

Μηχανική Μάθηση και Εφαρμογές



Ιωάννης Χατζηλυγερούδης

Αφυπηρετήσας Καθηγητής

Τμήμα Μηχανικών Η/Υ & Πληροφορικής



Επιχειρησιακό Πρόγραμμα
Ανάπτυξη Ανθρώπινου Δυναμικού,
Εκπαίδευση και Διά Βίου Μάθηση
Ειδική Υπηρεσία Διαχείρισης

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Περιγραμμά Παρουσίασης

- Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)
- Βασικές Προσεγγίσεις και Συστήματα ΤΝ
- Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι
- Επιβλεπόμενη Μάθηση
- Νευρωνικά Δίκτυα
- Βαθιά Μάθηση-Εφαρμογές

Περιγραμμά Παρουσίασης

- **Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)**
- Βασικές Προσεγγίσεις και Συστήματα ΤΝ
- Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι
- Επιβλεπόμενη Μάθηση
- Νευρωνικά Δίκτυα
- Βαθιά Μάθηση-Εφαρμογές

Τι είναι Τεχνητή Νοημοσύνη (1)

- Barr and Feigenbaum

«ΤΝ είναι ο τομέας της επιστήμης των υπολογιστών, που ασχολείται με τη σχεδίαση ευφυών υπολογιστικών συστημάτων, δηλ. συστημάτων που επιδεικνύουν χαρακτηριστικά που σχετίζονται με τη νοημοσύνη στην ανθρώπινη συμπεριφορά.»

- Marvin Minsky

«ΤΝ είναι η επιστήμη που κάνει τις μηχανές να κάνουν πράγματα που θα απαιτούσαν ευφυΐα αν γινόταν από ένα άνθρωπο.»

Τι είναι Τεχνητή Νοημοσύνη (2)

- Elaine Rich

«ΤΝ είναι η μελέτη του πως να κάνουμε τους Η/Υ να κάνουν πράγματα για τα οποία, προς το παρόν, οι άνθρωποι είναι καλύτεροι.»

- Patrick Henry Winston

«ΤΝ είναι η μελέτη των υπολογιστικών μεθόδων που καθιστούν δυνατά την αντίληψη, τον συλλογισμό και την ενέργεια.»

Τι είναι Τεχνητή Νοημοσύνη (3)

- Η δυσκολία έγκειται στον ορισμό του όρου «νοημοσύνη» (ή «ευφυΐα»). Άλλοι ορισμοί τον ορίζουν μέσω του εαυτού του (π.χ. ορισμός Minsky). Άλλοι τον παρακάμπτουν (π.χ. ορισμός Rich).
- Επίσης, στο ότι η ΤΝ είναι σημείο συνάντησης περισσότερων της μιας επιστήμης:
 - Επιστήμη των υπολογιστών (computer science)
 - Γνωστική επιστήμη (cognitive science)
 - Φιλοσοφία (philosophy)

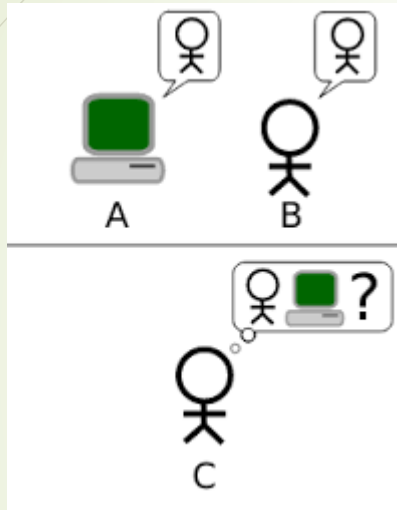
Τι είναι Τεχνητή Νοημοσύνη (4)

Οι Russell και Norvig διακρίνουν τους ορισμούς της ΤΝ σε τέσσερις κατηγορίες, με βάση το αν χαρακτηρίζουν ένα σύστημα ως ευφυές με κριτήριο το

- αν σκέφτεται σαν άνθρωπος (Μηχανισμός, Γνωστική Επιστήμη)
- αν ενεργεί σαν άνθρωπος (Συμπεριφορά, Turing test)
- αν σκέφτεται ορθολογικά (Μηχανισμός, Νόμοι Ορθής Σκέψης)
- αν ενεργεί ορθολογικά (Συμπεριφορά, Ορθολογικοί Πράκτορες)

Γενικά μπορούμε να πούμε ότι η ΤΝ ασχολείται με τη δημιουργία συστημάτων που λειτουργούν ή ενεργούν με τρόπο που αντιστοιχεί με τον τρόπο που λειτουργεί ή ενεργεί ένας άνθρωπος για **επίλυση προβλημάτων**.

Το παιχνίδι της μίμησης (Turing Test)



Το επιχείρημα του κινέζικου δωματίου (J. Searle)



Περιγραμμά Παρουσίασης

- Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)
- **Βασικές Προσεγγίσεις και Συστήματα ΤΝ**
- Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι
- Επιβλεπόμενη Μάθηση
- Νευρωνικά Δίκτυα
- Βαθιά Μάθηση-Εφαρμογές

Βασικές Προσεγγίσεις ΤΝ

- **Κλασσική ή Συμβολική προσέγγιση (symbolic approach)**

Στηρίζεται στην υπόθεση ότι η ευφυής συμπεριφορά παράγεται από τη διαχείριση συμβόλων, τα οποία παριστάνουν έννοιες και σχέσεις μεταξύ τους. Σχετίζεται με τη μέθοδο των λογικών συλλογισμών του Αριστοτέλη και μιμείται τη λειτουργία της ανθρώπινης σκέψης-συλλογιστικής.

Εκπρόσωποι: Λογική (Logic), Κανόνες (Rules), Οντολογίες (Ontologies)

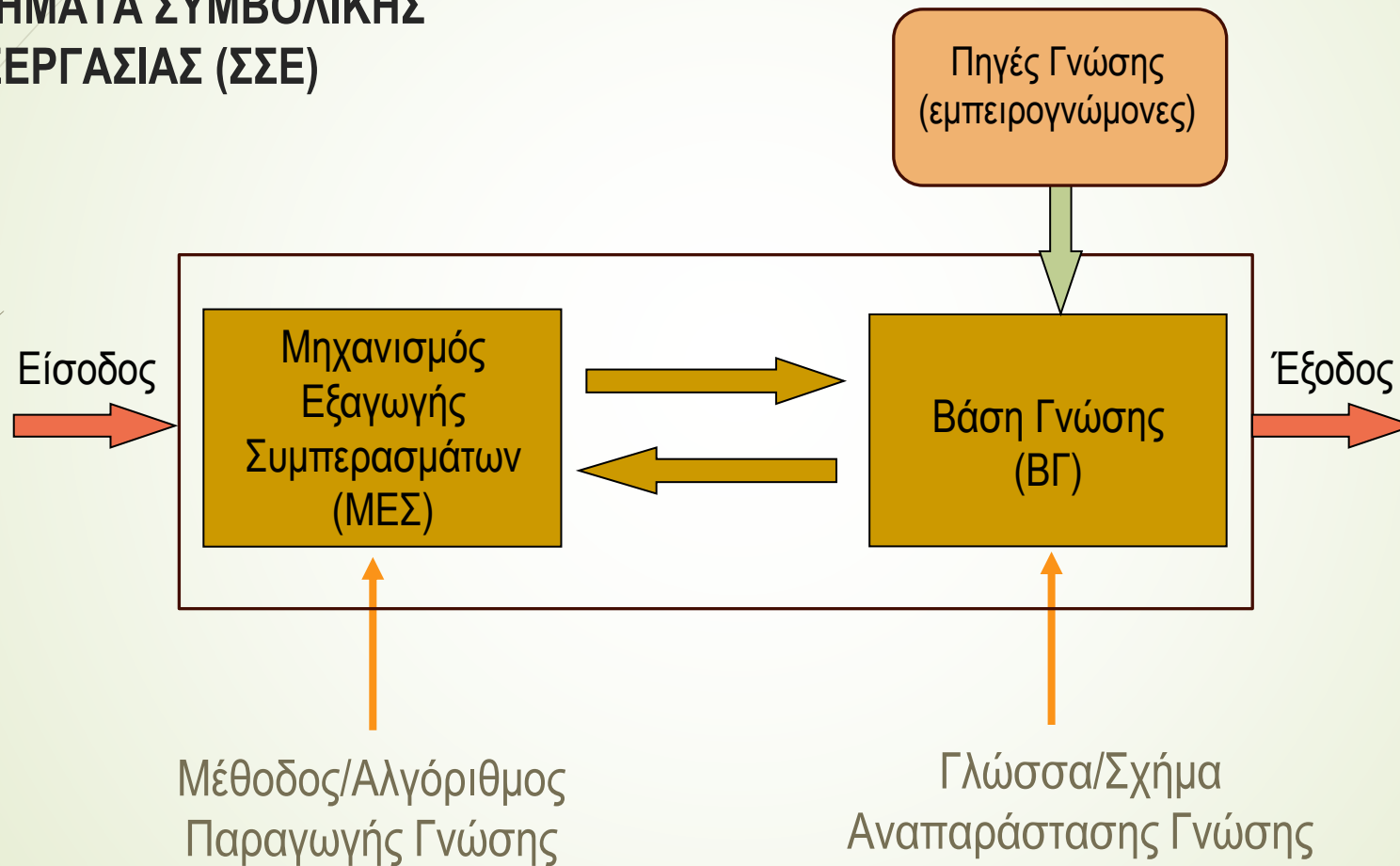
- **Μη συμβολική ή Υπολογιστική προσέγγιση (non-symbolic or computational approach)**

Η ευφυής συμπεριφορά παράγεται από τη μίμηση βιολογικών διεργασιών, όπως π.χ. η λειτουργία του εγκεφάλου, η εξέλιξη των ειδών, η συμπεριφορά των εντόμων κλπ.

