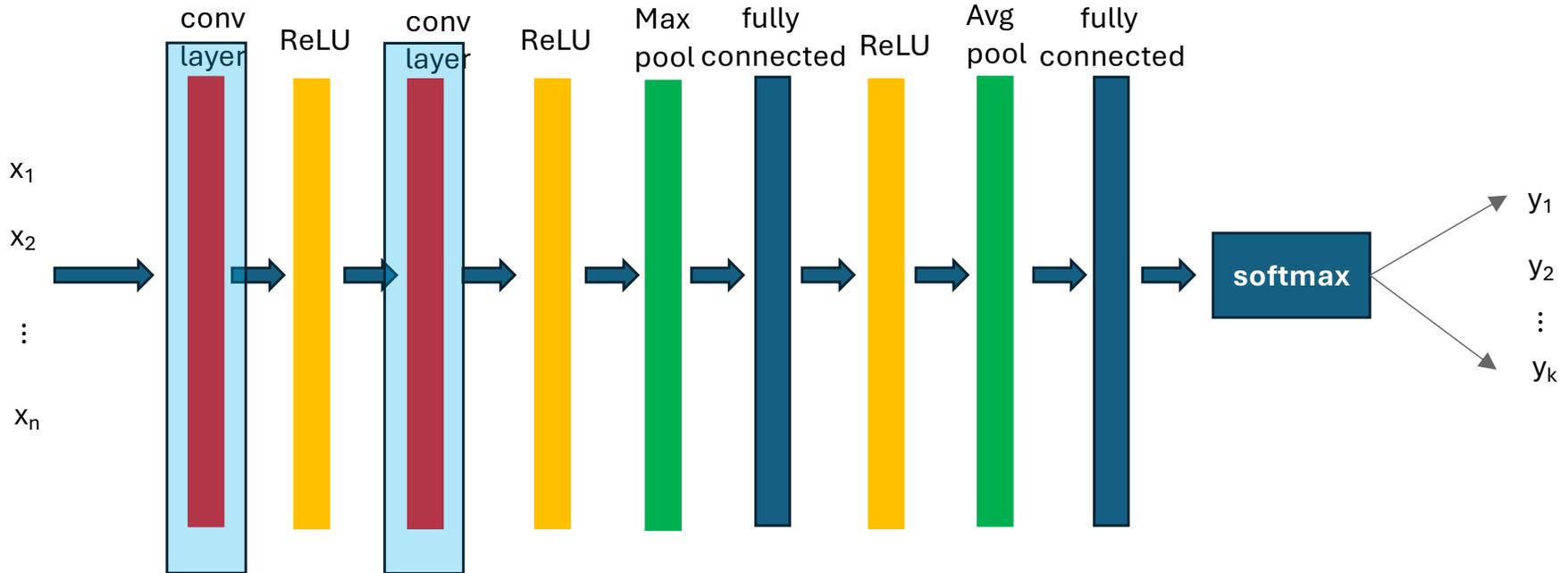
Εσωτερικό Γινόμενο

- Θα δούμε πώς συνδέονται αυτές οι έννοιες με τα συνελκτικά όταν το W και το X : (convolutional) νευρωνικά δίκτυα
- «Μοιάζουν» και στη συνέχεια, με τις
 - Είναι κοντά ενσωματώσεις (embeddings)
 - Σχηματίζουν μικρή γωνία $\angle WX$

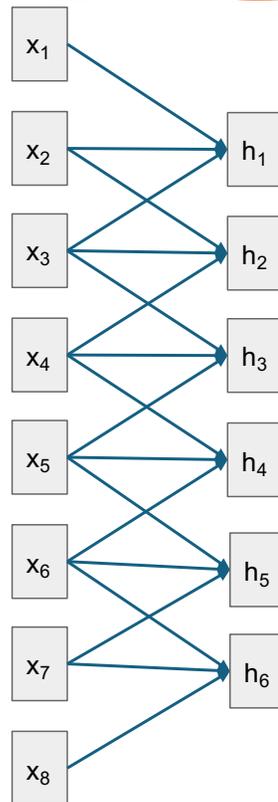
Convolutional Νευρωνικό Δίκτυο



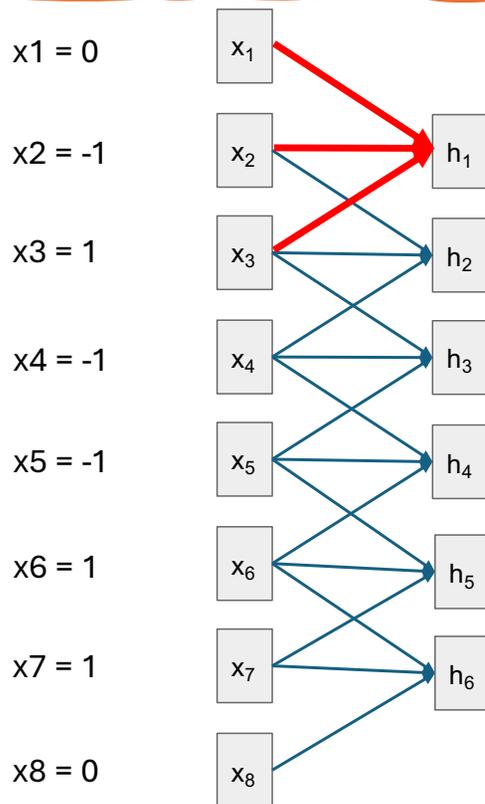
ConvNet: τι κάνουν τα συνελκτικά επίπεδα;



Convolutional layer: τι κάνει;

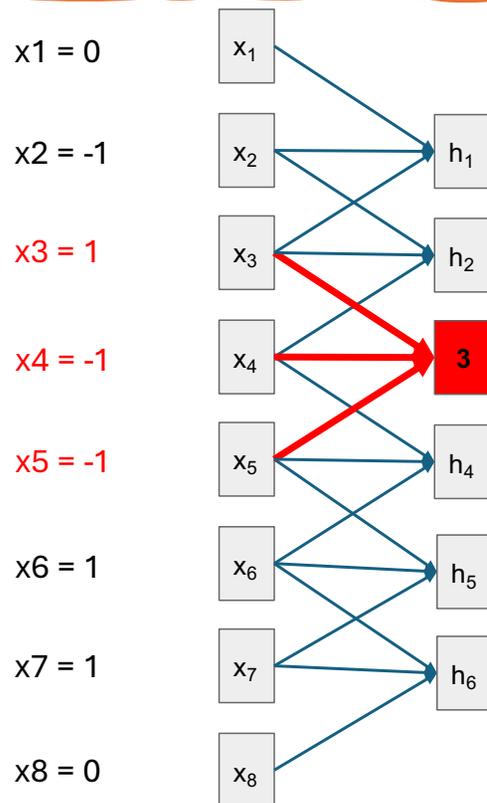


Convolutional layer: τι κάνει;



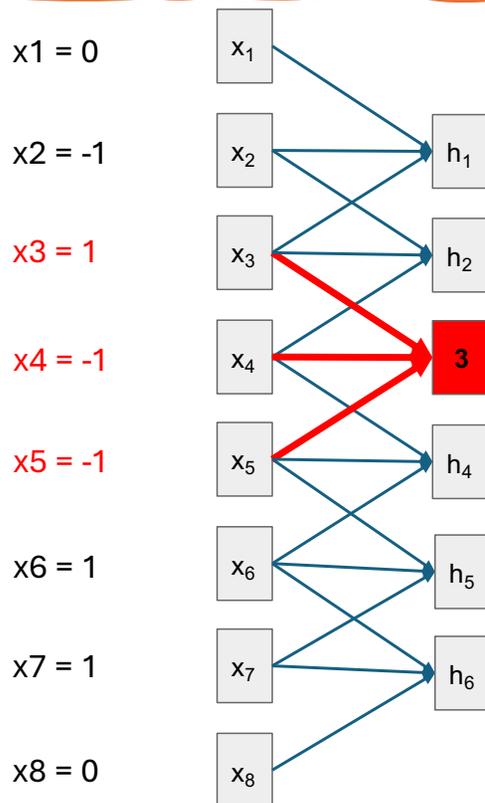
Ποιος νευρώνας στο πρώτο επίπεδο θα έχει την υψηλότερη τιμή (activation);

Convolutional layer: τι κάνει;



Ποιος νευρώνας στο πρώτο επίπεδο θα έχει την υψηλότερη τιμή (activation);

Convolutional layer: τι κάνει;

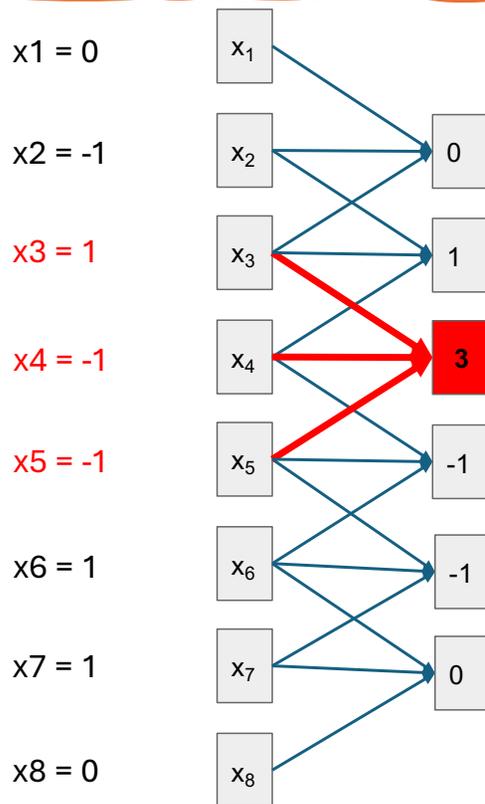


$$w_1 = 1$$
$$w_2 = -1$$
$$w_3 = -1$$

Ποιος νευρώνας στο πρώτο επίπεδο θα έχει την υψηλότερη τιμή (activation);

Οι αξίες (x_1, x_2, x_3) είναι αυτές που «μοιάζουν» πιο πολύ με το «kernel» (w_1, w_2, w_3)

Convolutional layer: τι κάνει;

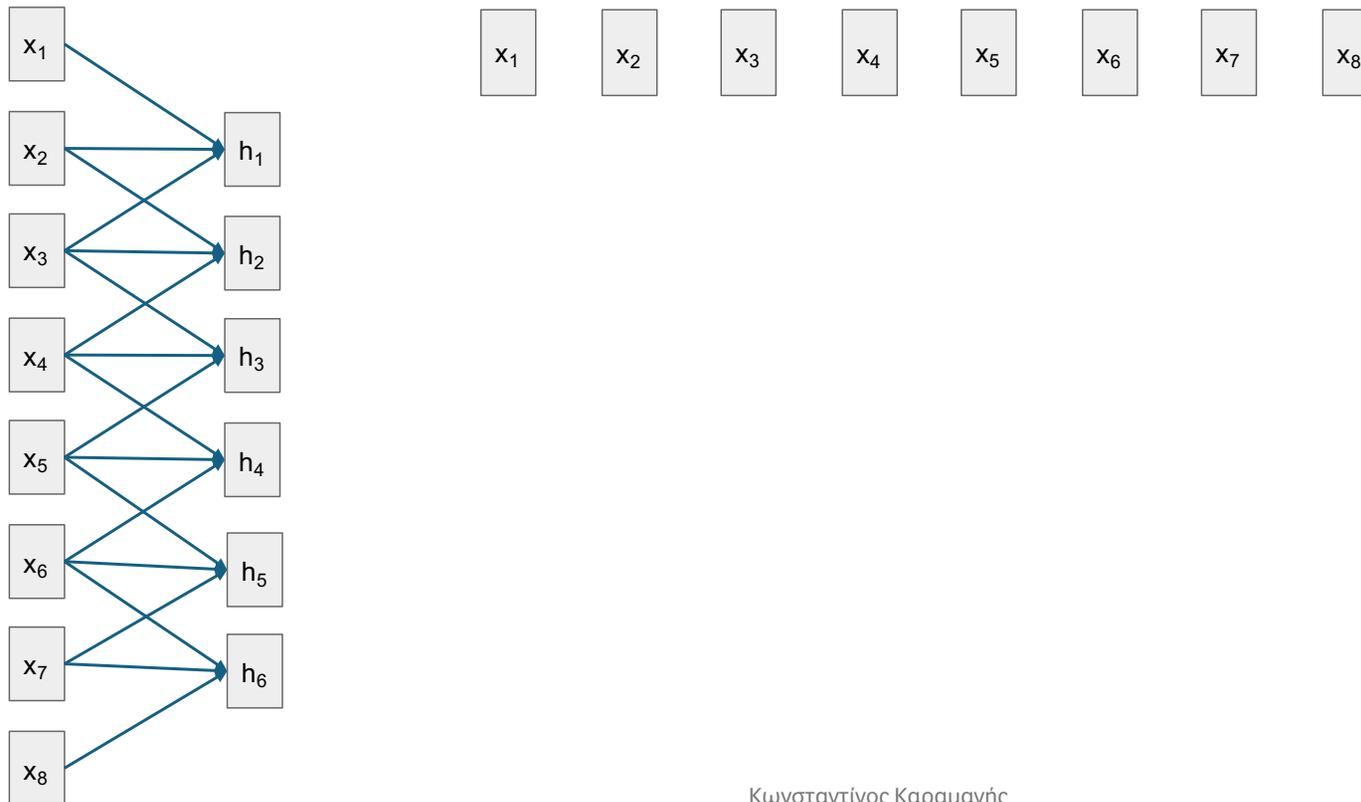


$$w_1 = 1$$
$$w_2 = -1$$
$$w_3 = -1$$

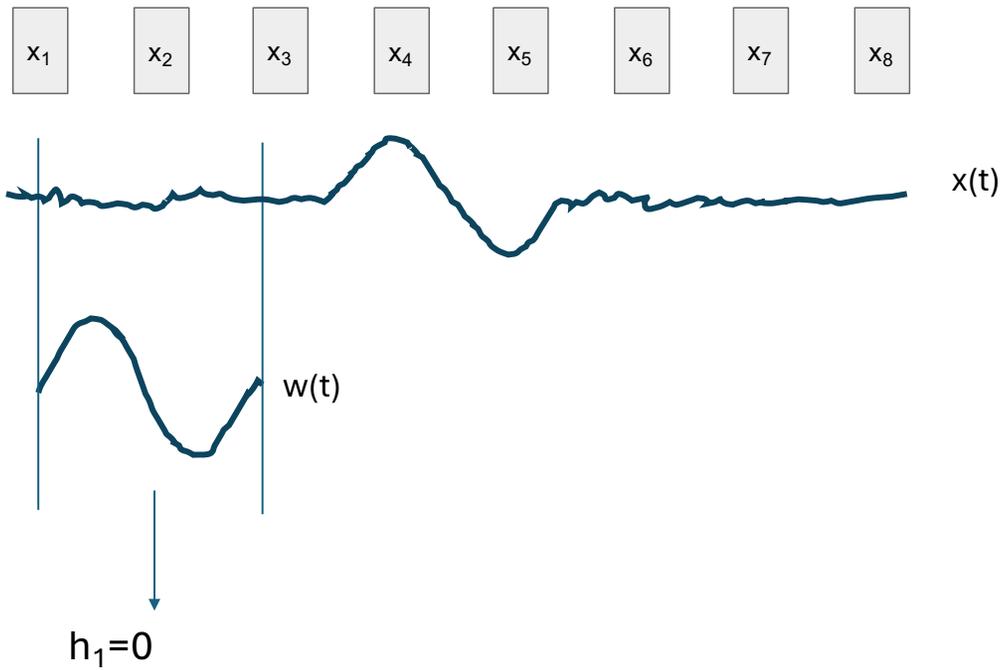
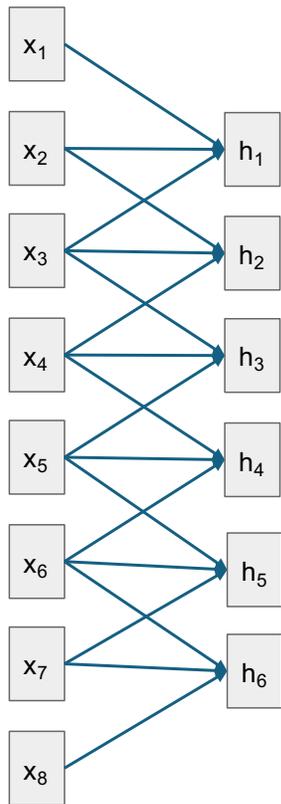
Ποιος νευρώνας στο πρώτο επίπεδο θα έχει την υψηλότερη τιμή (activation);

Οι τιμές (x_1, x_2, x_3) είναι αυτές που «μοιάζουν» πιο πολύ με το «kernel» (w_1, w_2, w_3)

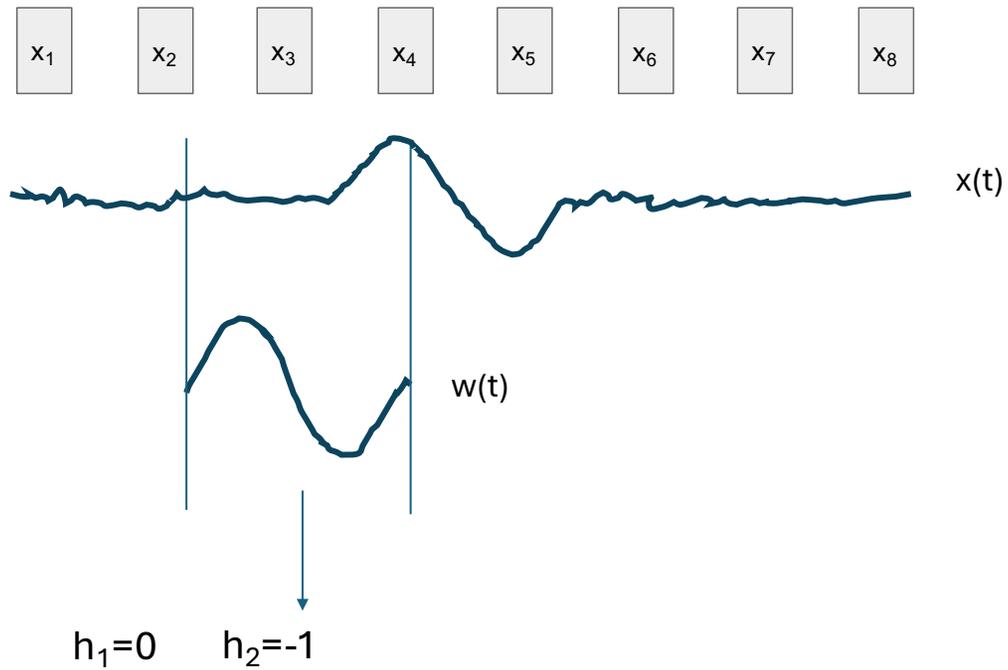
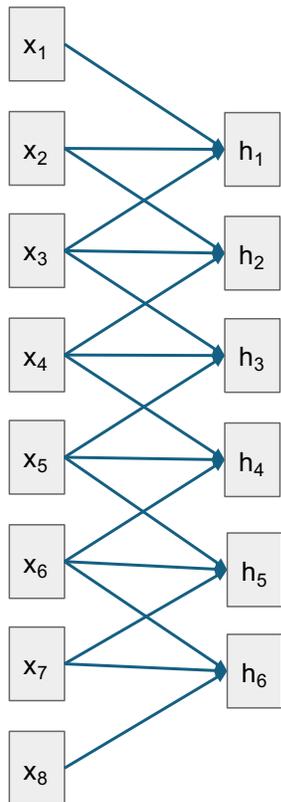
1D Convolution: 2^ο παράδειγμα



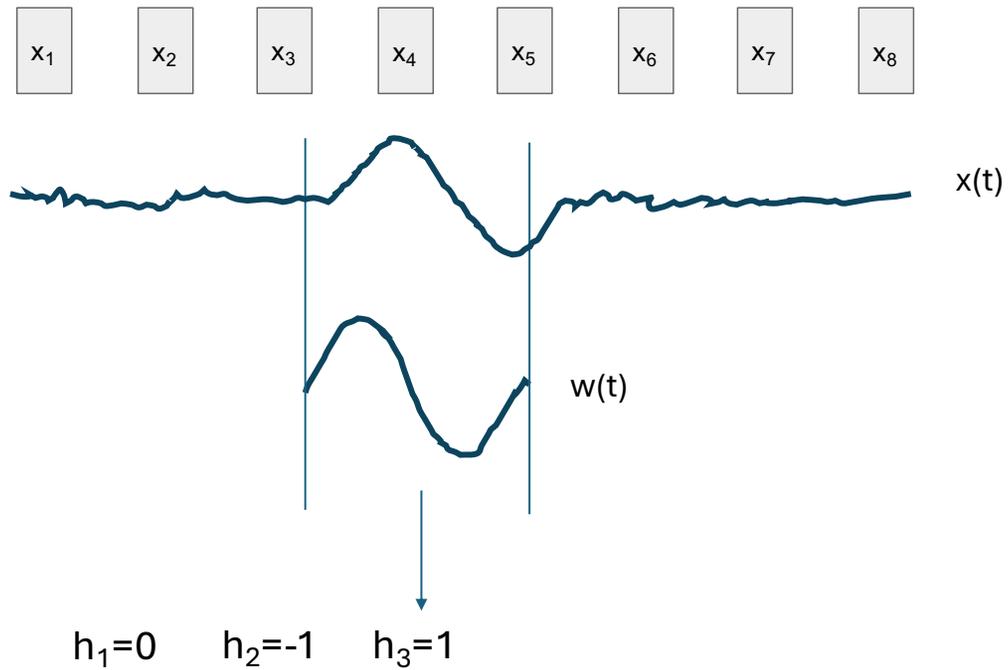
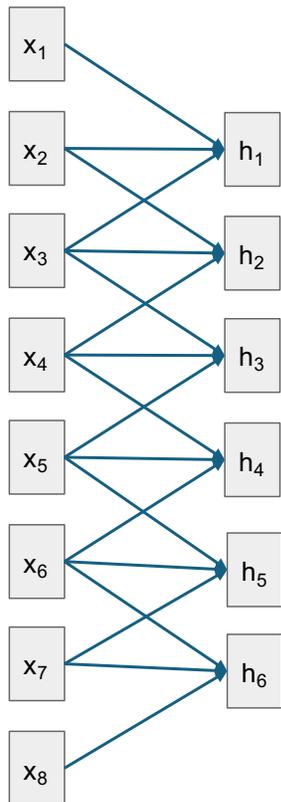
1D Convolution: 2^ο παράδειγμα



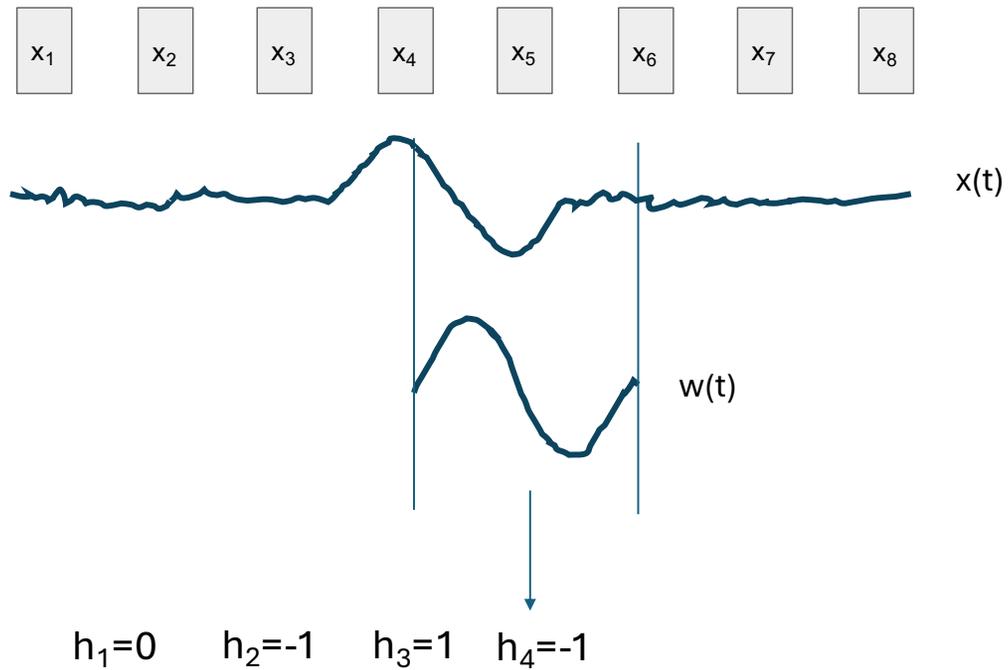
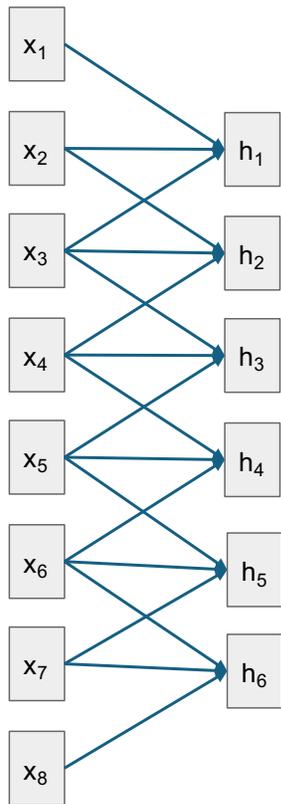
1D Convolution: 2^ο παράδειγμα



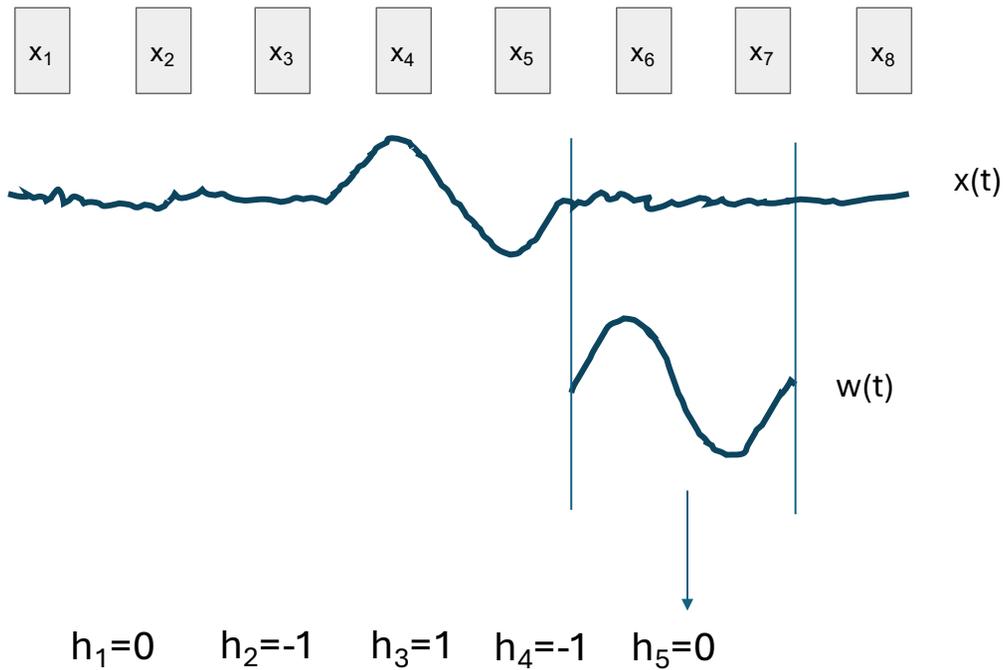
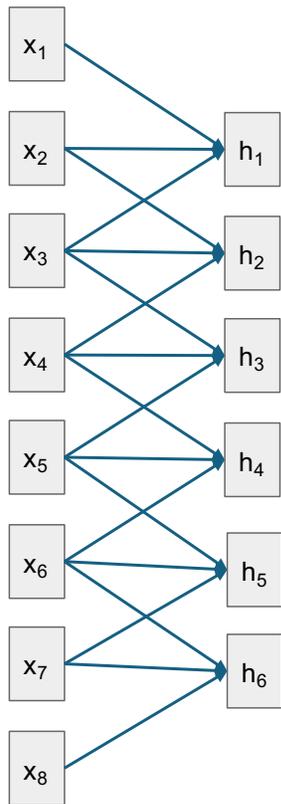
1D Convolution: 2^ο παράδειγμα



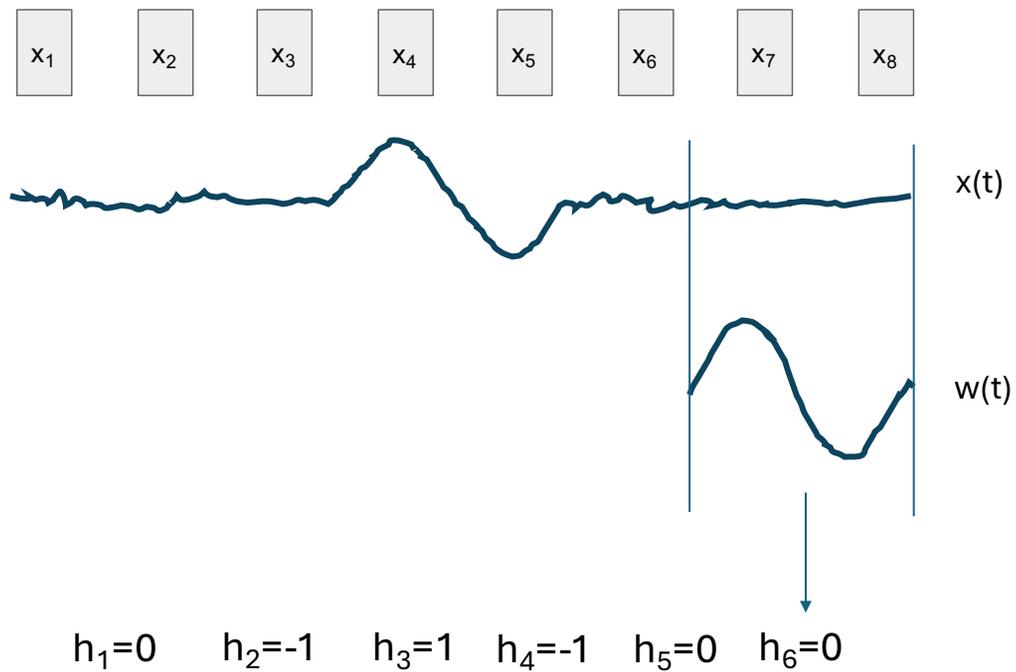
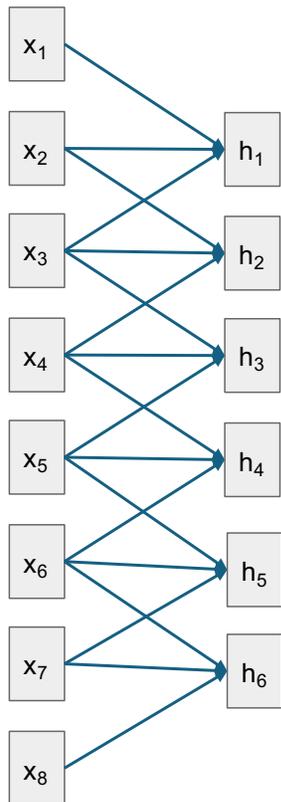
1D Convolution: 2^ο παράδειγμα



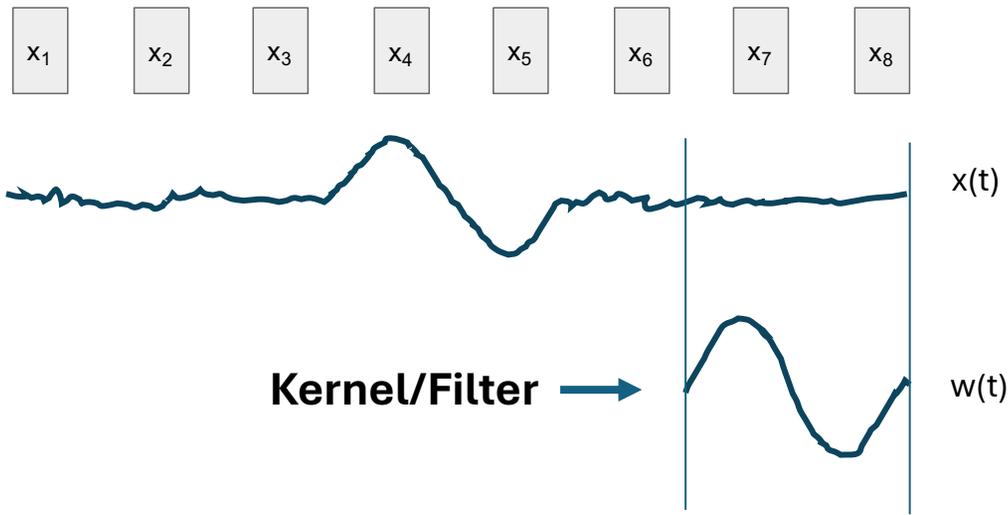
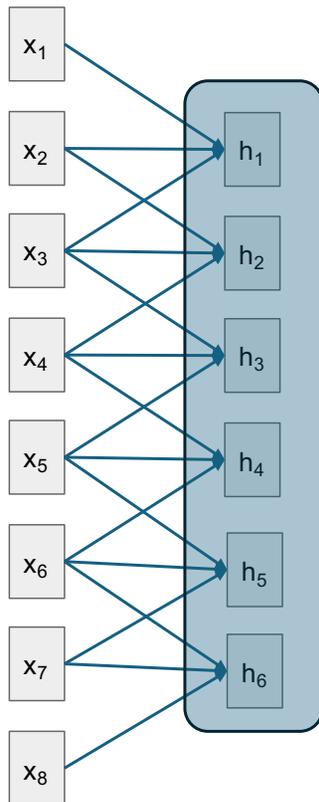
1D Convolution: 2^ο παράδειγμα