Εκπρόσωποι: Συνδεδετισμός (Connectionism) ή Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks), Γενετικοί Αλγόριθμοι (Genetic Algorithms), Μέθοδοι Νοημοσύνης Σμήνους (Swarm Intelligence)

Βασικές Αρχιτεκτονικές Συστημάτων ΤΝ (1)

ΣΥΣΤΗΜΑΤΑ ΣΥΜΒΟΛΙΚΗΣ ΕΠΕΞΕΡΓΑΣΙΑΣ (ΣΣΕ)



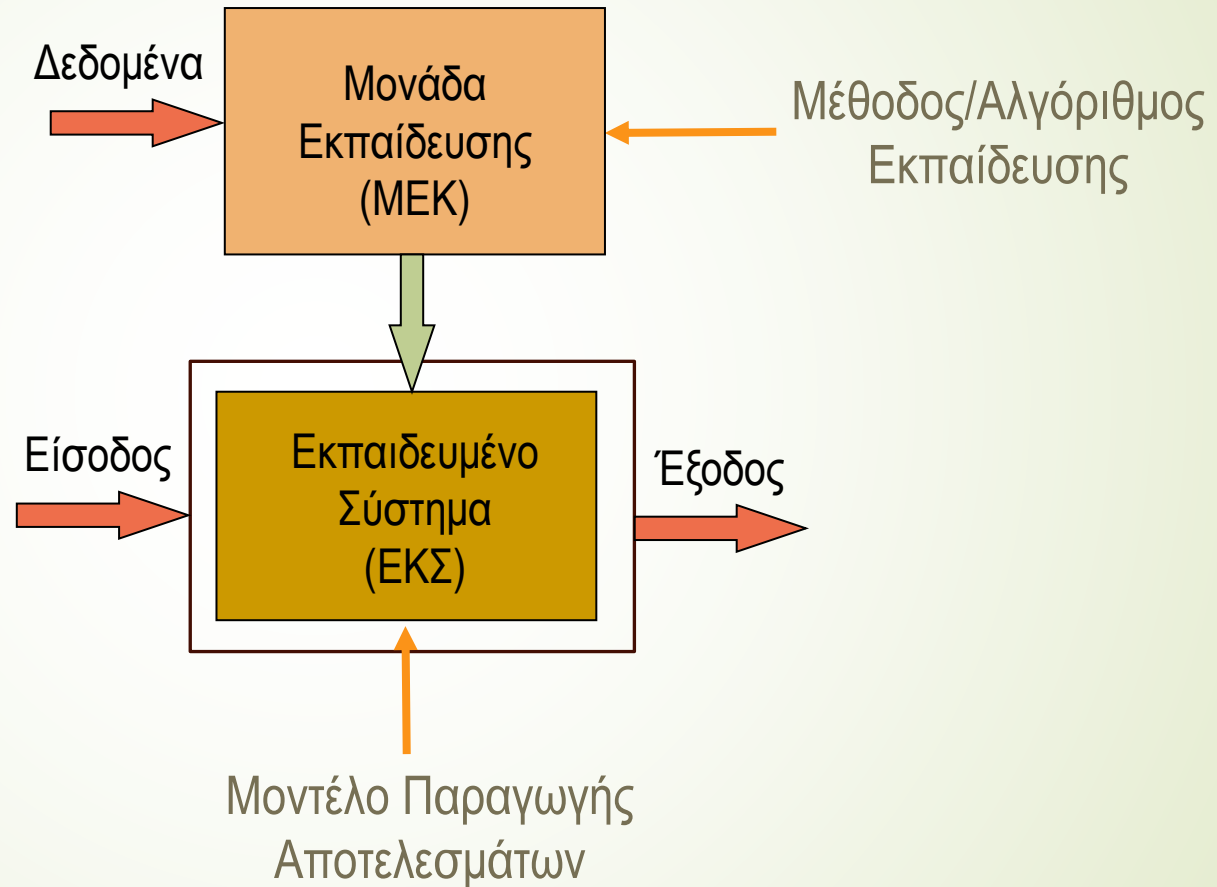
Βασικές Αρχιτεκτονικές Συστημάτων ΤΝ (2)

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΣΣΕ

- ✓ Απόκτηση γνώσης από εμπειρογνώμονες ή/και άλλες πηγές.
- ✓ Αποτύπωση της γνώσης σε μια γλώσσα αναπαράστασης γνώσης (Λογική, Κανόνες, Οντολογίες κλπ).
- ✓ Εξαγωγή συμπερασμάτων (inferencing) μέσω συμβολικού υπολογισμού (επεξεργασία συμβόλων).
- ✓ Δεν είναι απαραίτητη όλη η διαθέσιμη πληροφορία-γνώση, αλλά δεν μπορεί να είναι ελλιπής.
- ✓ Διαδραστική καθοδήγηση της εισόδου δεδομένων στο σύστημα.
- ✓ Επεξήγηση των συμπερασμάτων (εξόδου).
- ✓ Τμηματοποιημένη δομή (ΒΓ).

Βασικές Αρχιτεκτονικές Συστημάτων ΤΝ (3)

ΣΥΣΤΗΜΑΤΑ ΜΗ ΣΥΜΒΟΛΙΚΗΣ
ΕΠΕΞΕΡΓΑΣΙΑΣ (ΣΜΣΕ)

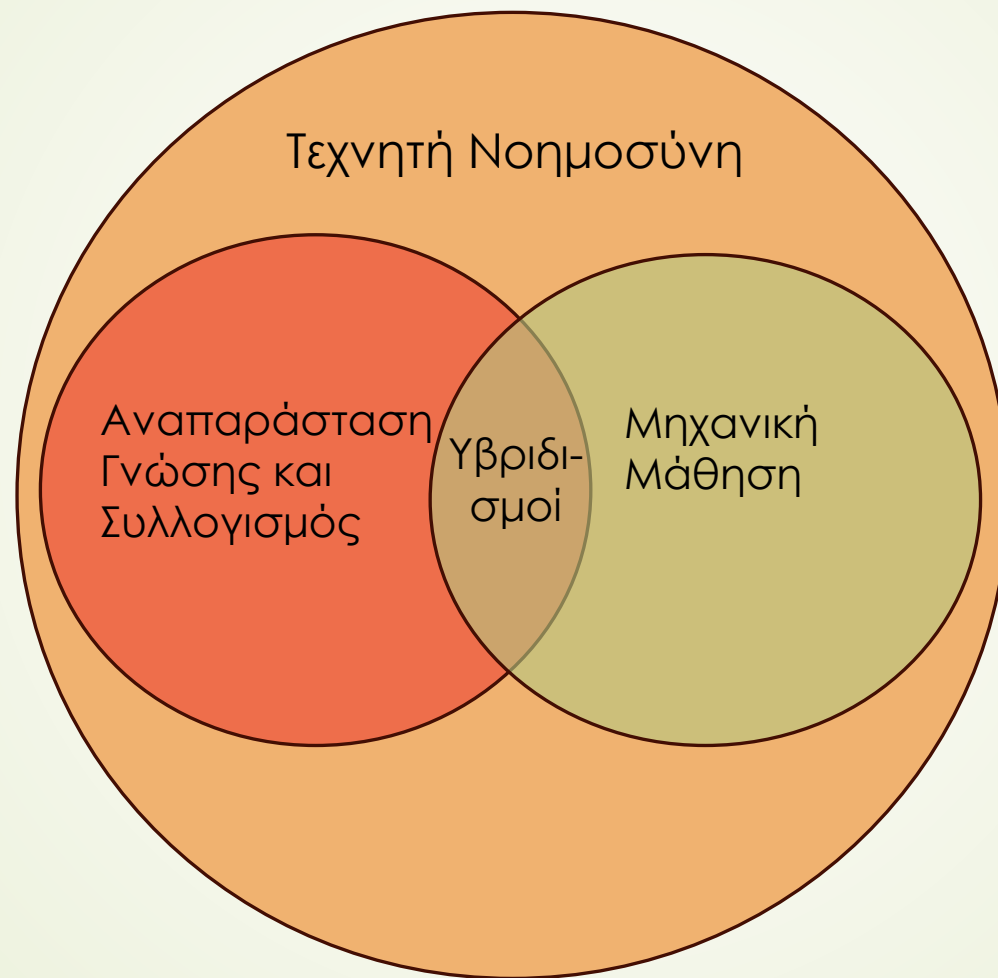


Βασικές Αρχιτεκτονικές Συστημάτων ΤΝ (4)

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΣΜΣΕ

- ✓ Απαιτείται εκπαίδευση του συστήματος με τη χρήση δεδομένων.
- ✓ Παραγωγή αποτελεσμάτων μέσω ενός εκπαιδευμένου μοντέλου (στατιστικού, δέντρου, κανόνων κλπ).
- ✓ Παραγωγή αποτελεσμάτων μπορεί να γίνει και από ελλιπή γνώση (στα περισσότερα μοντέλα).
- ✓ Δεν υπάρχει διαδραστική καθοδήγηση της εισόδου δεδομένων στο σύστημα.
- ✓ Δεν παρέχεται επεξήγηση των συμπερασμάτων (εξόδου)-το σύστημα λειτουργεί σαν μαύρο κουτί.