1D Convolution: 2^ο παράδειγμα



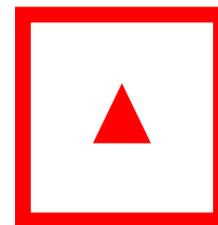
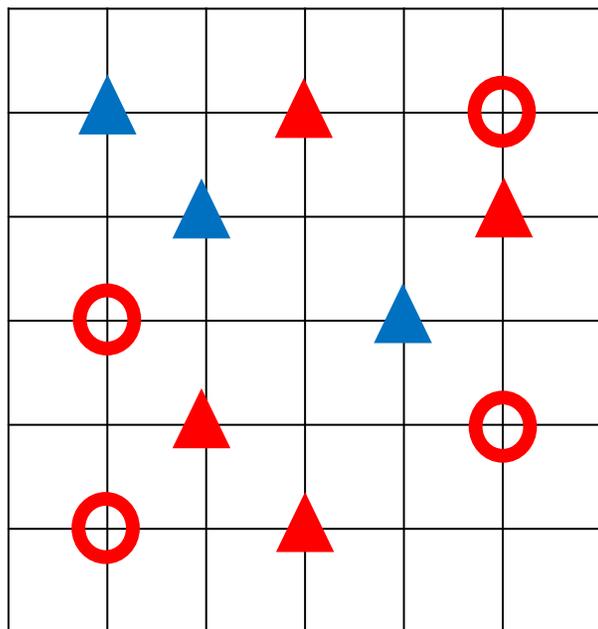
1D Convolution: 2^ο παράδειγμα



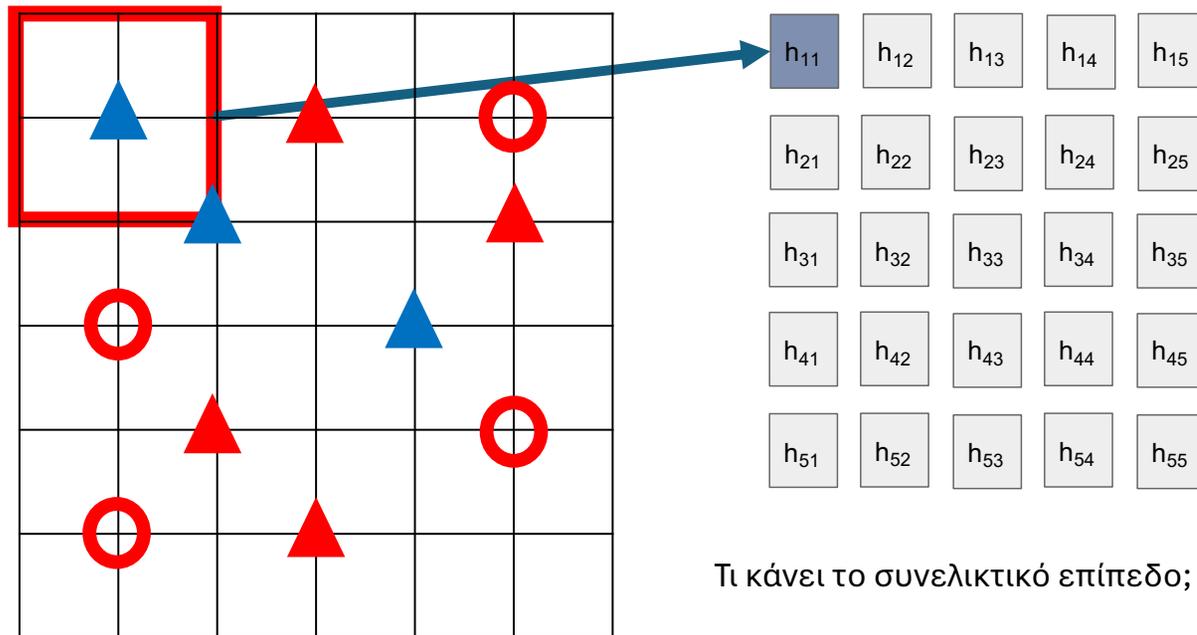
$h_1=0$ $h_2=-1$ $h_3=1$ $h_4=-1$ $h_5=0$ $h_6=0$

← Επόμενο επίπεδο

Convolutional επίπεδο για εικόνες



Convolutional επίπεδο για εικόνες

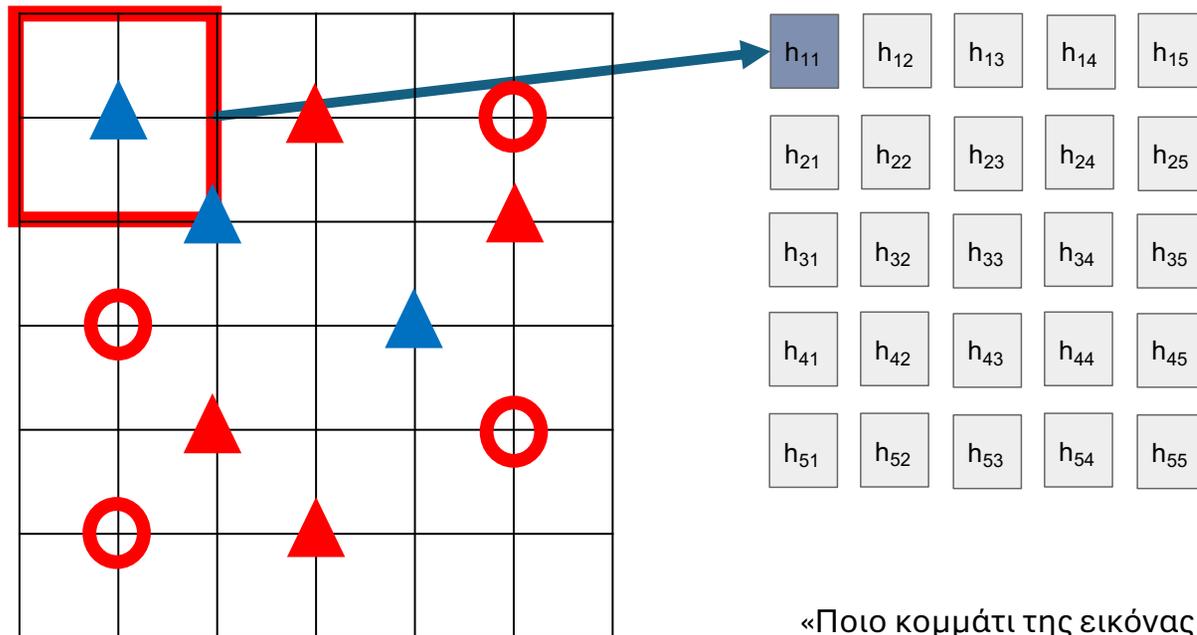


Τι κάνει το συνελκτικό επίπεδο;

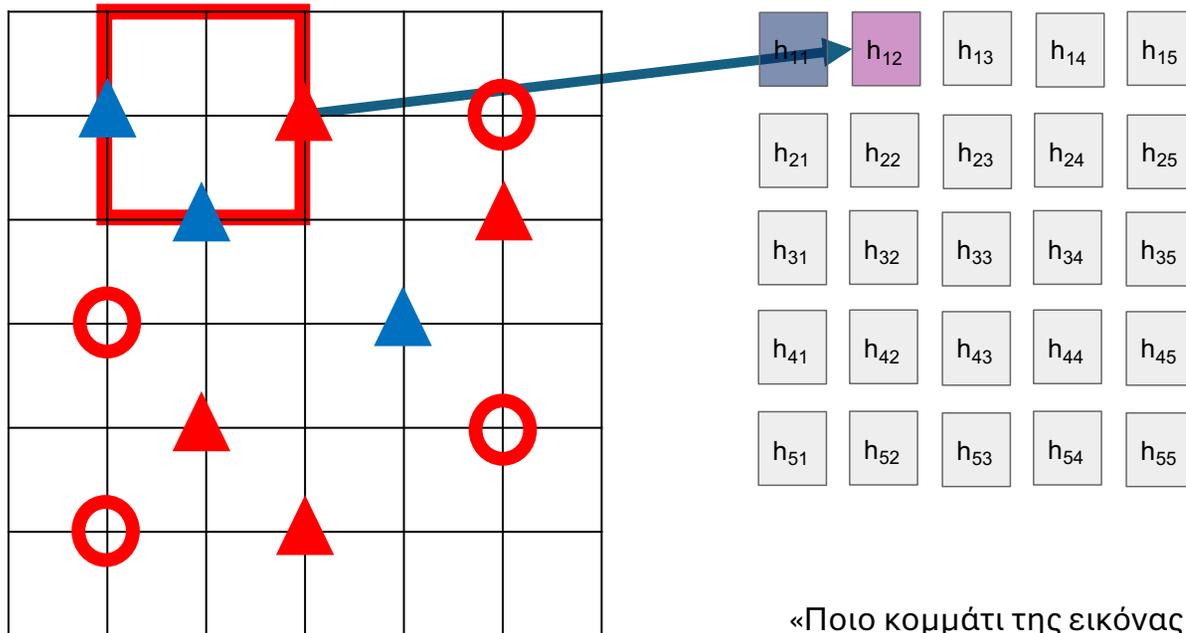
Προσπαθεί να βρει συγκεκριμένα σχήματα σε κάθε κομμάτι της εικόνας.

Κωνσταντίνος Καραμανής

Convolutional επίπεδο για εικόνες



Convolutional επίπεδο για εικόνες

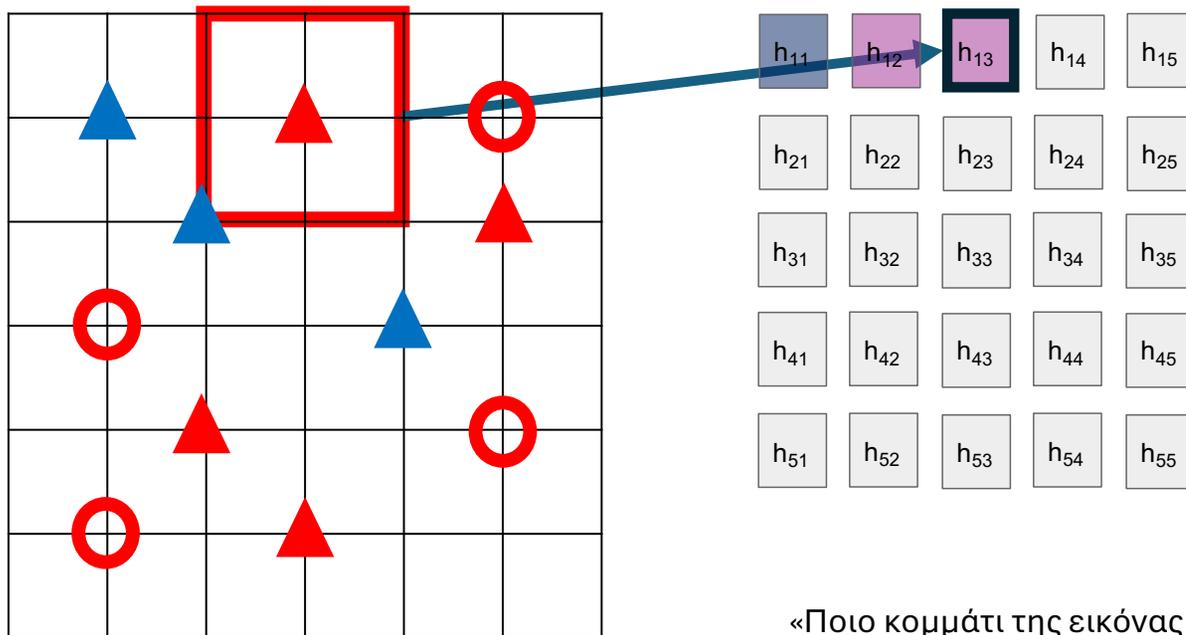


«Ποιο κομμάτι της εικόνας περιέχει ένα

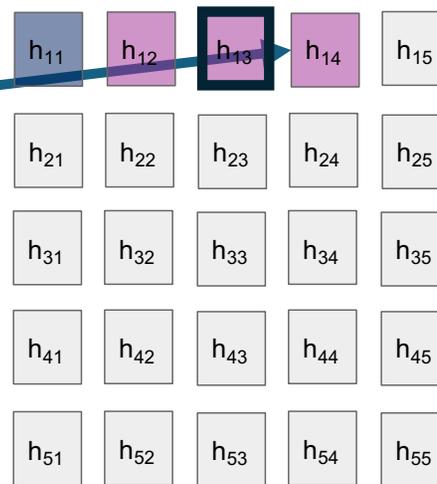
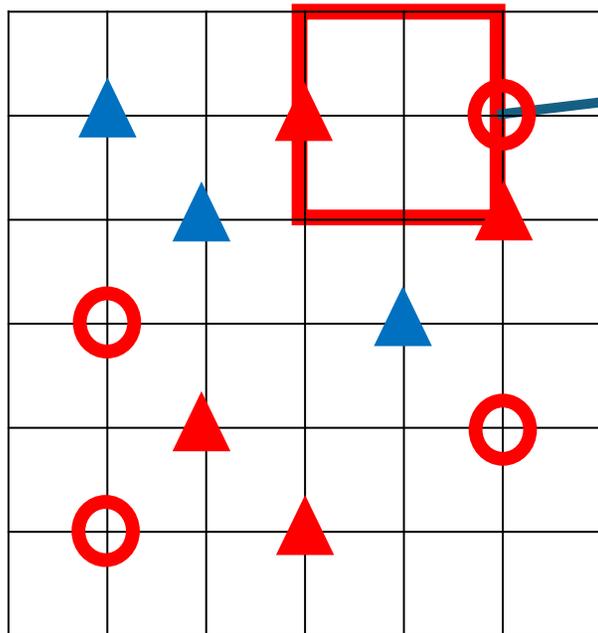


»

Convolutional επίπεδο για εικόνες



Convolutional επίπεδο για εικόνες

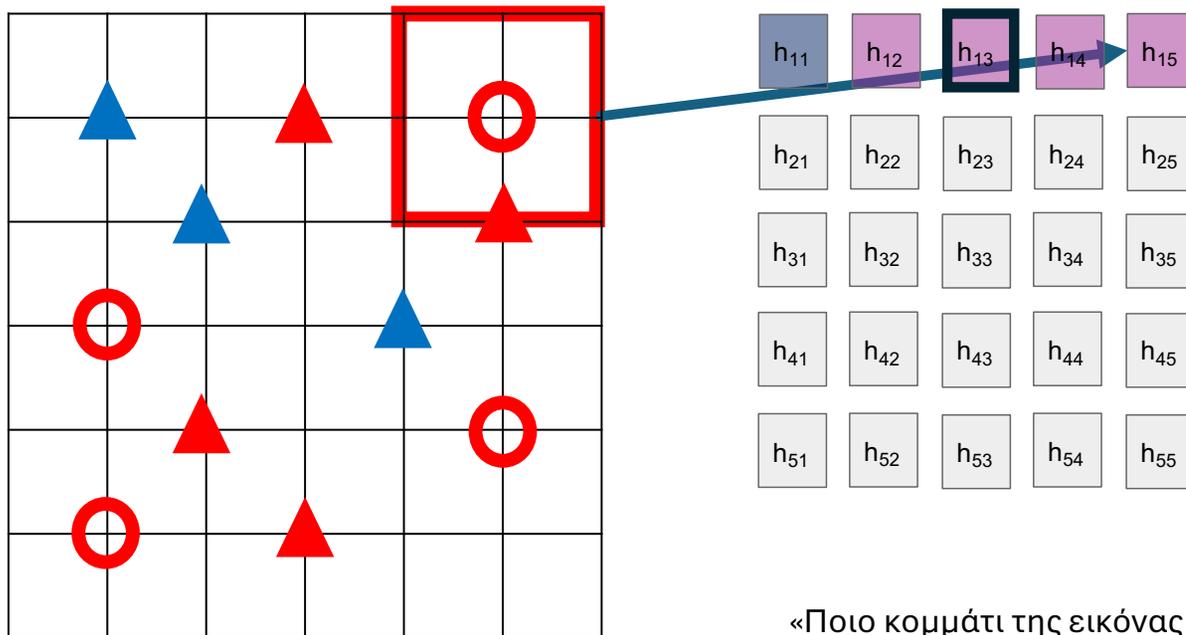


«Ποιο κομμάτι της εικόνας περιέχει ένα

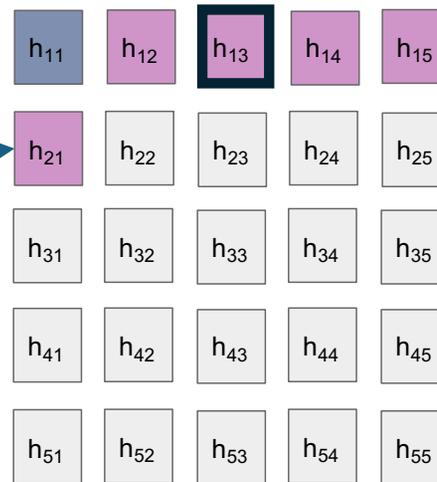
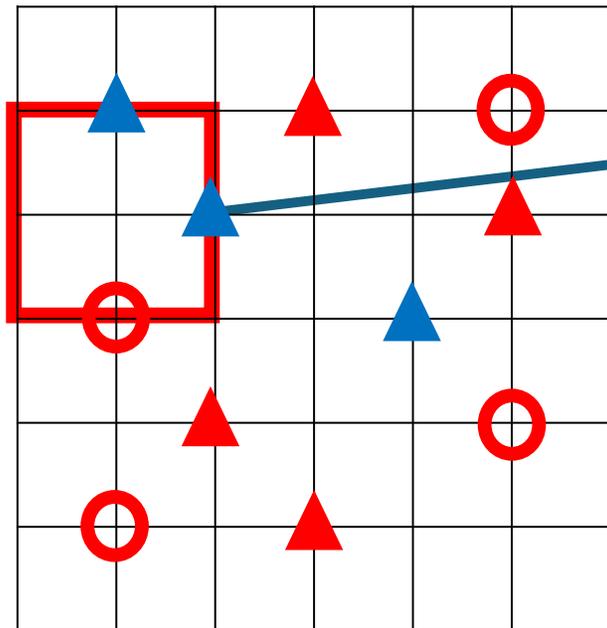


»

Convolutional επίπεδο για εικόνες



Convolutional επίπεδο για εικόνες

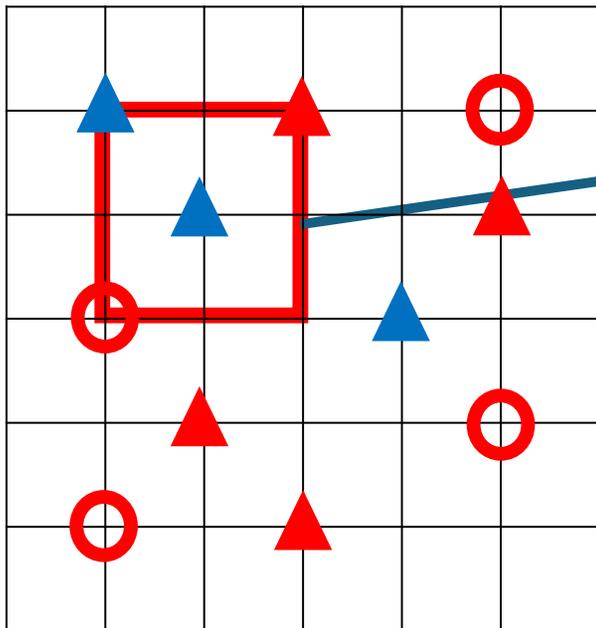


«Ποιο κομμάτι της εικόνας περιέχει ένα



»

Convolutional επίπεδο για εικόνες



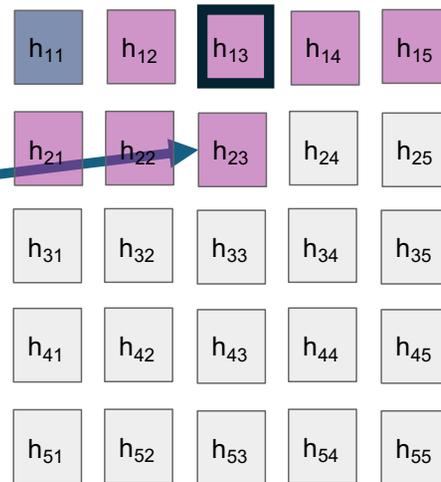
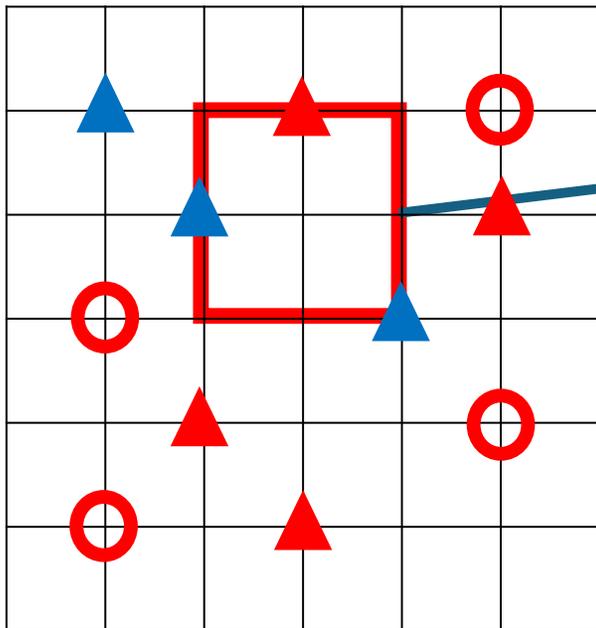
h ₁₁	h ₁₂	h ₁₃	h ₁₄	h ₁₅
h ₂₁	h ₂₂	h ₂₃	h ₂₄	h ₂₅
h ₃₁	h ₃₂	h ₃₃	h ₃₄	h ₃₅
h ₄₁	h ₄₂	h ₄₃	h ₄₄	h ₄₅
h ₅₁	h ₅₂	h ₅₃	h ₅₄	h ₅₅

«Ποιο κομμάτι της εικόνας περιέχει ένα



»

Convolutional επίπεδο για εικόνες

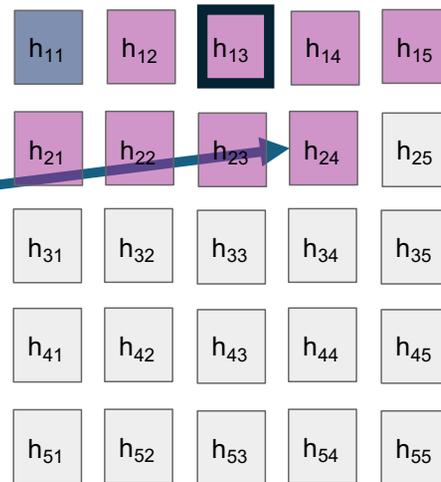
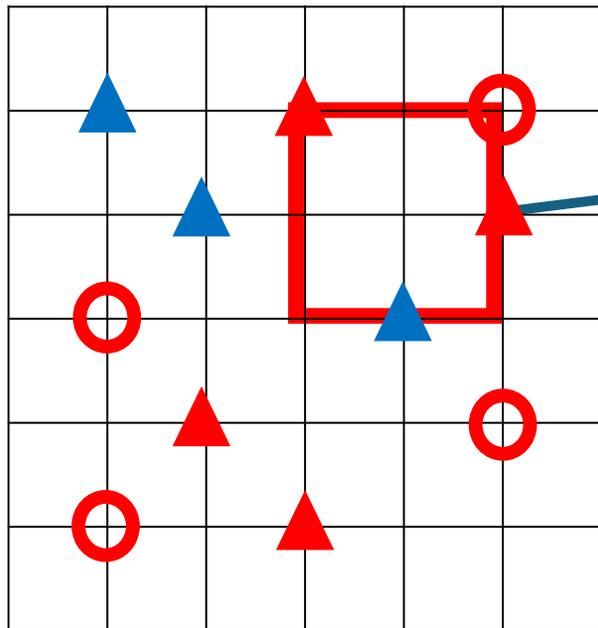


«Ποιο κομμάτι της εικόνας περιέχει ένα



»

Convolutional επίπεδο για εικόνες

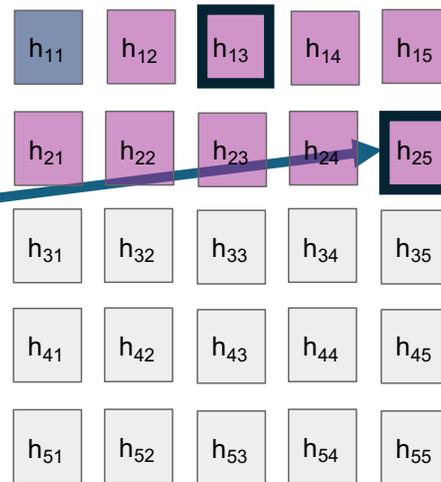
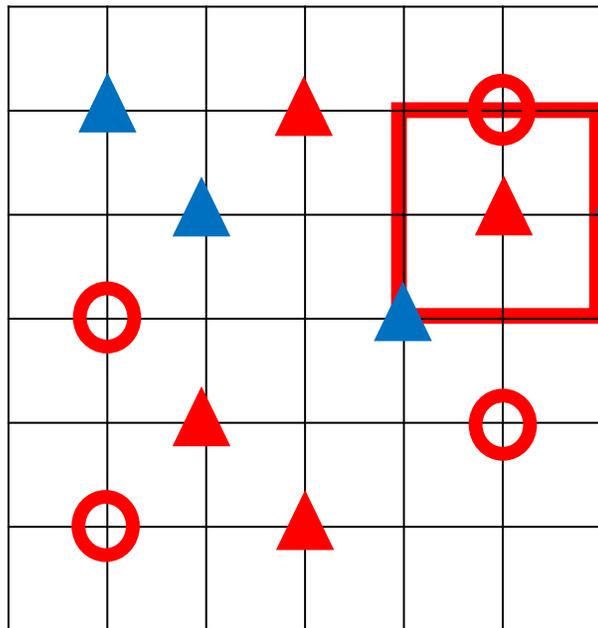


«Ποιο κομμάτι της εικόνας περιέχει ένα

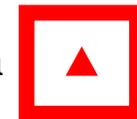


»

Convolutional επίπεδο για εικόνες

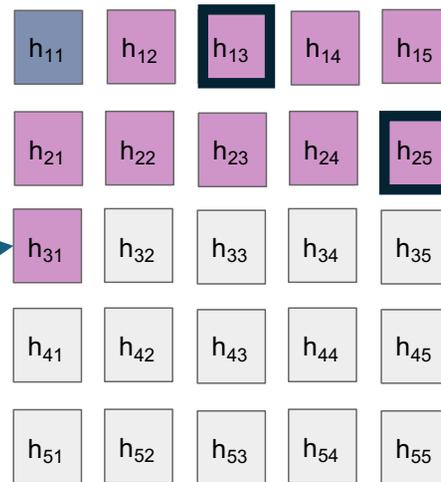
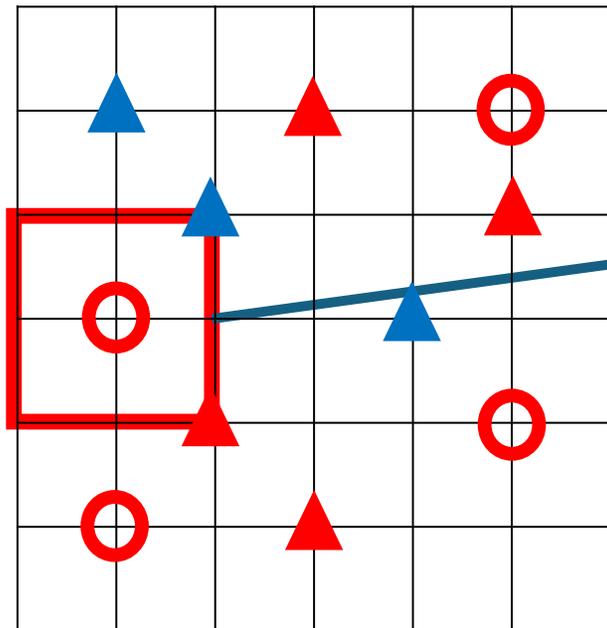


«Ποιο κομμάτι της εικόνας περιέχει ένα

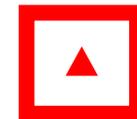


»

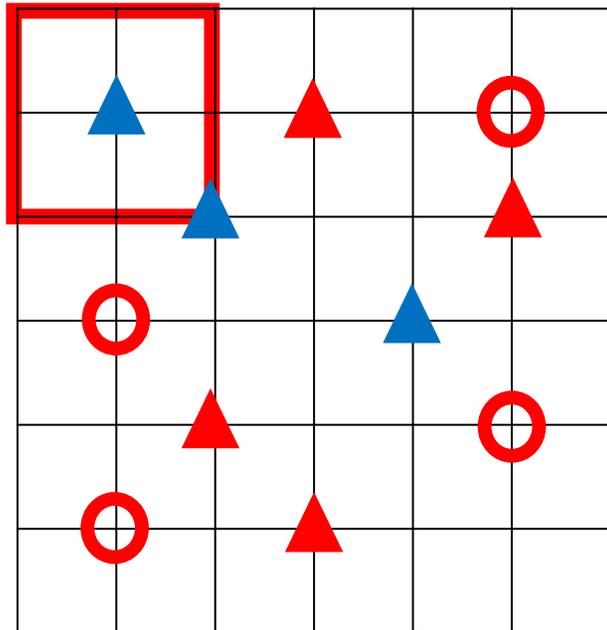
Convolutional επίπεδο για εικόνες



«Πιο κομμάτι της εικόνας περιέχει ένα



Convolutional επίπεδο για εικόνες



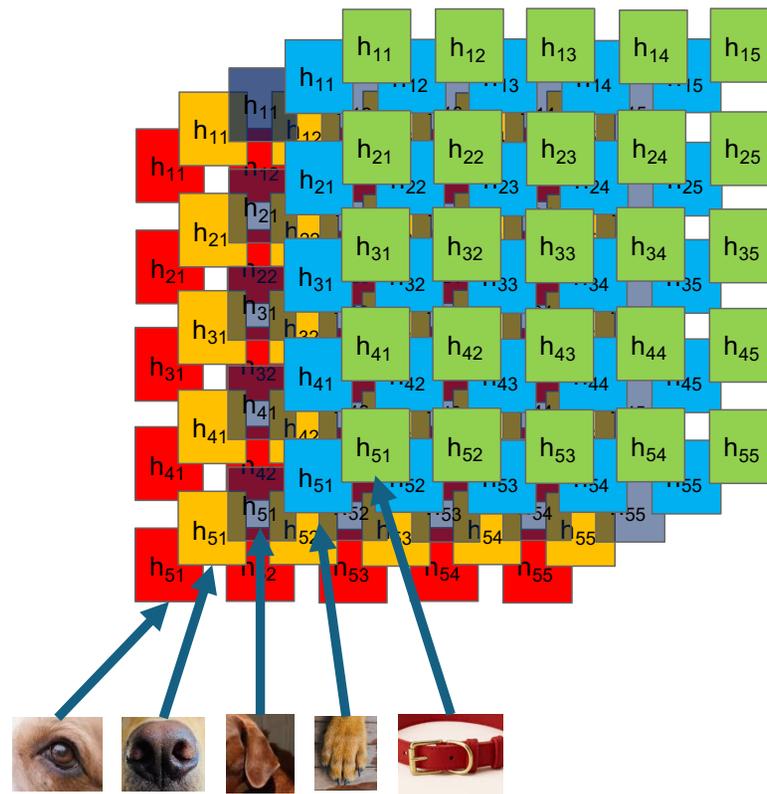
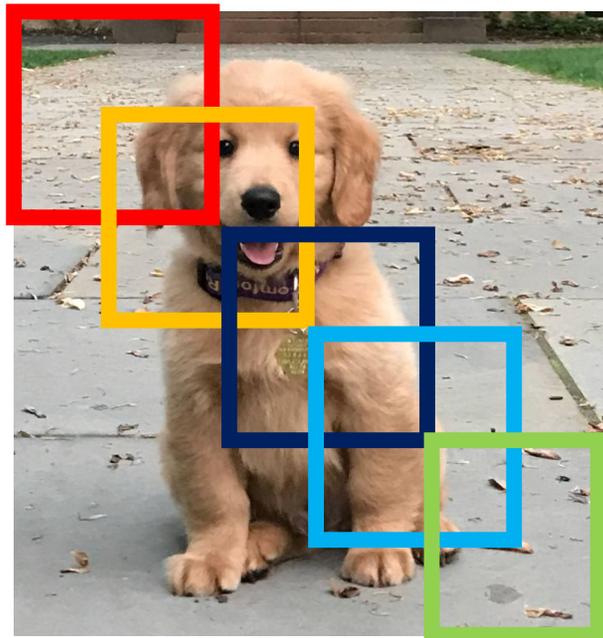
h_{11}	h_{12}	h_{13}	h_{14}	h_{15}
h_{21}	h_{22}	h_{23}	h_{24}	h_{25}
h_{31}	h_{32}	h_{33}	h_{34}	h_{35}
h_{41}	h_{42}	h_{43}	h_{44}	h_{45}
h_{51}	h_{52}	h_{53}	h_{54}	h_{55}