Βασικές Περιοχές ΤΝ



Περιγραμμά Παρουσίασης

- Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)
- Βασικές Προσεγγίσεις και Συστήματα ΤΝ
- **Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι**
- Επιβλεπόμενη Μάθηση
- Νευρωνικά Δίκτυα
- Βαθιά Μάθηση-Εφαρμογές

Μηχανική Μάθηση-ΜΜ (Machine Learning-ML)

Είναι μια περιοχή της ΤΝ που ασχολείται με **αλγορίθμους** που δημιουργούν **υπολογιστικά μοντέλα** τα οποία έχουν τη δυνατότητα να **μαθαίνουν** να επιλύουν προβλήματα και να **βελτιώνονται** μέσω χρήσης **εμπειρικών δεδομένων** χωρίς να έχουν προγραμματιστεί εκ των προτέρων γι' αυτό.

“A computer program is said to learn from experience E with respect to some task T and some performance measure P , if its performance on T , as measured by P , improves with experience E .” (Mitchel, 1997)

“Ένα πρόγραμμα (ή σύστημα) υπολογιστή λέμε ότι μαθαίνει από μια εμπειρία E σε σχέση με κάποια εργασία T και κάποια μετρική απόδοσης P , αν η απόδοσή του στην T , όπως μετριέται από την P , βελτιώνεται μέσω της εμπειρίας E . (Mitchel, 1997)

Τι Προβλήματα μπορεί να επιλύσει η ΜΜ;

- **Κατηγοριοποίησης ή Ταξινόμησης (Classification)**

Αναφέρονται στην εύρεση της κλάσης, ανάμεσα σε δύο ή περισσότερες, στην οποία ανήκει μια περίπτωση σχετικών δεδομένων. Π.χ. η εύρεση της ασθένειας της καρδιάς με βάση τα ιατρικά δεδομένα ενός ασθενούς. Μπορεί να εμφανίζονται ως προβλήματα **διάγνωσης** ή **πρόβλεψης**.

- **Συναρτησιακής Προσέγγισης ή Παλινδρόμησης (Regression)**

Αναφέρονται στην εύρεση της τιμής μιας άγνωστης συνάρτησης (που αντιπροσωπεύεται από ένα σύνολο δεδομένων) όταν δίνονται οι τιμές των παραμέτρων της. Π.χ. πρόβλεψη των θανάτων των ιχθύων σε ένα ιχθυοτροφείο με βάση τα δεδομένα μια περιόδου εκτροφής. Εμφανίζονται συνήθως ως προβλήματα **πρόβλεψης**.

- **Ομαδοποίησης (Clustering)**

Αναφέρονται στην εύρεση ομάδων συγγενών περιπτώσεων δεδομένων (κλάσεων) σε μια περιοχή γνώσης μέσω ενός συνόλου σχετικών περιπτώσεων. Π.χ. εύρεση κατηγοριών κειμένων ομοειδούς περιεχομένου μέσω ενός συνόλου διάφορων κειμένων.

- **Σχεδίασης ενεργειών (Planning)**

Αναφέρονται στην εύρεση μιας σειράς ενεργειών για την επίτευξη ενός στόχου. Π.χ. η εκμάθηση ενός ρομπότ για την μεταφορά ενός αντικειμένου.

Μέθοδοι Μάθησης (1)

- **Επιβλεπόμενη Μάθηση (Supervised Learning)**

Στους αλγορίθμους επιβλεπόμενης μάθησης δίνεται η είσοδος και η (σωστή) έξοδος. Με βάση αυτές τις πληροφορίες παράγουν το μοντέλο του συστήματος. Οι αλγόριθμοι αυτοί είναι κατάλληλοι για **προβλήματα ταξινόμησης και συναρτησιακής προσέγγισης**.

- **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)**

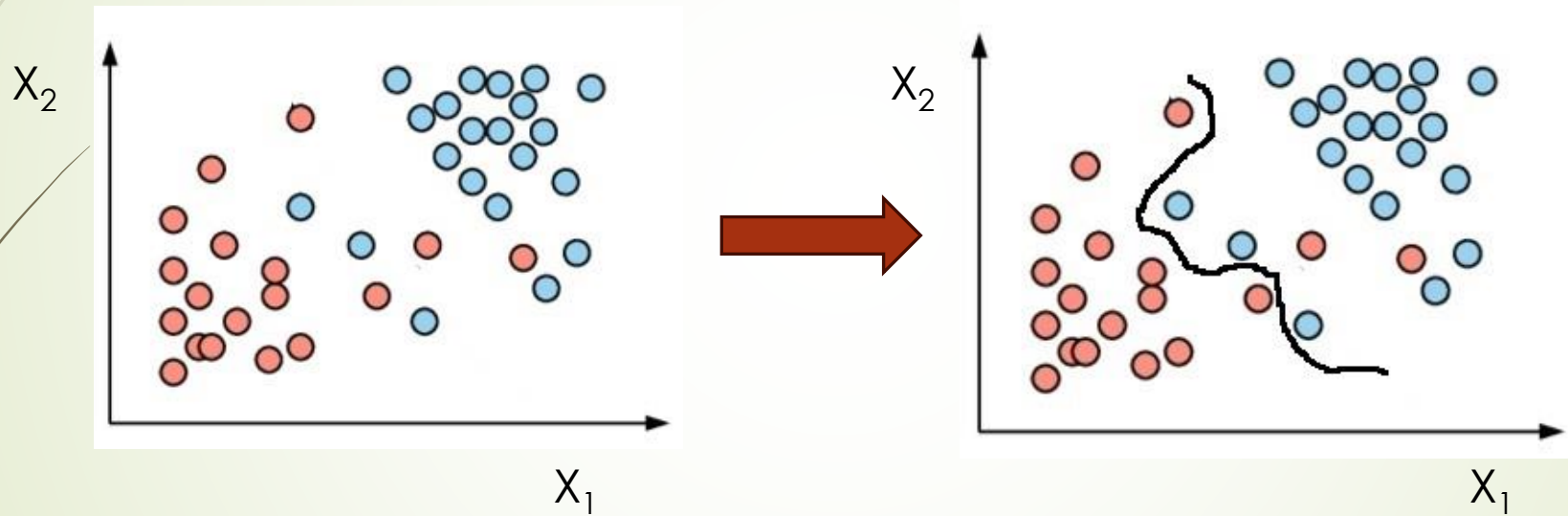
Στους αλγορίθμους μη επιβλεπόμενης μάθησης δίνεται η είσοδος χωρίς την (σωστή) έξοδο, ούτε άλλη πληροφόρηση. Οι αλγόριθμοι αυτοί καλούνται να βρουν παραδείγματα δεδομένων με ίδια ή παρόμοια έξοδο. Είναι κατάλληλοι για **προβλήματα ομαδοποίησης**.

- **Ενισχυτική Μάθηση (Reinforcement Learning)**

Στους αλγορίθμους ενισχυτικής μάθησης χρησιμοποιείται η έννοια της ανταμοιβής (reward). Όσο πιο πετυχημένη είναι η έξοδος τόσο μεγαλύτερη είναι και η ανταμοιβή. Αυτό τελικά οδηγεί σε πετυχημένες εξόδους. Τέτοιοι αλγόριθμοι είναι κατάλληλοι για **προβλήματα σχεδίασης ενεργειών**.

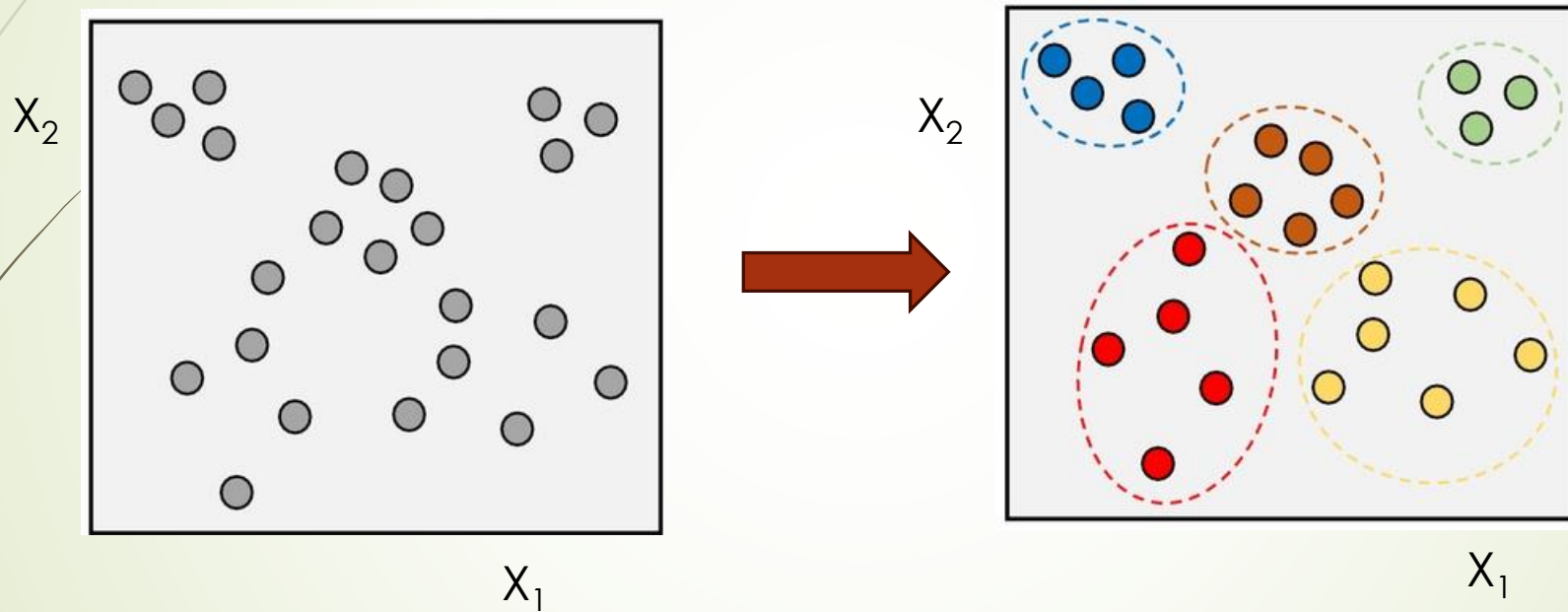
Μέθοδοι Μάθησης (2)