«Ποιο κομμάτι της εικόνας περιέχει ένα

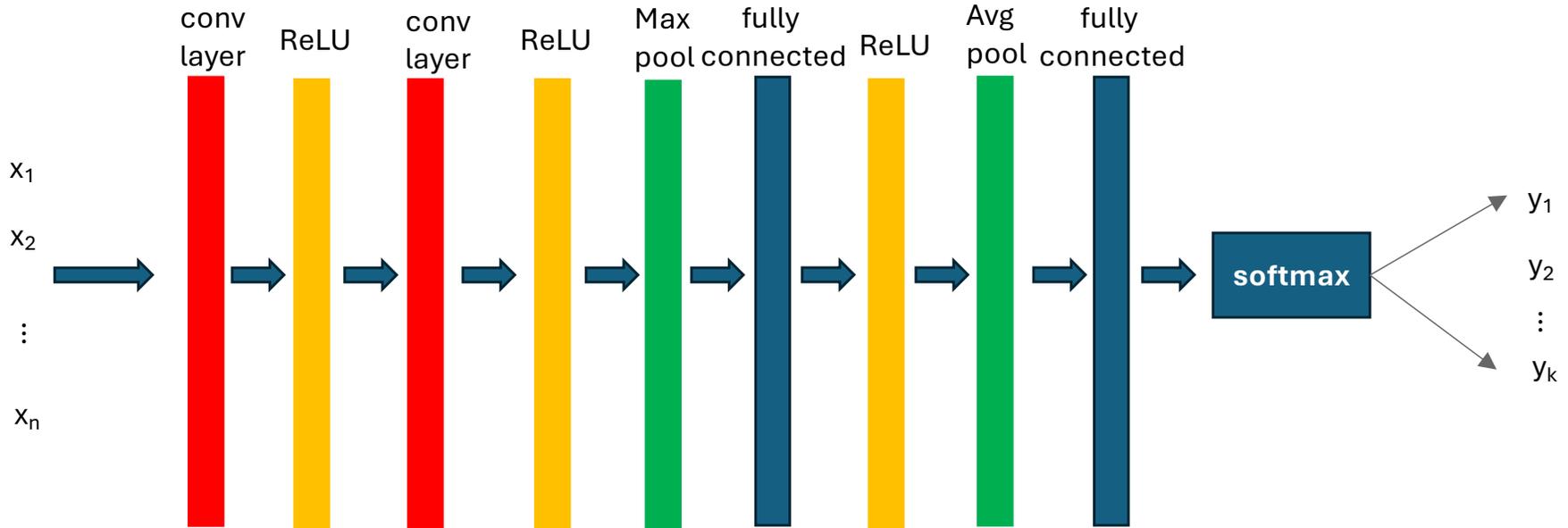


»

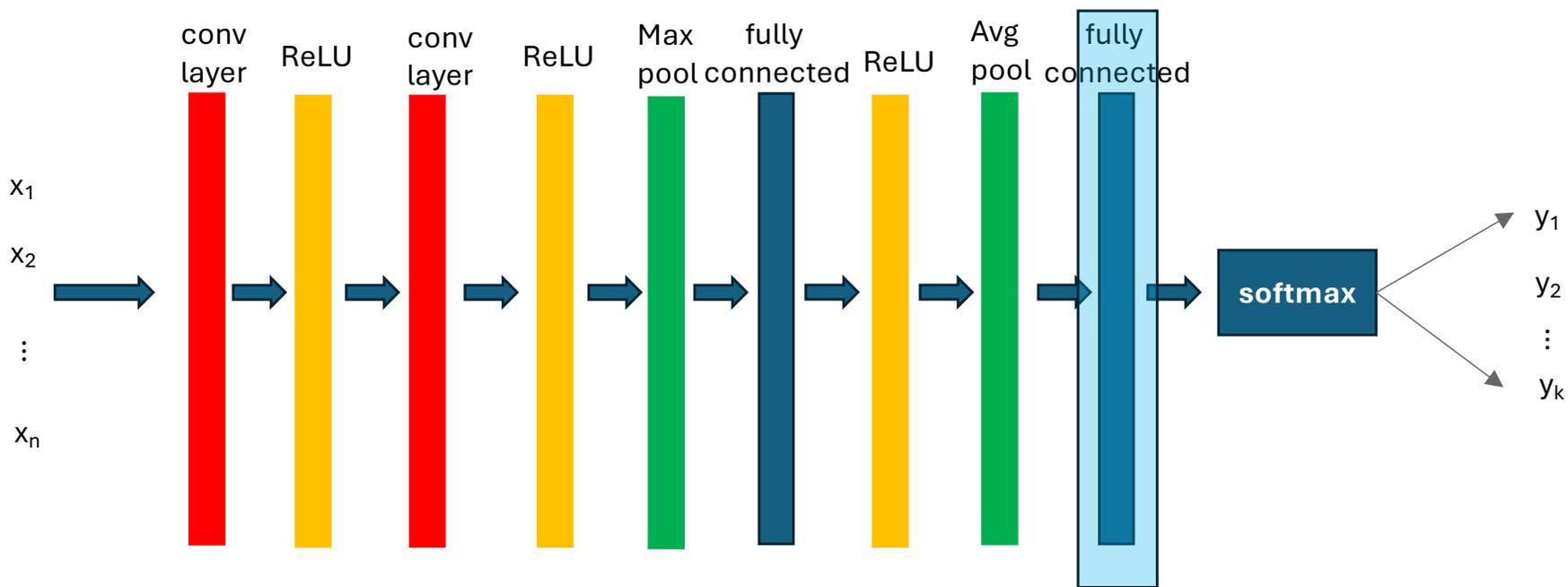
Convolutional επίπεδο για εικόνες



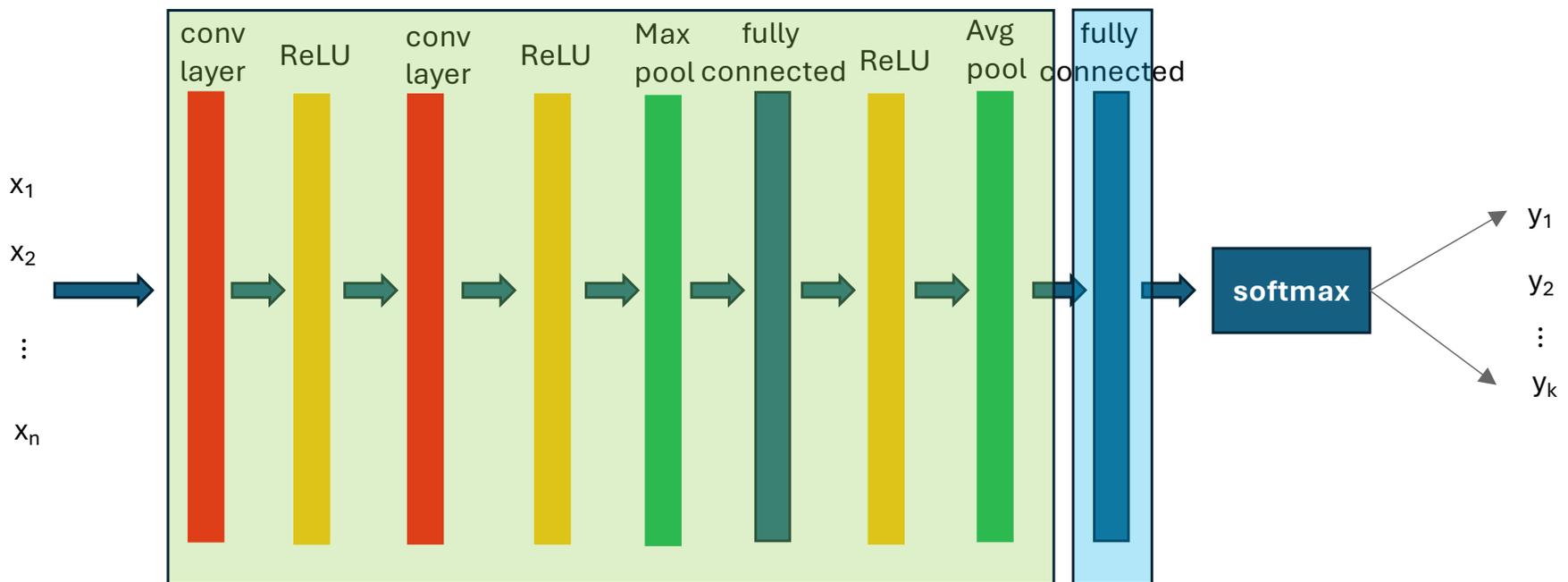
Convolutional Νευρωνικό Δίκτυο



Τι κάνει το τελευταίο επίπεδο;

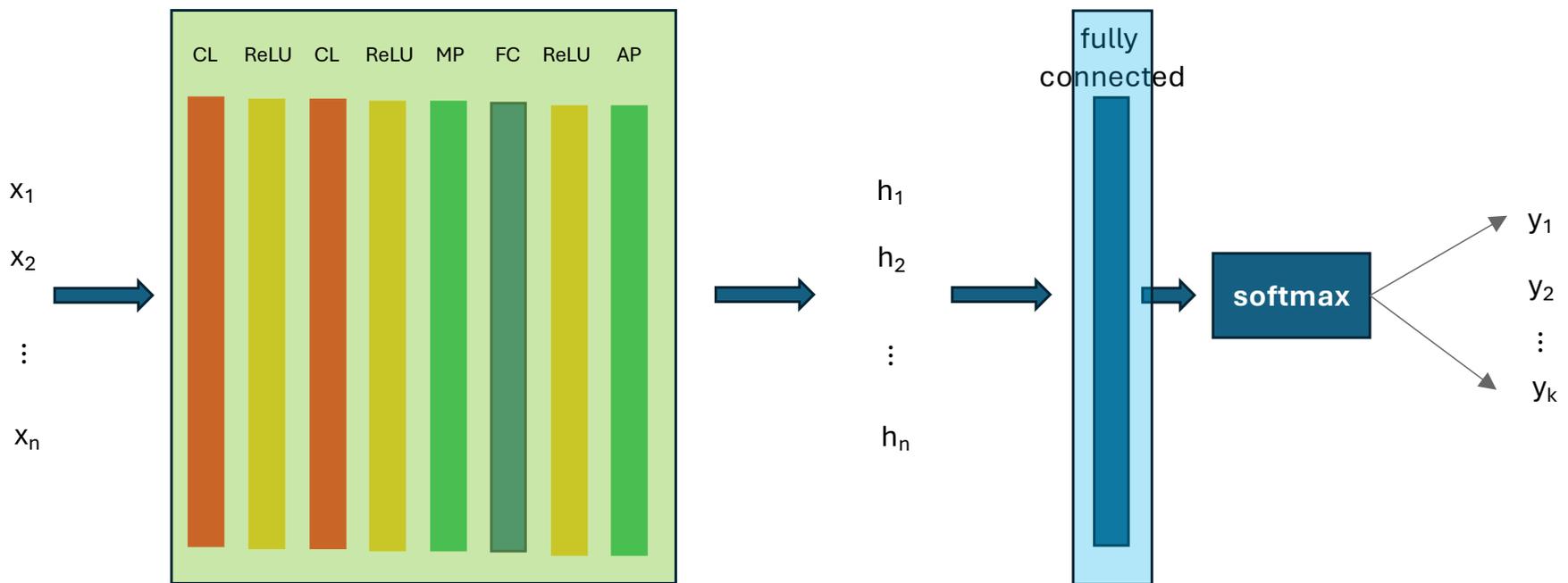


Τι κάνει το τελευταίο επίπεδο;

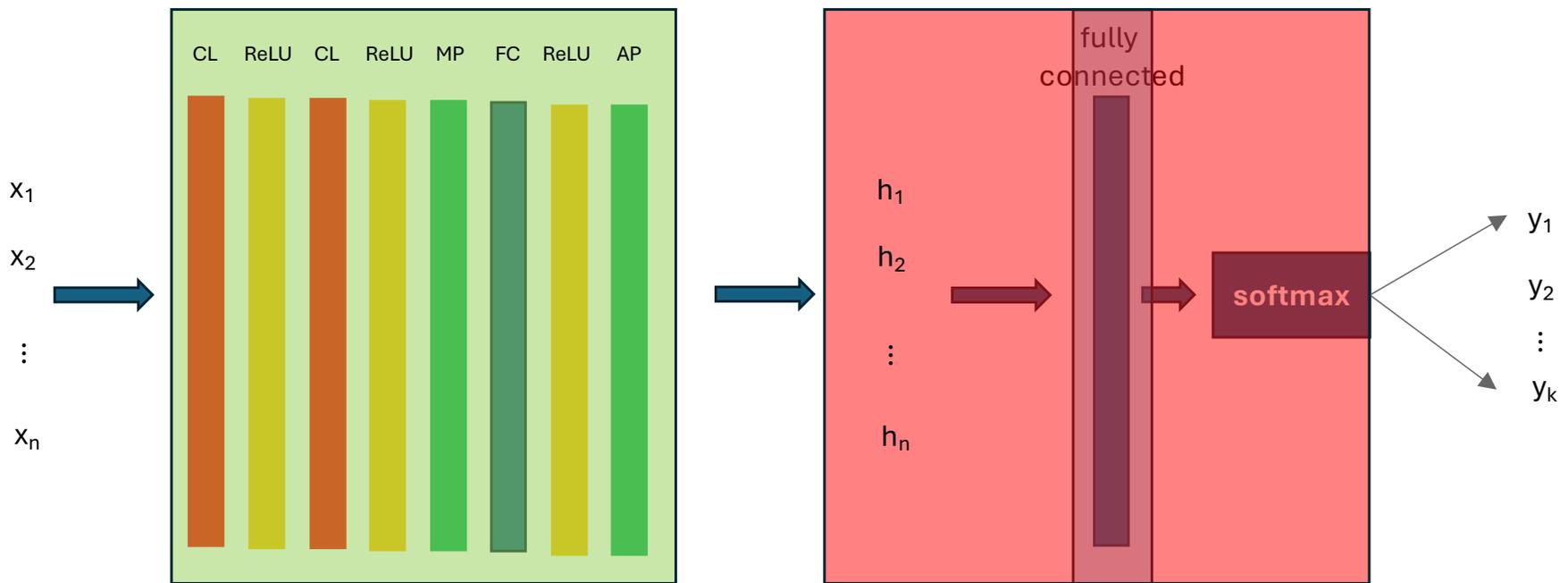


Κωνσταντίνος Καραμανής

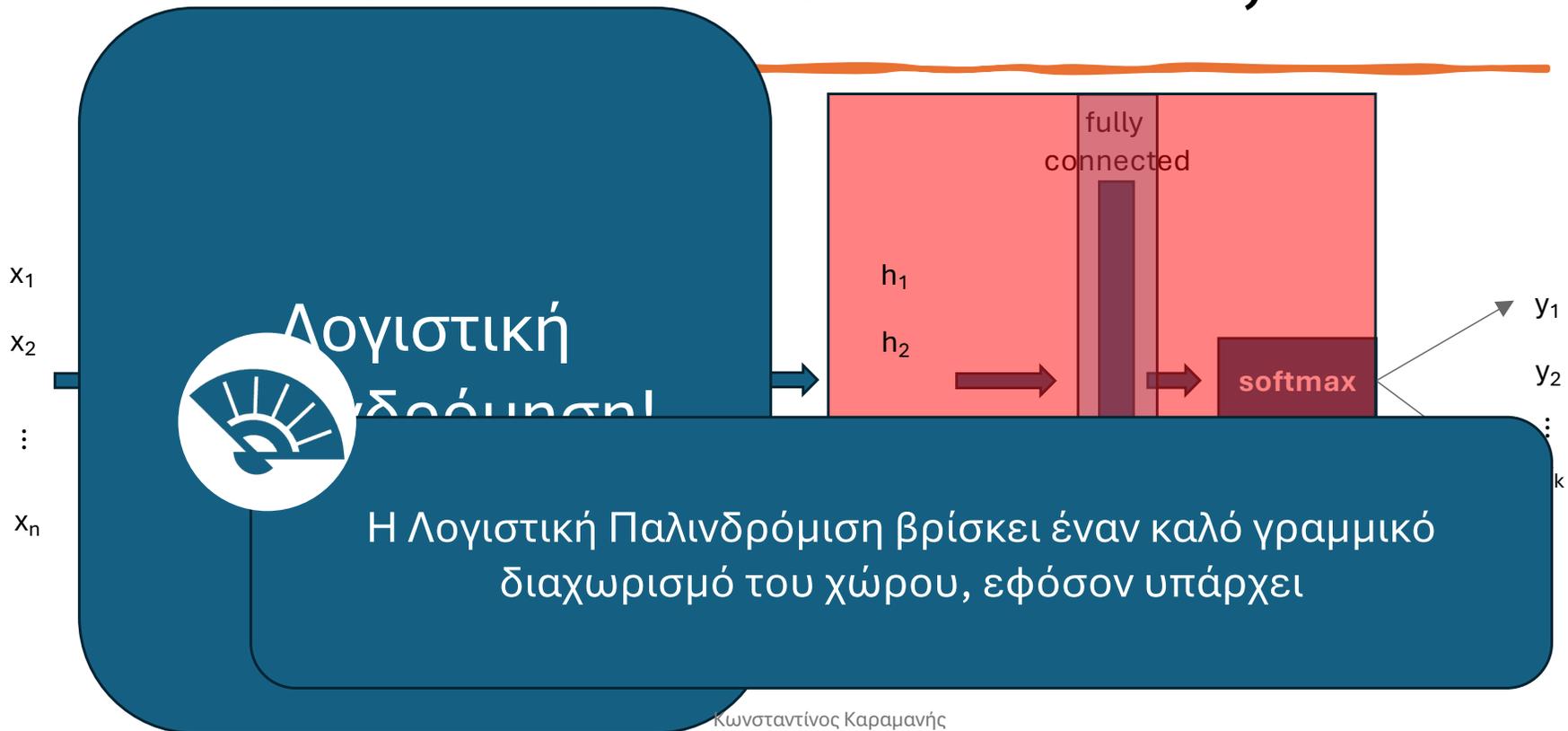
Τι κάνει το τελευταίο επίπεδο;



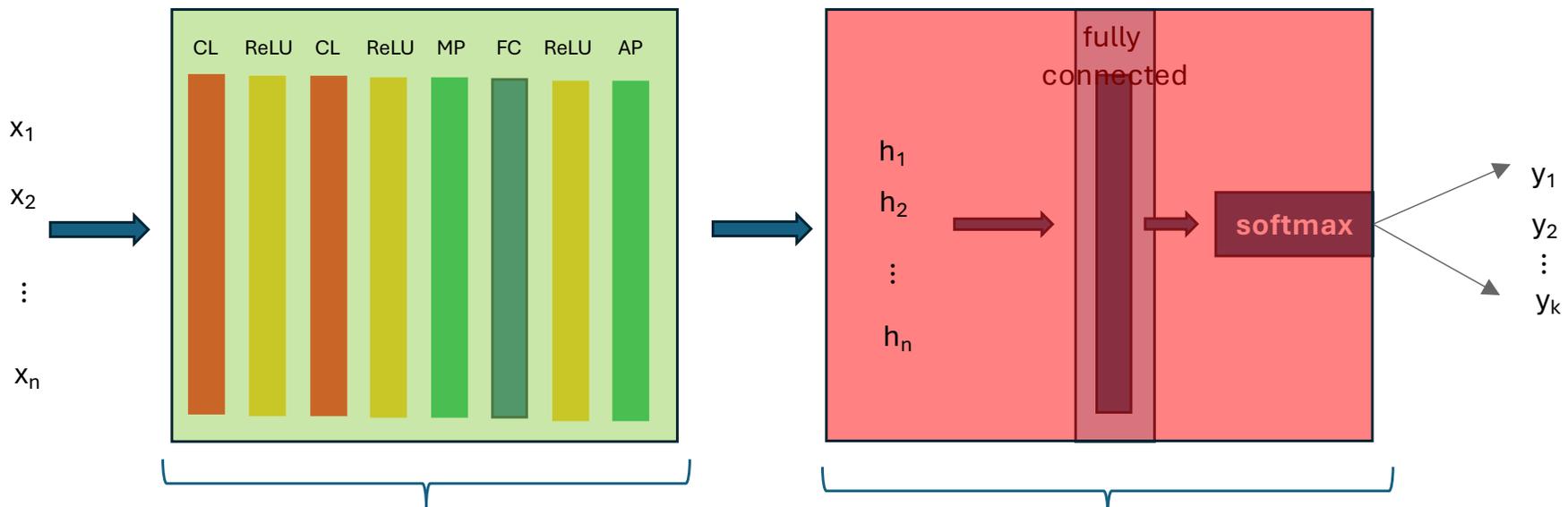
Τι κάνει το τελευταίο επίπεδο;



Τι κάνει το τελευταίο επίπεδο;



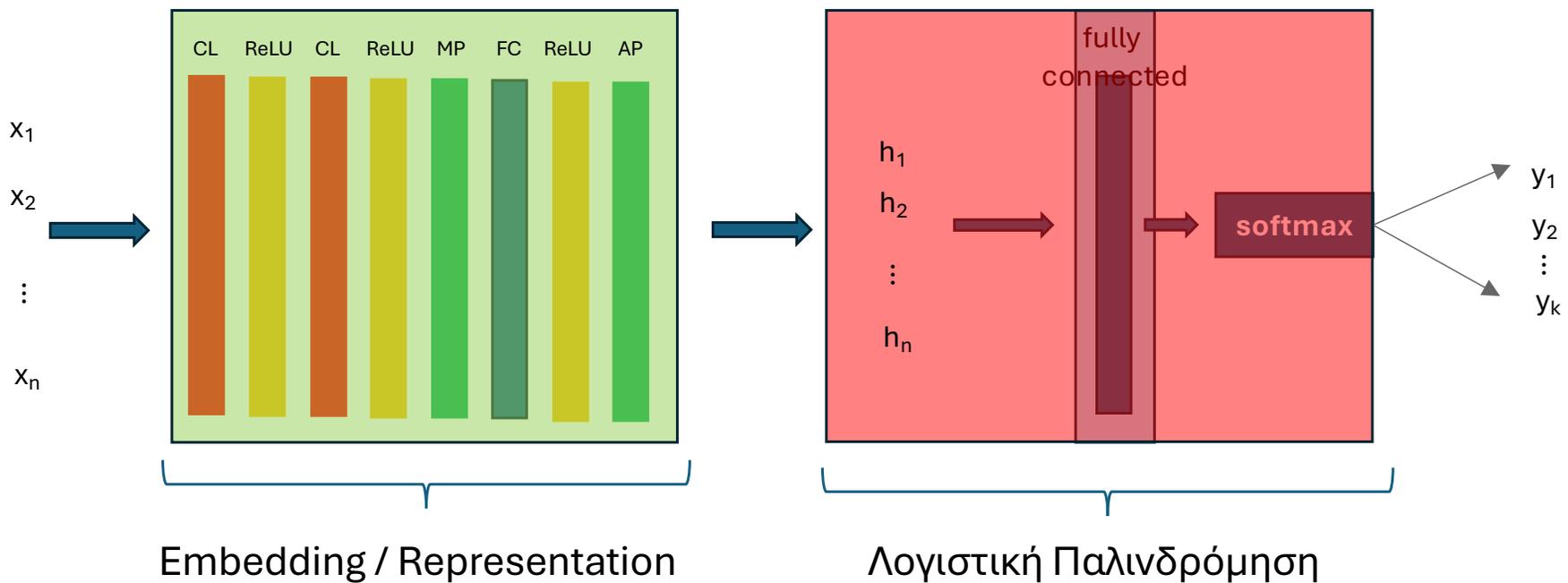
Τι κάνει το τελευταίο επίπεδο;



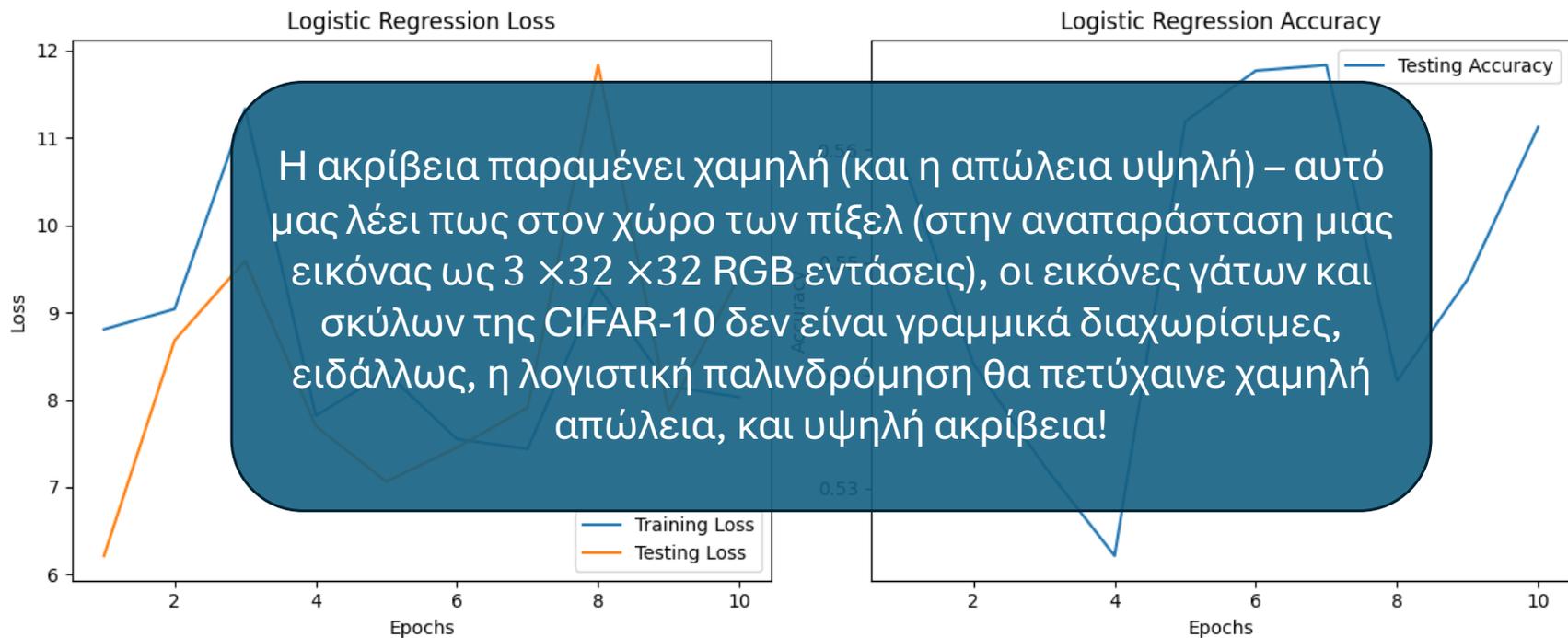
Βρίσκει καλό μια καλή ενσωμάτωση (embedding) ή αναπαράσταση (representation) για τα δεδομένα, όπου είναι (σχεδόν) γραμμικά διαχωρίσιμα (linearly separable)

Λύνει ένα πρόβλημα λογιστικής παλινδρόμησης στον χώρο των επεξεργασμένων δεδομένων (representation / embedding)

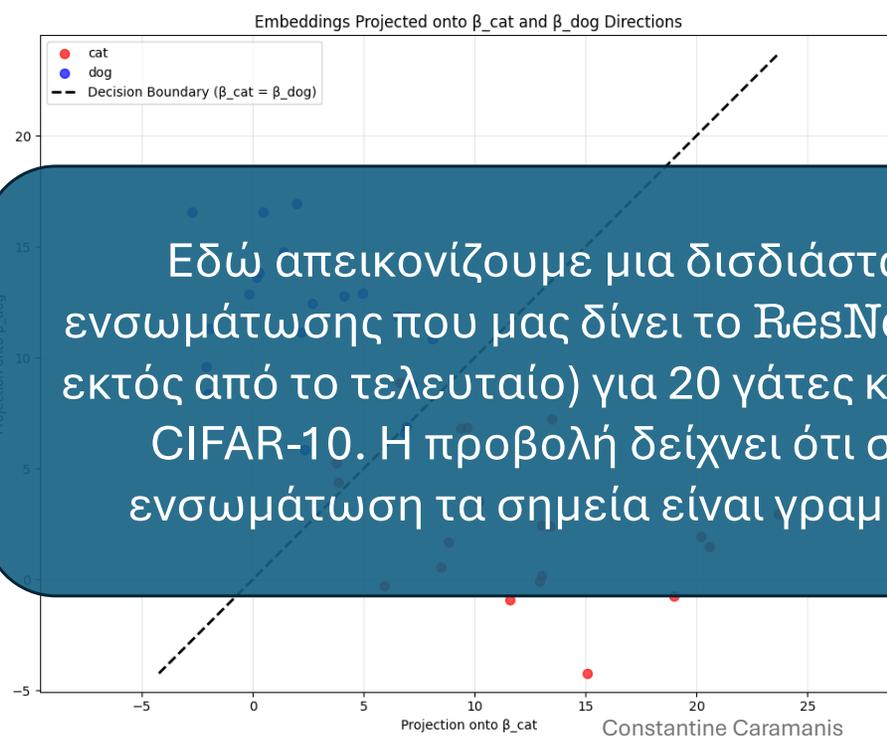
Τι κάνει το τελευταίο επίπεδο;



Λογιστική Παλινδρόμηση: Γάτες και Σκύλοι

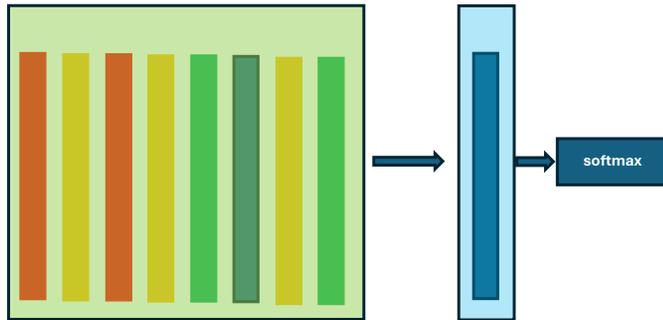


ResNet18: Ενσωματώσεις Γάτων/Σκύλων



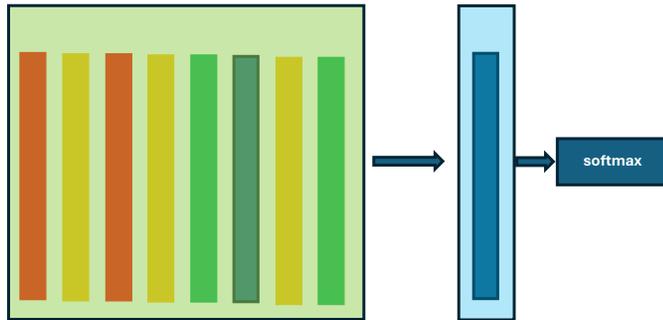
Εδώ απεικονίζουμε μια δισδιάστατη προβολή της ενσωμάτωσης που μας δίνει το ResNet18 (όλα τα επίπεδα εκτός από το τελευταίο) για 20 γάτες και 20 σκύλους από το CIFAR-10. Η προβολή δείχνει ότι στη συγκεκριμένη ενσωμάτωση τα σημεία είναι γραμμικά διαχωρίσιμα.