ΕΠΙΒΛΕΠΟΜΕΝΗ ΜΑΘΗΣΗ (ταξινόμηση)



Μέθοδοι Μάθησης (3)

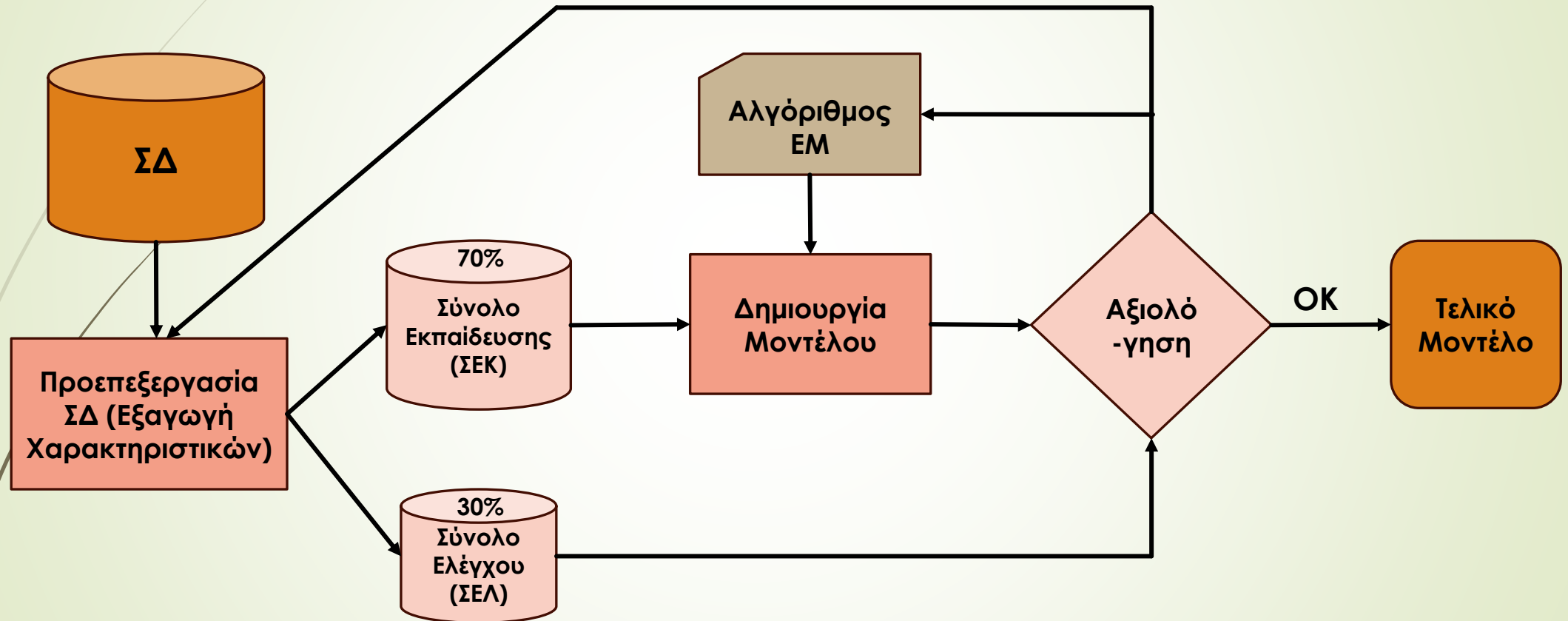
ΜΗ ΕΠΙΒΛΕΠΟΜΕΝΗ ΜΑΘΗΣΗ (ομαδοποίηση)



Περιγραμμά Παρουσίασης

- Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)
- Βασικές Προσεγγίσεις και Συστήματα ΤΝ
- Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι
- **Επιβλεπόμενη Μάθηση**
- Νευρωνικά Δίκτυα
- Βαθιά Μάθηση-Εφαρμογές ΜΜ

Επιβλεπόμενη Μάθηση (ΕΜ)



Αλγόριθμοι Επιβλεπόμενης Μάθησης

- **Naïve Bayes (Απλοϊκή Μέθοδος Bayes)**

Δημιουργεί ένα πιθανοτικό μοντέλο με βάση τον Νόμο του Bayes, ώστε να ταξινομεί σωστά τα δεδομένα. Η κλάση με την μεγαλύτερη πιθανότητα επιλέγεται ως απάντηση κάθε φορά.

- **Decision Trees (Δέντρα Αποφάσεων)**

Δημιουργούν δέντρα, των οποίων τα φύλλα αντιστοιχούν σε κλάσεις, μέσω των οποίων φιλτράρονται οι προς αξιολόγηση περιπτώσεις, ώστε να καταλήγουν στο σωστό φύλλο.

- **Rules Extraction (Εξαγωγής Κανόνων)**

Παράγουν ένα σύνολο κανόνων μέσω των οποίων εξάγονται οι σωστές κλάσεις.

- **Support Vector Machines (Μηχανές Διανυσμάτων Υποστήριξης)**

Βρίσκει το κατάλληλο υπερεπίπεδο στον N-διάστατο χώρο ώστε να ταξινομούνται σωστά τα δεδομένα (N = αριθμός χαρακτηριστικών).

- **Neural Networks (Νευρωνικά Δίκτυα)**

Αποτελούνται από ένα σύνολο εκπαιδευμένων υπολογιστικών νευρώνων συνδεδεμένων μεταξύ τους, έτσι ώστε να υλοποιούν μια κατάλληλη (πολύπλοκη συνήθως) συνάρτηση, η οποία να ταξινομεί σωστά τα δεδομένα.

Σύνολο Δεδομένων-ΣΔ (Dataset)

- Αποτελείται από εμπειρικά δεδομένα (π.χ. δεδομένα από τις καρτέλες ασθενών με νοσήματα καρδιάς από ένα νοσοκομείο).
- Χρειάζεται **προεπεξεργασία** των εμπειρικών δεδομένων (αφαίρεση περιττών πληροφοριών, διαχείριση ελλিপών τιμών, **επιλογή χαρακτηριστικών** κ.α.)
- Ένα προεπεξεργασμένο ΣΔ αποτελείται από **παραδείγματα** (examples). Κάθε παράδειγμα είναι μια ακολουθία τιμών. Κάθε τιμή/θέση στην ακολουθία αντιστοιχεί σε μια ιδιότητα ή **χαρακτηριστικό** (feature) που σχετίζεται με το πρόβλημα στο οποίο αναφέρεται το ΣΔ.
- Το τελευταίο χαρακτηριστικό στην ακολουθία ονομάζεται **χαρακτηριστικό στόχου** ή **κλάση** και (πρέπει να) αποτελεί το αποτέλεσμα (έξοδο) του μοντέλου αν χρησιμοποιηθούν ως είσοδοι οι τιμές των υπόλοιπων χαρακτηριστικών της ακολουθίας. Τα δεδομένα στα οποία υπάρχει χαρακτηριστικό στόχου, λέγονται **επισημασμένα** (labelled).
- Τα χαρακτηριστικά μπορεί να είναι διαφόρων τύπων: κατηγορικά (categorical) ή διακριτά (discrete), αριθμητικά (numerical) ή συνεχή (continuous).

Σύνολο Δεδομένων-Παράδειγμα (1)

Χαρακτηριστικά

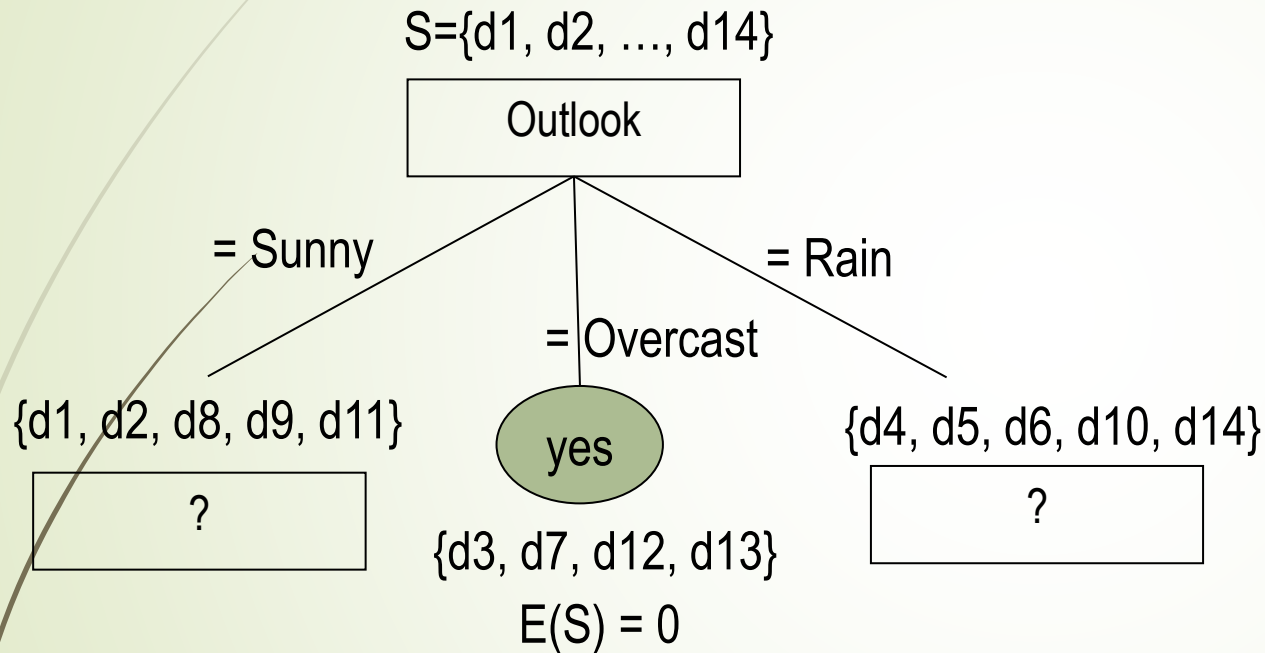
Χαρακτηριστικό-Στόχος

No	Outlook	Temp.	Humid.	Wind	PlayTennis
1	Sunny	30°C	High	Weak	No
2	Sunny	29°C	High	Strong	No
3	Overcast	27°C	High	Weak	Yes
4	Rainy	20°C	High	Weak	Yes
5	Rainy	13°C	Normal	Weak	Yes
6	Rainy	15°C	Normal	Strong	No
7	Overcast	12°C	Normal	Strong	Yes
8	Sunny	18°C	High	Weak	No
9	Sunny	10°C	Normal	Weak	Yes
10	Rainy	22°C	Normal	Weak	Yes
...

Σύνολο Δεδομένων-Παράδειγμα (2)

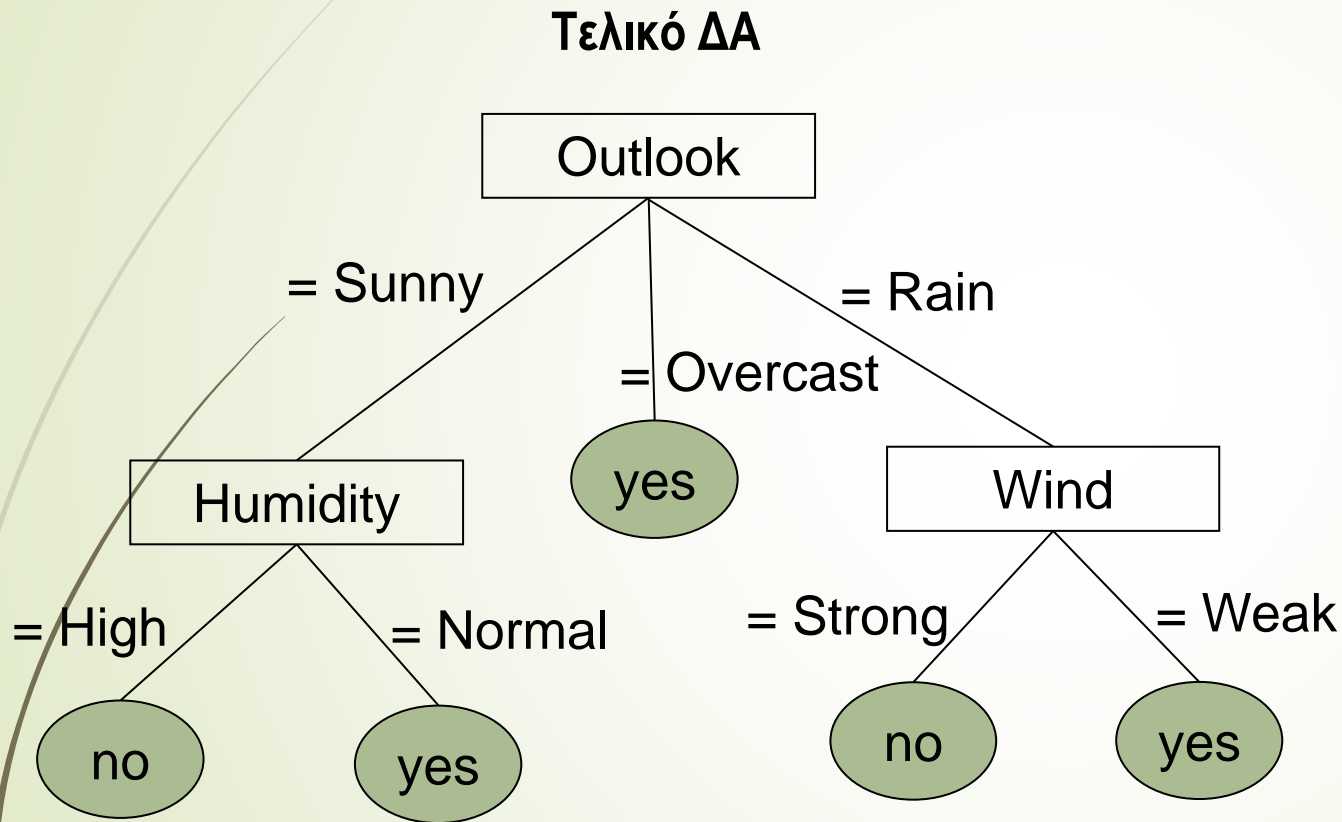
No	Outlook	Temp.	Humid.	Wind	PlayTennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rainy	Mild	High	Weak	Yes
5	Rainy	Cool	Normal	Weak	Yes
6	Rainy	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rainy	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rainy	Mild	High	Strong	No

Αλγόριθμοι Δέντρων Αποφάσεων (ID3, C4.5)



- Χρησιμοποιούν μια μετρική (κέρδος πληροφορίας, λόγος κέρδους) για να επιλέξουν σε ποιο κόμβο του δέντρου θα μπει ποιο χαρακτηριστικό.
- Ο στόχος είναι να διασπάσουν το αρχικό ΣΔ σε ομοιογενή υποσύνολα.

Αλγόριθμοι Δέντρων Αποφάσεων (ID3, C4.5)



- Παρατηρείστε ότι δεν χρησιμοποιείται καθόλου το χαρακτηριστικό "Temperature".

Μετατροπή Δέντρων σε Κανόνες

Εξαγόμενοι Κανόνες

if outlook is sunny and
and humidity is high
then playtennis is no

if outlook is rain and
and wind is strong
then playtennis is no

if outlook is sunny and
and humidity is normal
then playtennis is yes

if outlook is rain and
and wind is weak
then playtennis is yes

if outlook is overcast
then playtennis is yes

Αξιολόγηση Αλγορίθμων-Μοντέλων (1)

Μετρικές (περιπτώσεις δυαδικής εξόδου)

- ✓ **Accuracy** = $(TP+TN)/(TP+FN+FP+TN)$
(Ορθότητα)
- ✓ **Sensitivity** = $TP/(TP+FN)$ (Ευαισθησία)
- ✓ **Specificity** = $TN/(TN+FP)$ (Εξειδίκευση).
- ✓ **Precision** = $TP/(TP+FP)$ (Ακρίβεια)
(Τιμές στο διάστημα $[0,1]$)

- ✓ TP (true positives): πλήθος περιπτώσεων που ανήκουν σε μια κλάση C και ταξινομήθηκαν σ' αυτή .
- ✓ FN (false negatives): πλήθος περιπτώσεων που ανήκουν σε μια κλάση C και δεν ταξινομήθηκαν σ' αυτή .
- ✓ FP (false positives): πλήθος περιπτώσεων που δεν ανήκουν σε μια κλάση C και ταξινομήθηκαν σ' αυτή .
- ✓ TN (true negatives): πλήθος περιπτώσεων που δεν ανήκουν σε μια κλάση C και δεν ταξινομήθηκαν σ' αυτή .

Αξιολόγηση Αλγορίθμων-Μοντέλων (2)

Μετρικές (περιπτώσεις εξόδων πολλαπλών κλάσεων)

- ✓ **Recall** = $TP/(TP+FN)$ (Ανάκληση)
- ✓ **Precision** = $TP/(TP+FP)$ (Ακρίβεια)
- ✓ **F_measure** = $(2*precision*recall)/(precision+recall)$

(ή **F1_score**)

(Τιμές στο διάστημα $[0, 1]$)