Convolutional Νευρωνικό Δίκτυο

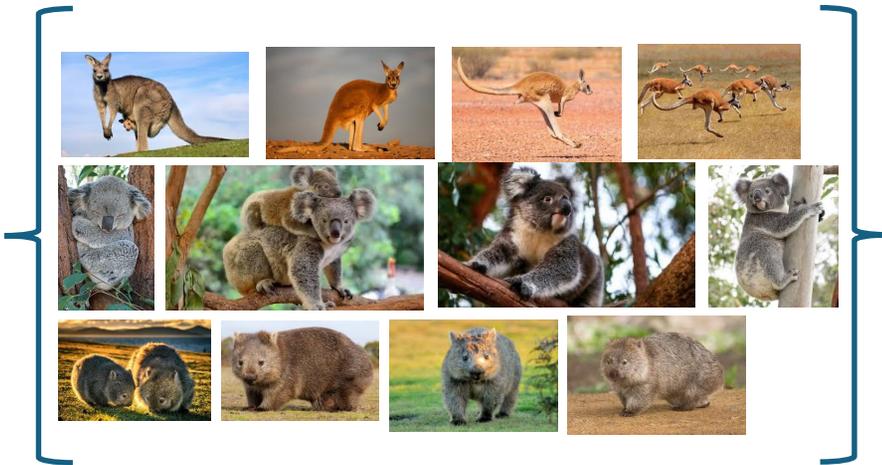


Δεδομένα {φασκλώμυς, κοάλα, καγκουρό}

Convolutional Νευρωνικό Δίκτυο

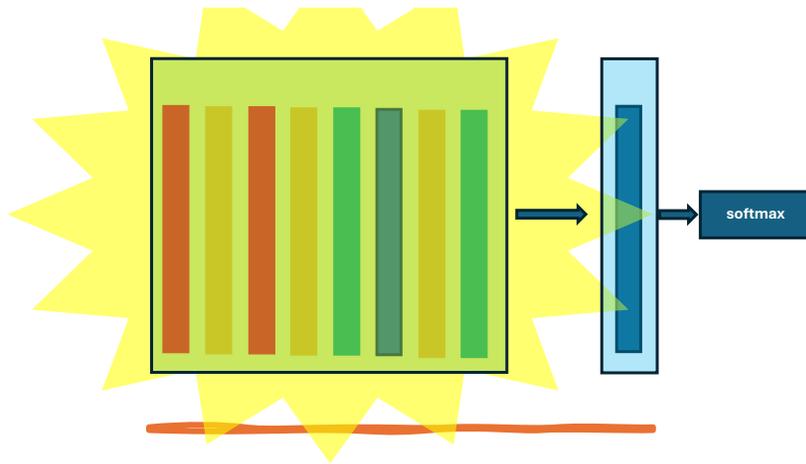


Δεδομένα {φασκλώμους, κοάλα, καγκουρό}

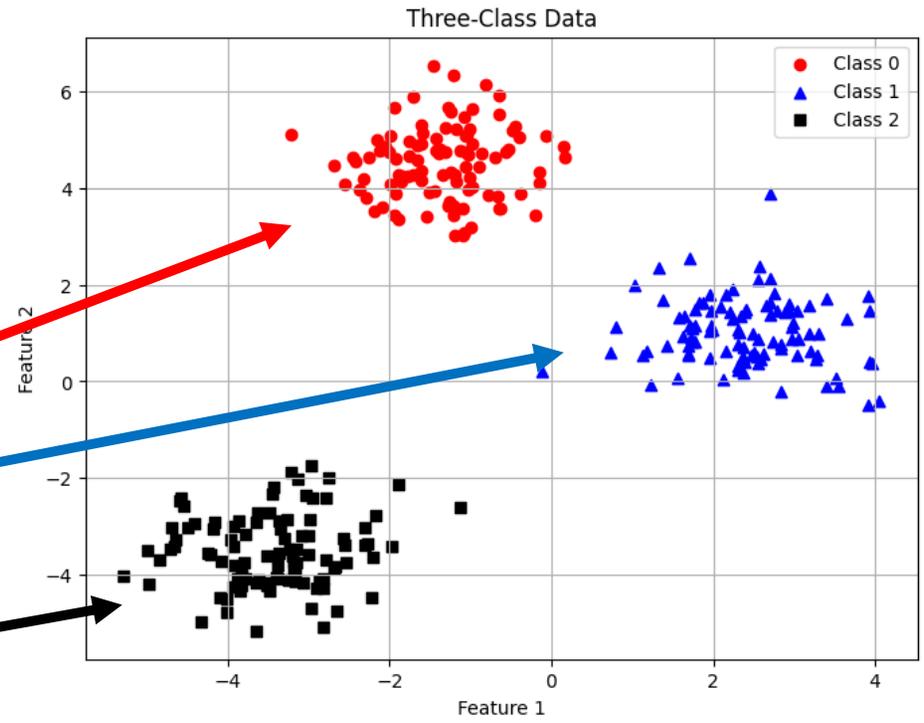
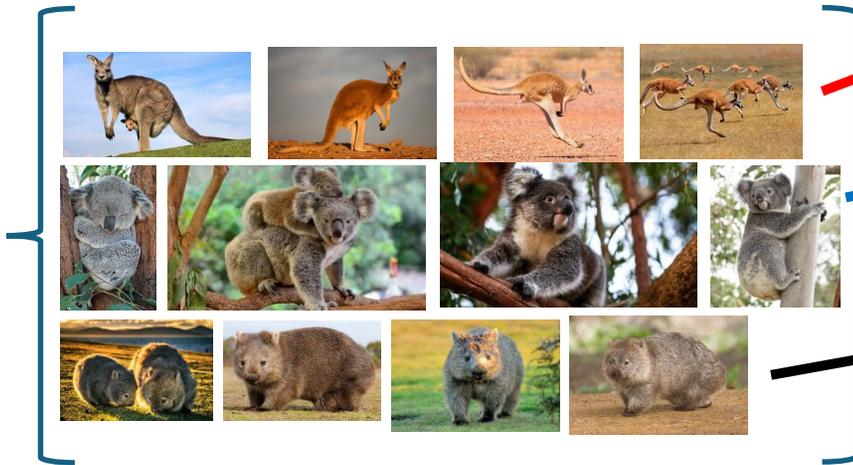


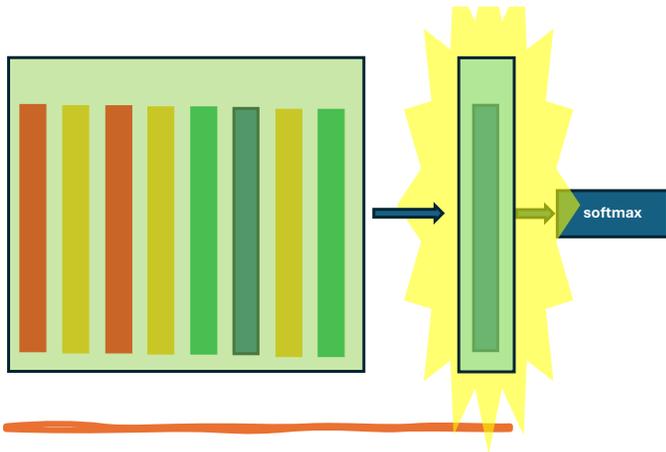
Κωνσταντίνος Καραμανής

Convolutional Νευρωνικό Δίκτυο

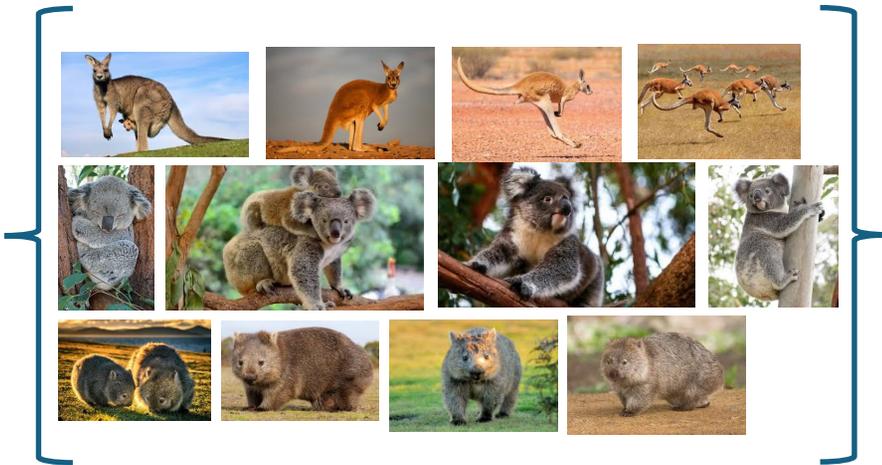


Δεδομένα {φασκλώμους, κοάλα, καγκουρό}



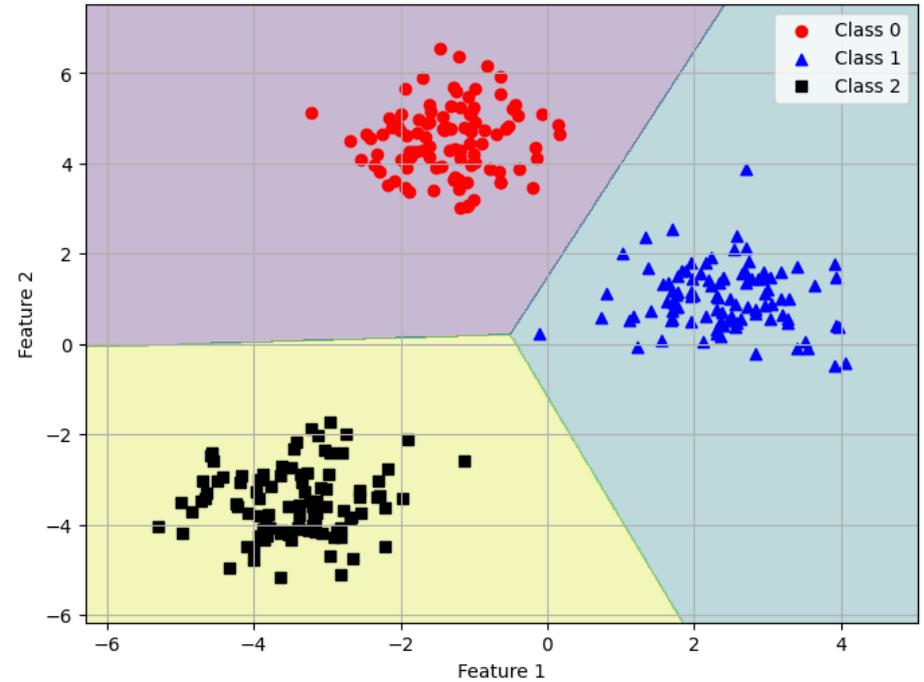


Δεδομένα {φασκλώμυς, κοάλα, καγκουρό}

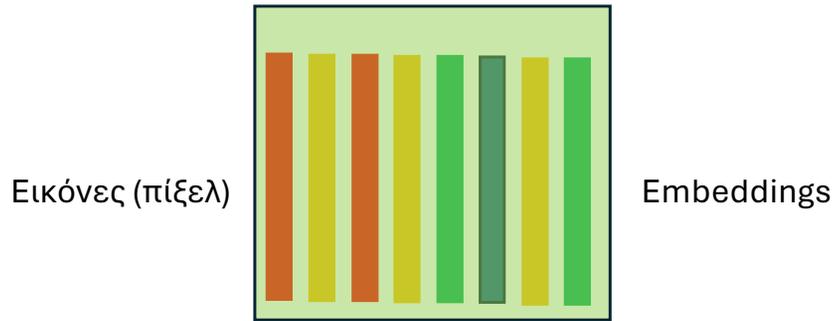


Convolutional Νευρωνικό Δίκτυο

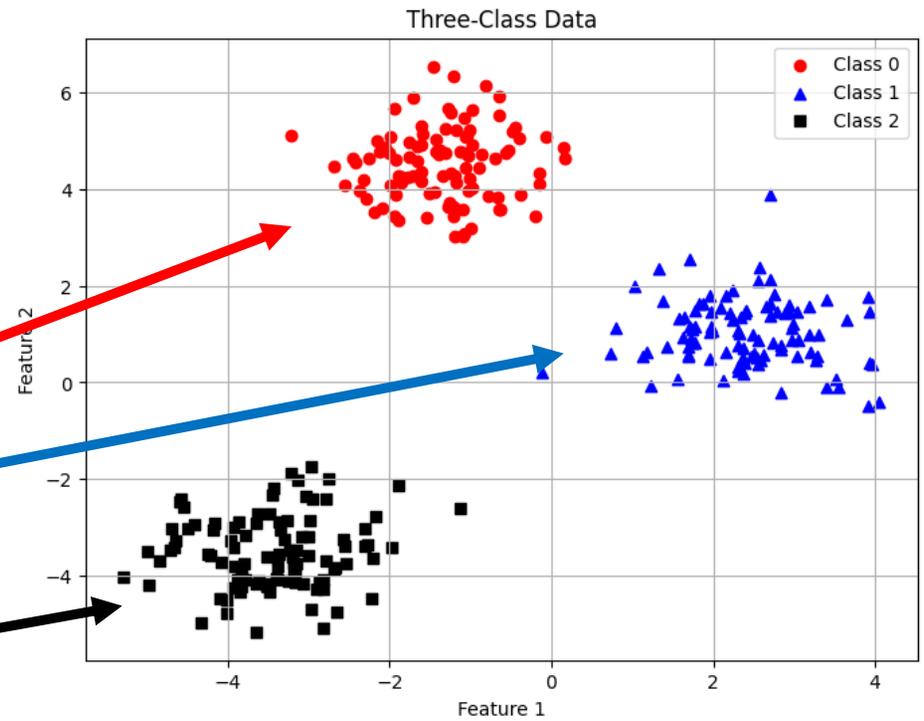
Three-Class Logistic Regression Decision Boundaries



Embeddings (Ενσωματώσεις)



Δεδομένα {φασκλώμους, κοάλα, καγκουρό}



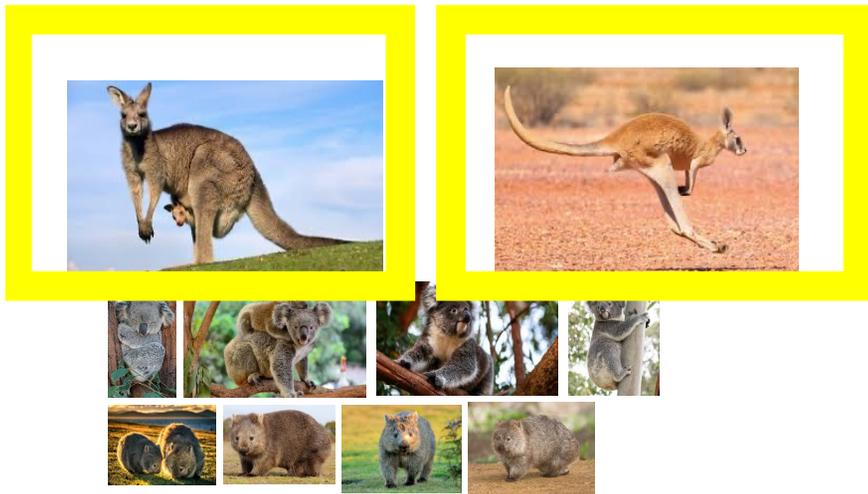
Embeddings: απόσταση στον γεωμετρικό χώρο αντικαθιστά την απόσταση που αντιλαμβανόμαστε



Embeddings: απόσταση στον γεωμετρικό χώρο αντικαθιστά την απόσταση που αντιλαμβανόμαστε