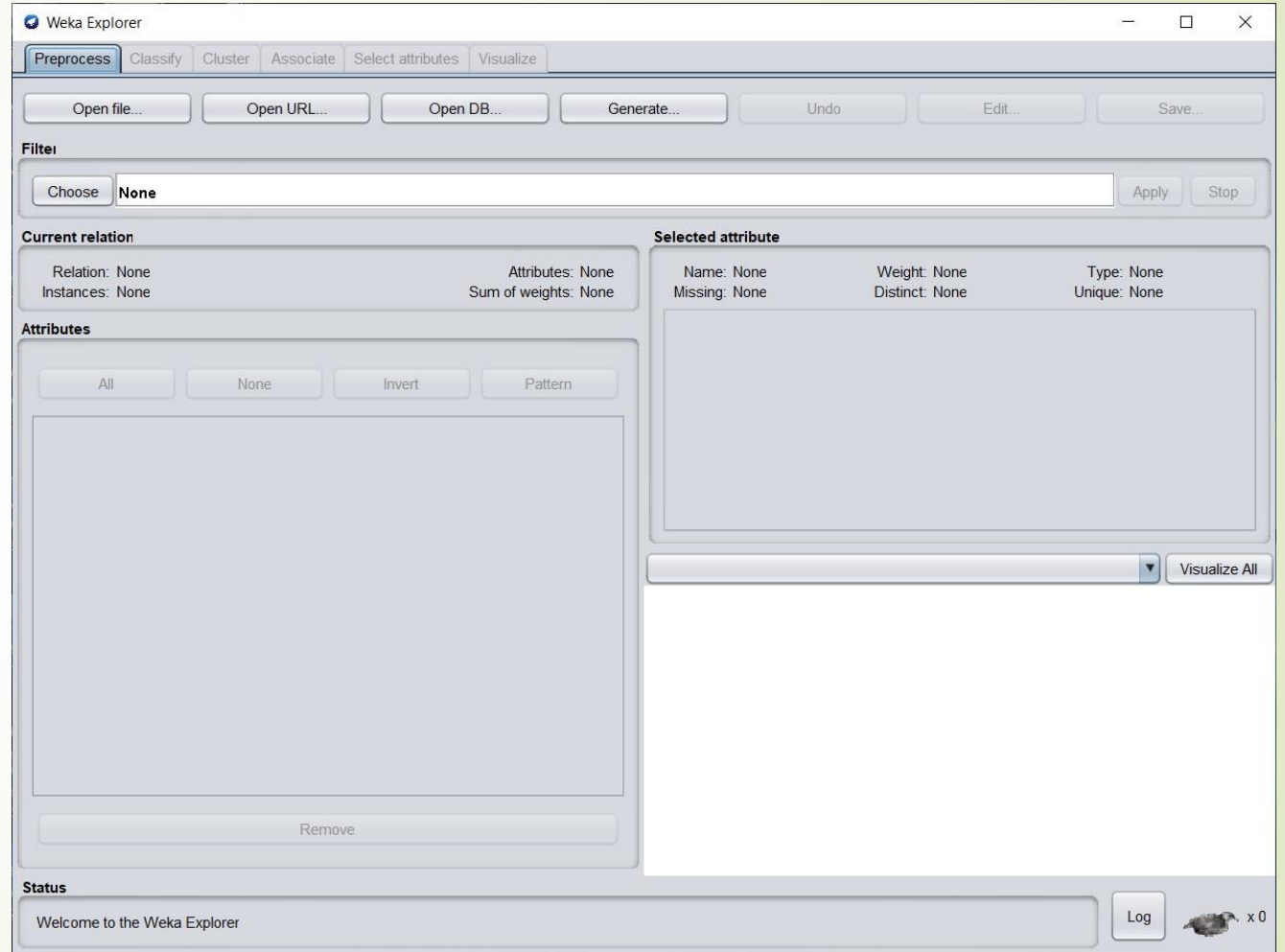
Περιγραμμά Παρουσίασης

- Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)
- Βασικές Προσεγγίσεις και Συστήματα ΤΝ
- Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι
- Επιβλεπόμενη Μάθηση
- **Εργαλείο WEKA**
- Νευρωνικά Δίκτυα
- Βαθιά Μάθηση-Εφαρμογές

Εργαλείο WEKA

- Υπάρχουν διάφορα εργαλεία υλοποίησης και χρήσης αλγορίθμων μηχανικής μάθησης (WEKA, KNIME, RapidMiner κλπ), τα περισσότερα των οποίων είναι ελεύθερα.
- Το WEKA (Waikato Environment for Knowledge Analysis) είναι ένα αρκετά εύχρηστο εργαλείο με πολλές δυνατότητες. Είναι δημιουργία του University of Waikato, New Zealand (<https://www.cs.waikato.ac.nz/ml/weka/>) και είναι ελεύθερο λογισμικό. Μπορείτε να το κατεβάσετε από δω: <https://sourceforge.net/projects/weka/>
- Έχει υλοποιημένους πάρα πολλούς αλγορίθμους MM που είναι έτοιμοι για χρήση. Κάθε αλγόριθμος έχει κάποιες «υπερπαραμέτρους», ο καθορισμός των οποίων παίζει ρόλο στο αποτέλεσμα.
- Επίσης, διαθέτει υλοποιημένες και μεθόδους προεπεξεργασίας του ΣΔ.
- Το ΣΔ πρέπει να διαμορφωθεί με ένα συγκεκριμένο τρόπο (αρχεία τύπου .arff) για να μπορεί να χρησιμοποιηθεί από τους υλοποιημένους αλγορίθμους.

Εργαλείο WEKA



Σύνολο Δεδομένων-WEKA

```
@relation weather.symbolic
```

```
@attribute outlook {sunny, overcast, rainy}
```

```
@attribute temperature {hot, mild, cool}
```

```
@attribute humidity {high, normal}
```

```
@attribute windy {strong, weak}
```

```
@attribute playTennis {yes, no}
```

```
@data
```

```
sunny, hot, high, weak, no
```

```
sunny, hot, high, strong, no
```

```
overcast, hot, high, weak, yes
```

```
rainy, mild, high, weak, yes
```

```
rainy, cool, normal, weak, yes
```

```
rainy, cool, normal, strong, no
```

```
overcast, cool, normal, strong, yes
```

```
sunny, mild, high, weak, no
```

```
sunny, cool, normal, weak, yes
```

```
rainy, mild, normal, weak, yes
```

```
sunny, mild, normal, strong, yes
```

```
overcast, mild, high, strong, yes
```

```
overcast, hot, normal, weak, yes
```

```
rainy, mild, high, strong, no
```

Αποτελέσματα-WEKA

```
=== Summary ===
```

```
Correctly Classified Instances      9      64.2857 %
Incorrectly Classified Instances     5      35.7143 %
Kappa statistic                     0.3137
Mean absolute error                  0.369
Root mean squared error              0.5137
Relative absolute error              78.7529 %
Root relative squared error         106.2217 %
Total Number of Instances           14
```

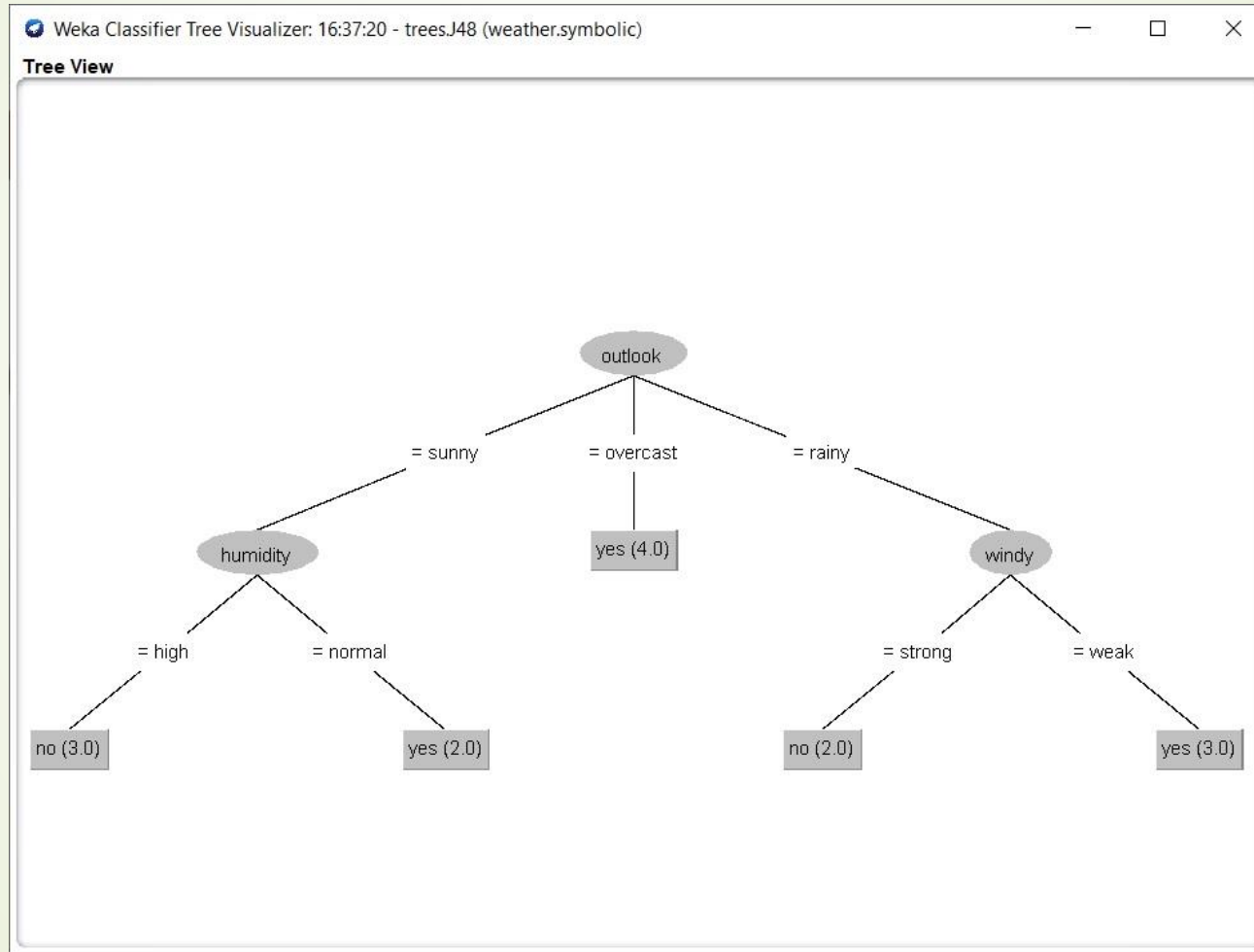
```
=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
	0,556	0,200	0,833	0,556	0,667	0,344	0,667	0,742
	0,800	0,444	0,500	0,800	0,615	0,344	0,667	0,471
Weighted Avg.	0,643	0,287	0,714	0,643	0,648	0,344	0,667	0,645

```
=== Confusion Matrix ===
```

```
a b  <-- classified as
5 4  | a = yes
1 4  | b = no
```

Δέντρο-WEKA

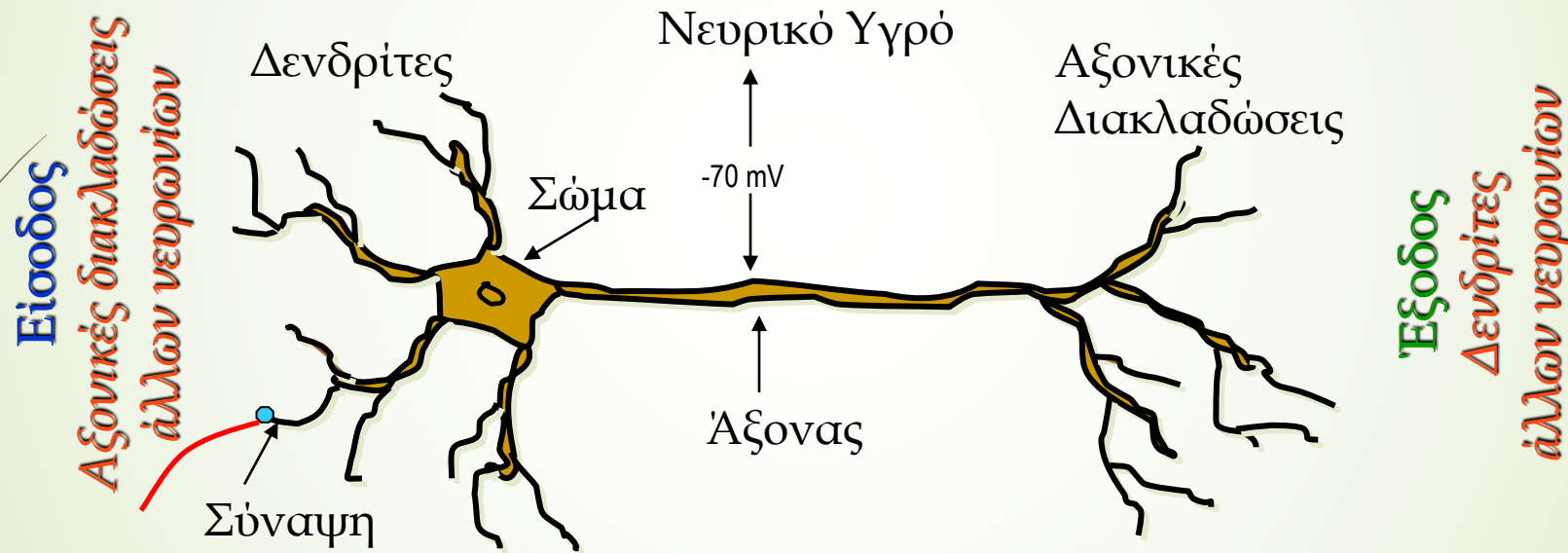


Περιγραμμά Παρουσίασης

- Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)
- Βασικές Προσεγγίσεις και Συστήματα ΤΝ
- Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι
- Επιβλεπόμενη Μάθηση
- **Νευρωνικά Δίκτυα**
- Βαθιά Μάθηση-Εφαρμογές ΜΜ

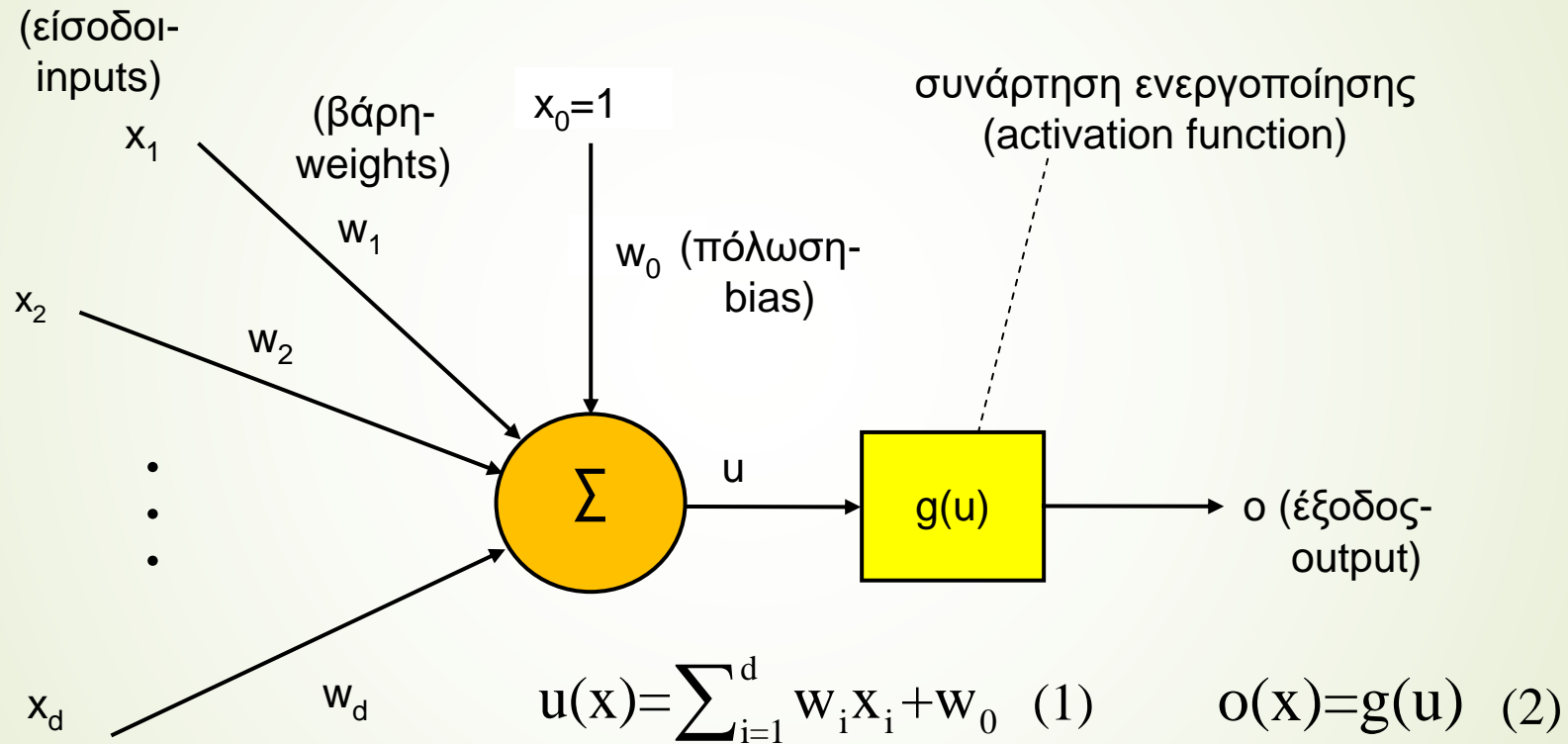
Νευρωνικά Δίκτυα-Φυσικό Μοντέλο

Προήλθαν από την προσπάθεια να προσομοιώσουμε τη λειτουργία του ανθρώπινου εγκεφάλου, ο οποίος αποτελείται από δισεκατομμύρια νευρώνες:

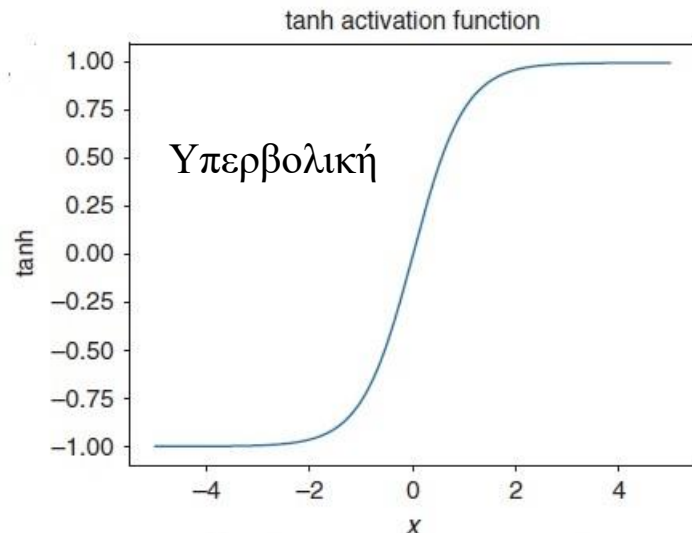
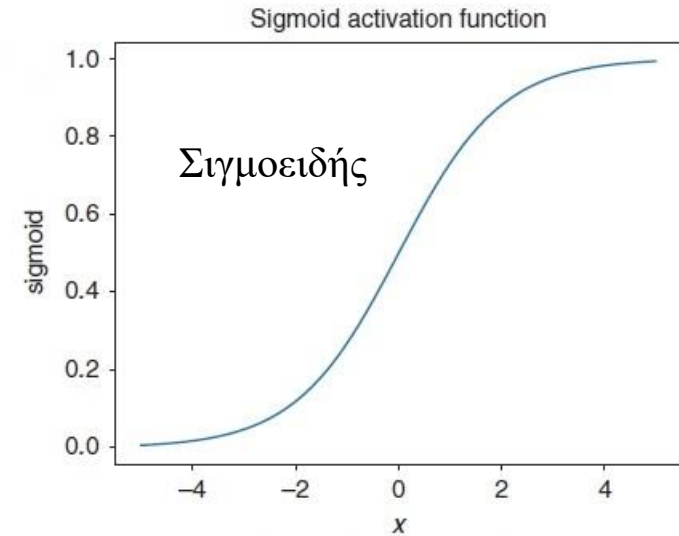
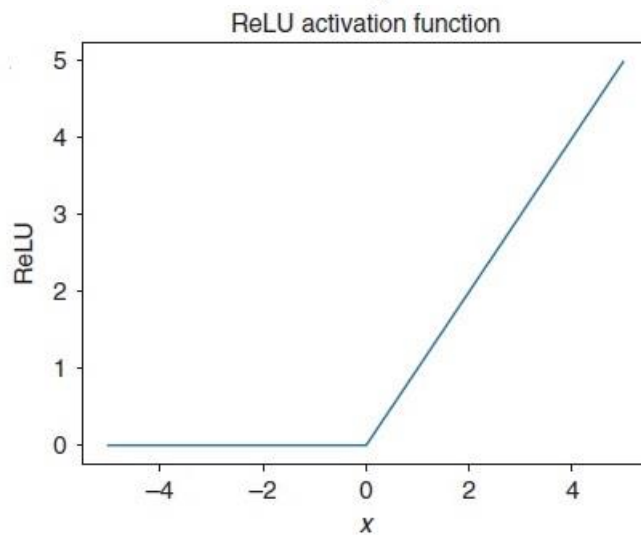
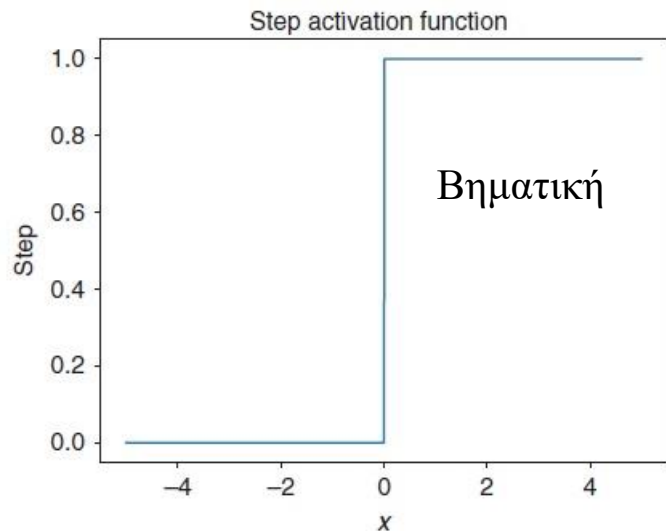


Παρατήρηση: Το σήμα που μεταδίδεται μέσω των συνάψεων στους δενδρίτες επηρεάζεται από την ισχύ της σύναψης, που ονομάζεται **συναπτικό δυναμικό**.

Νευρωνικά Δίκτυα-Υπολογιστικό Μοντέλο Νευρώνα