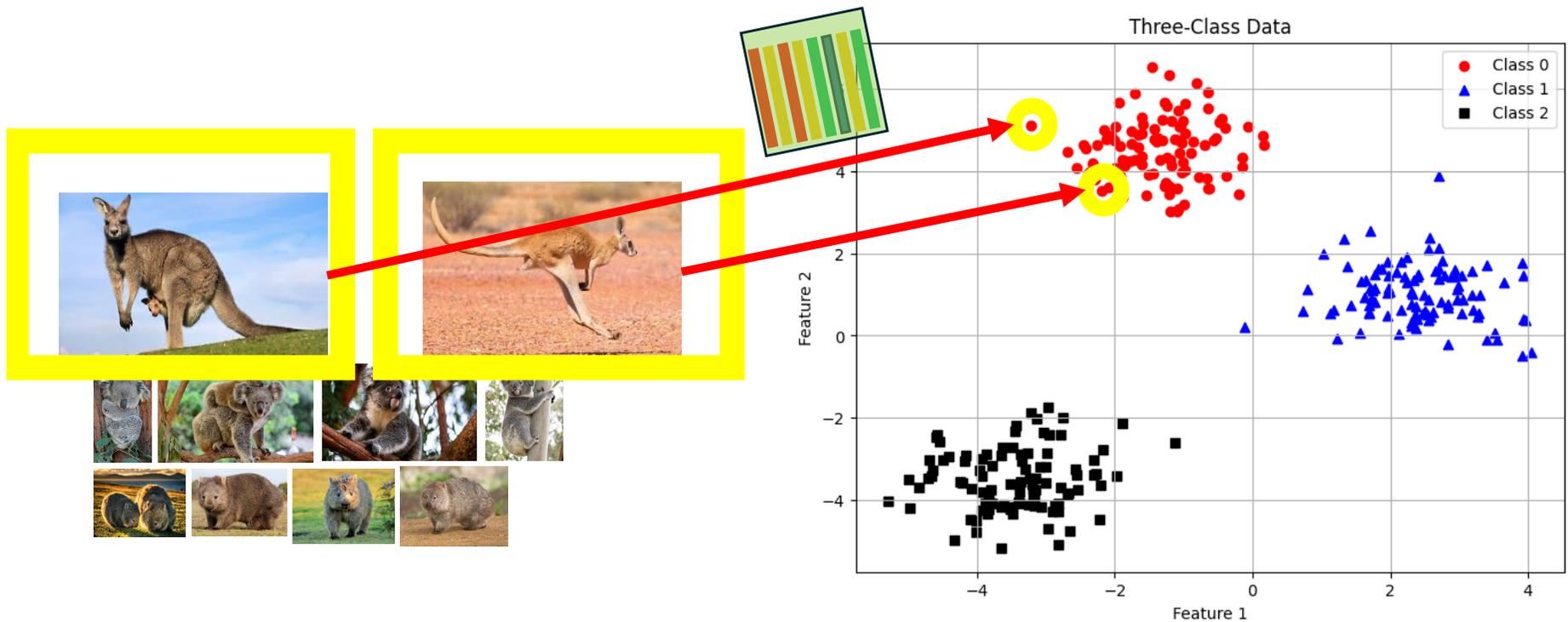


Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση

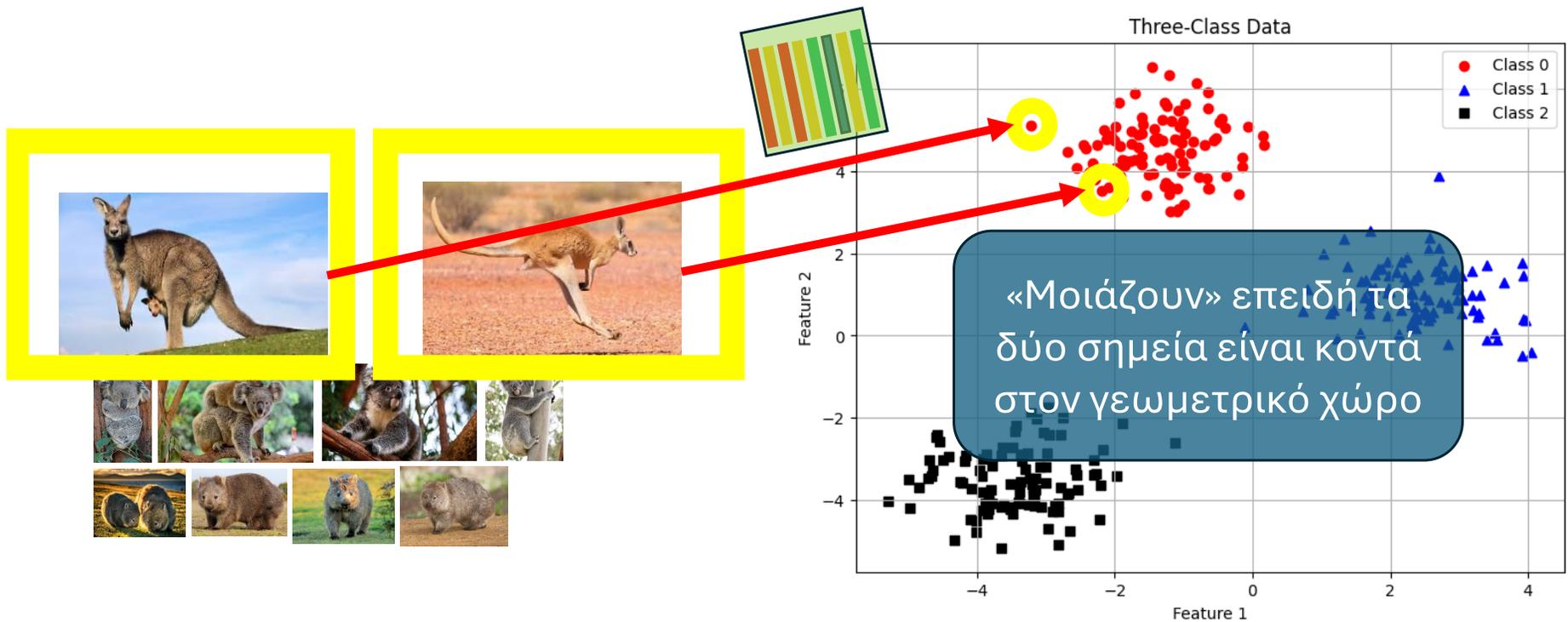


Πίξελ προς πίξελ δεν μοιάζουν καθόλου οι δύο εικόνες. «Μοιάζουν» επειδή απεικονίζουν και οι δύο καγκουρό

Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση

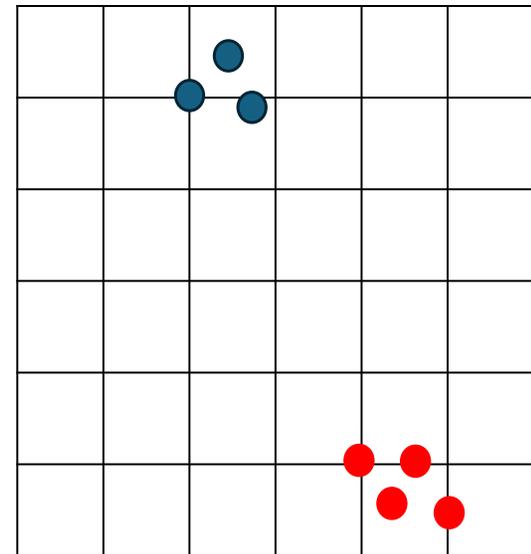
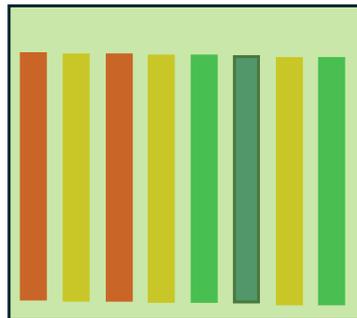


Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση

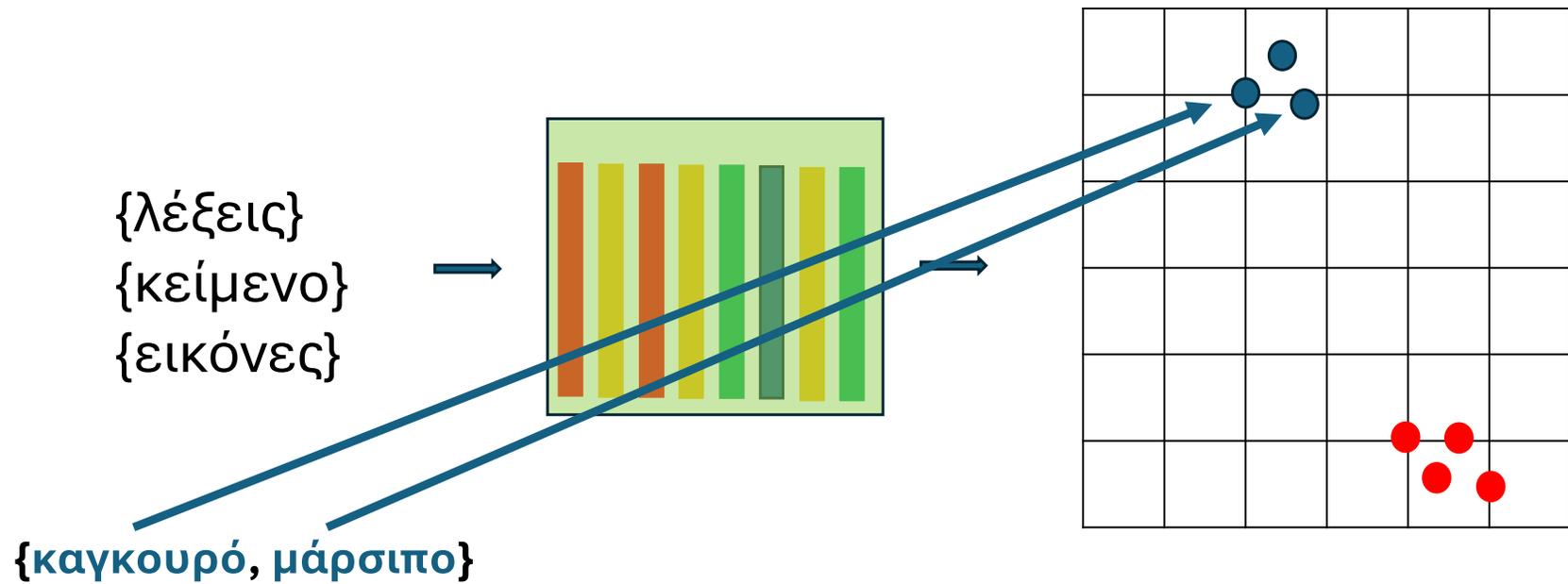


Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση

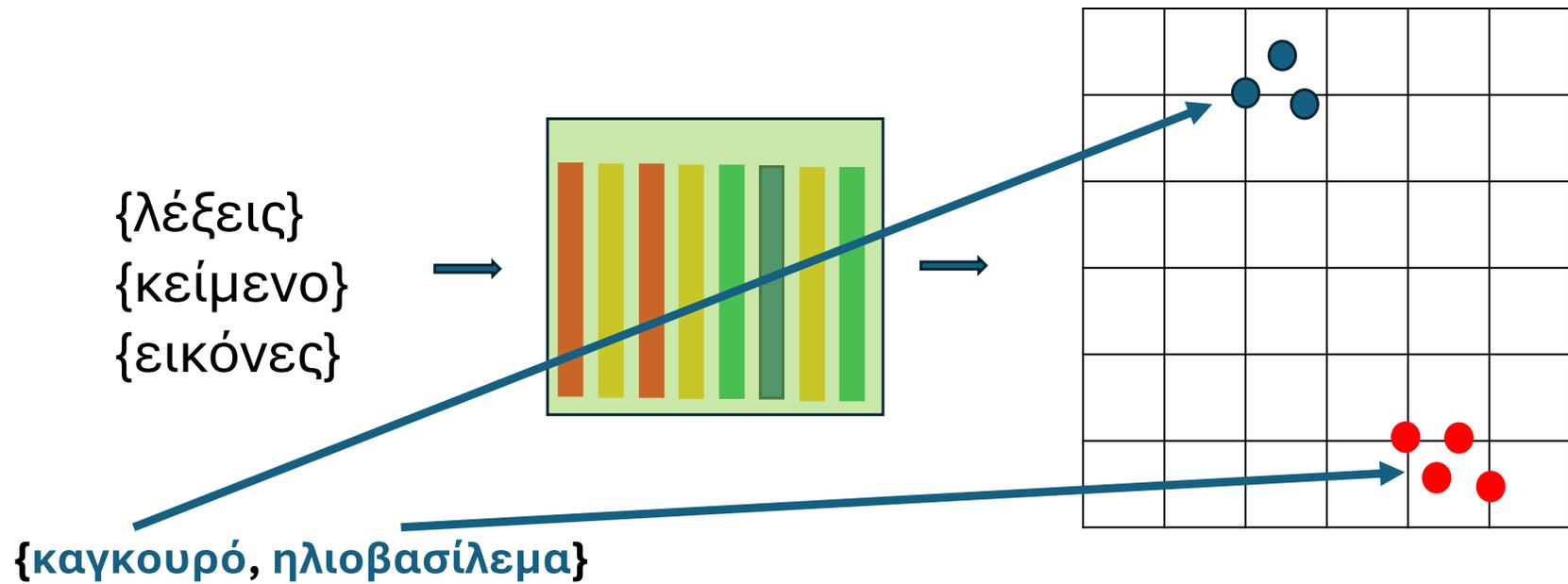
{λέξεις}
{κείμενο}
{εικόνες}



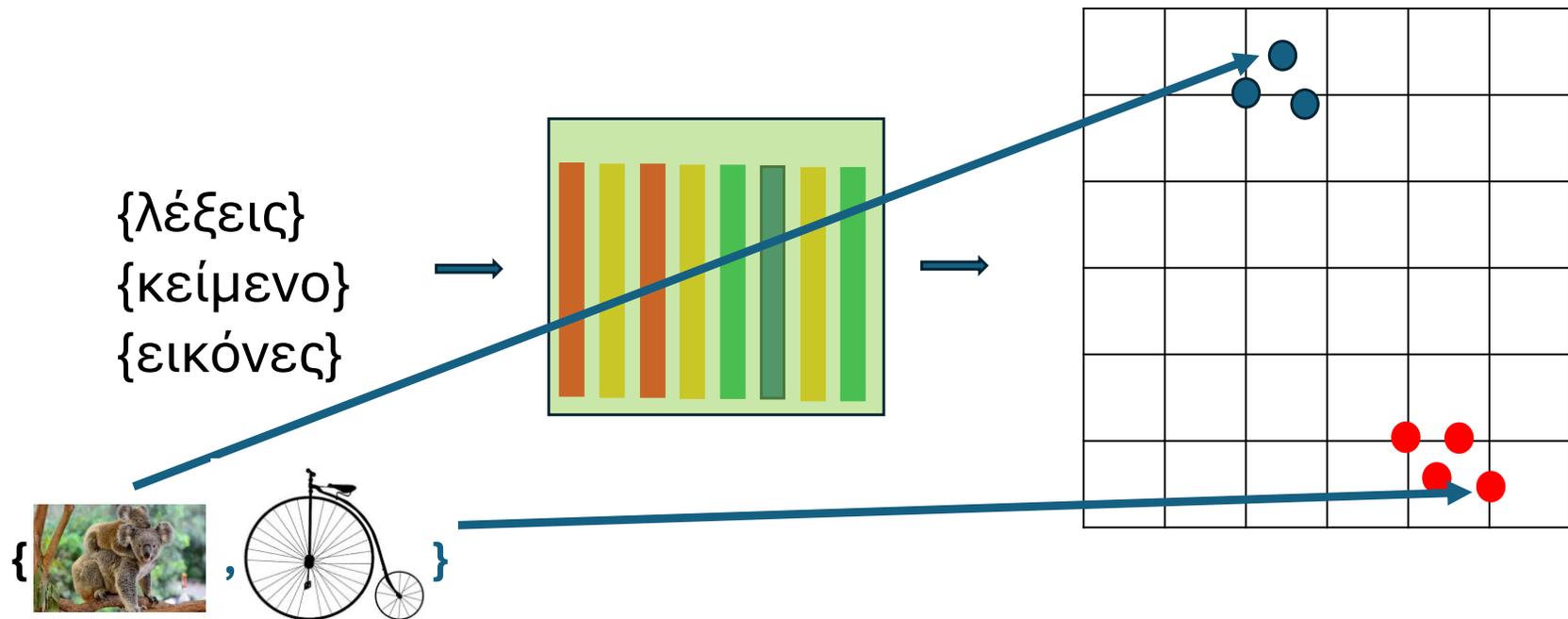
Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση



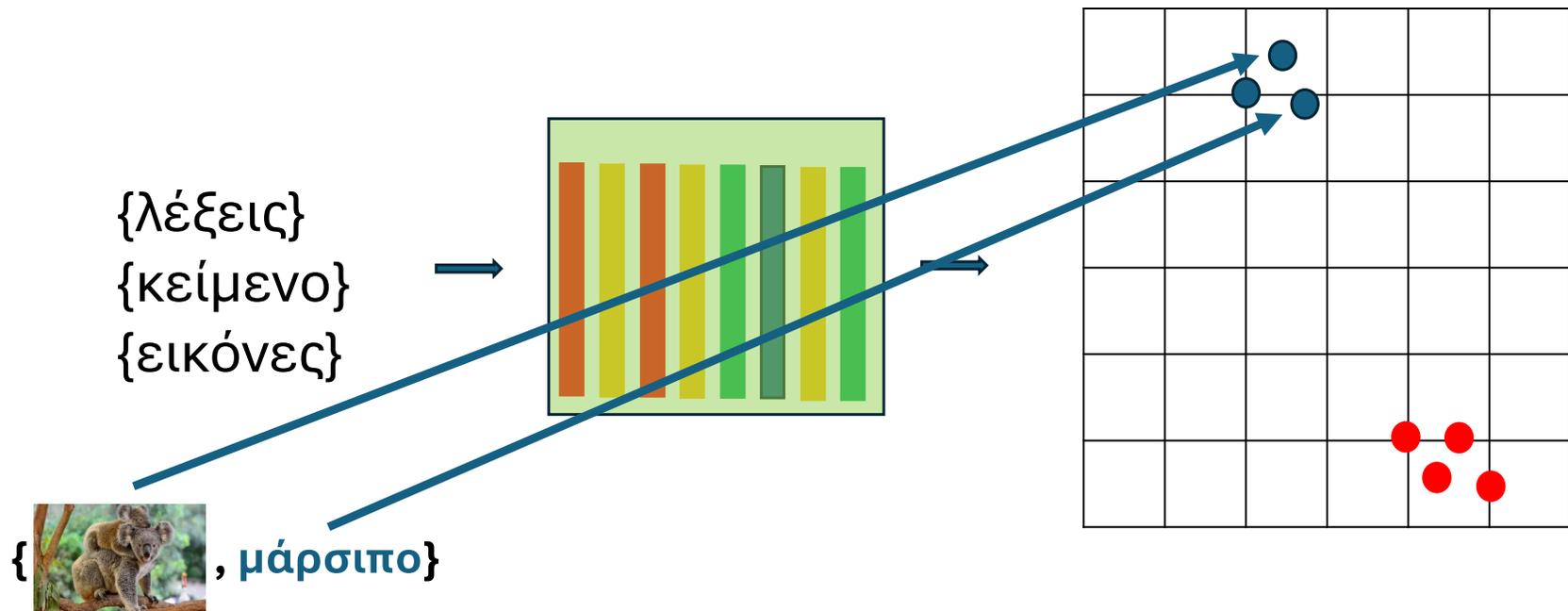
Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση



Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση



Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση



Embeddings/Ενσωματώσεις: η απόσταση στον γεωμετρικό χώρο αντικαθιστά την «σημασιολογική» απόσταση

Πώς δημιουργούμε καλές ενσωματώσεις (embeddings) για {λέξεις}, {κείμενο}, {εικόνες} εικόνες, λέξεις, κείμενο, ή για δεδομένα που συνδυάζουν και τις δύο μορφές;