Νευρωνικά Δίκτυα-Συναρτήσεις Ενεργοποίησης



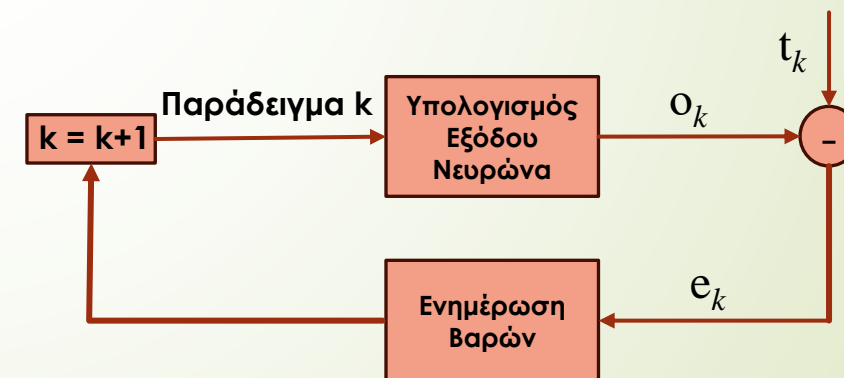
Νευρωνικά Δίκτυα-Εκπαίδευση Νευρώνα Perceptron

Στόχος της εκπαίδευσης ενός νευρώνα είναι ο **προσδιορισμός των βαρών** του, ώστε η λειτουργία του να αναπαριστάνει με τον καλύτερο τρόπο το (κρυμμένο) μοντέλο (π.χ. μια συνάρτηση) των δεδομένων του ΣΔ.

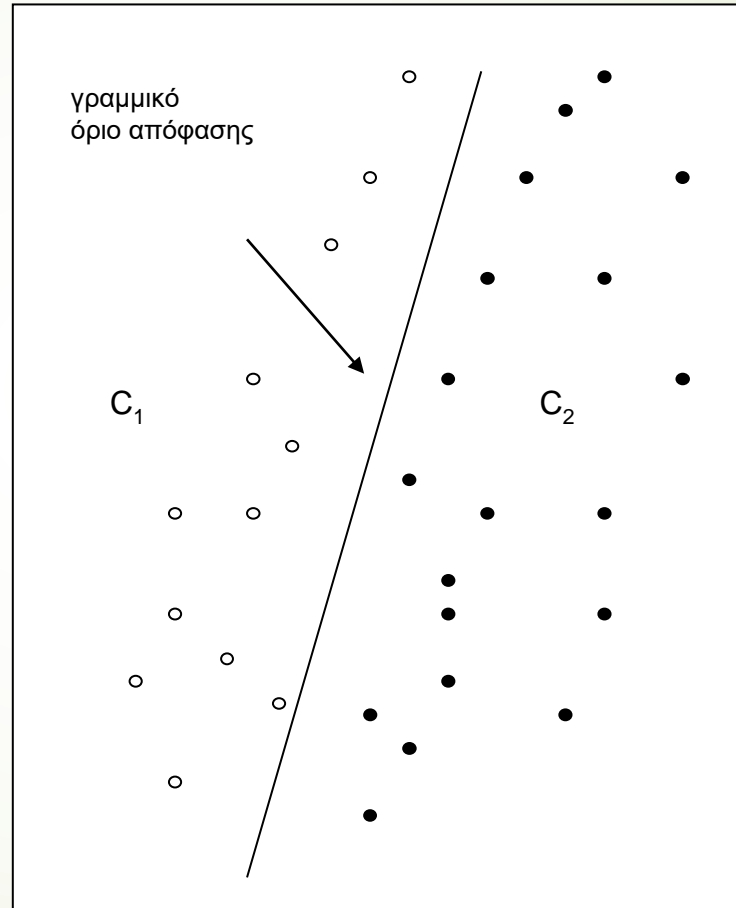
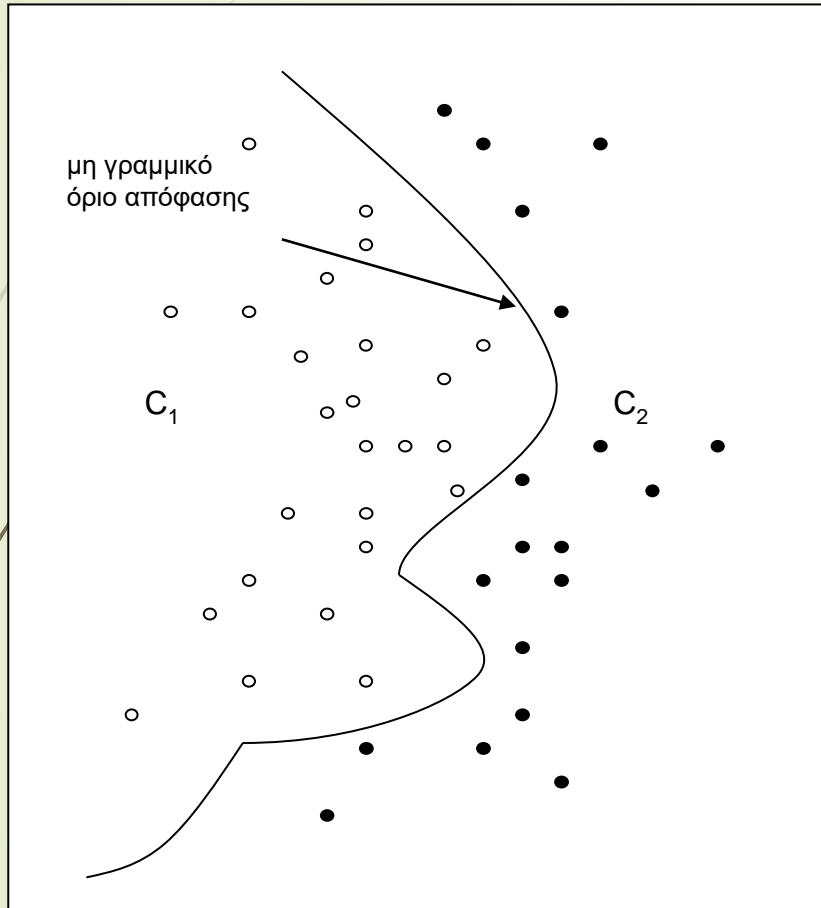
1. $k = 1$
2. Βάζουμε μηδενικές τιμές βαρών και ως είσοδο το παράδειγμα k του ΣΔ
3. Υπολογίζεται η έξοδος του νευρώνα με βάση τις εξισώσεις (1) και (2) και τη βηματική συνάρτηση ως συνάρτηση ενεργοποίησης
4. Υπολογισμός σφάλματος εξόδου, $e_k = t_k - o_k$, όπου t_k η επιθυμητή έξοδος
5. Ενημέρωση των βαρών του νευρώνα: $w_k(new) = w_k + \eta * e_k * x_k$ (η : ρυθμός μάθησης)
6. $k = k + 1$, μετάβαση στο βήμα 2.

Τερματισμός: αμετάβλητα βάρη
(αμελητέες μεταβολές)

Δεν συμβαίνει πάντα
(περίπτωση μη γραμμικού ΣΔ)



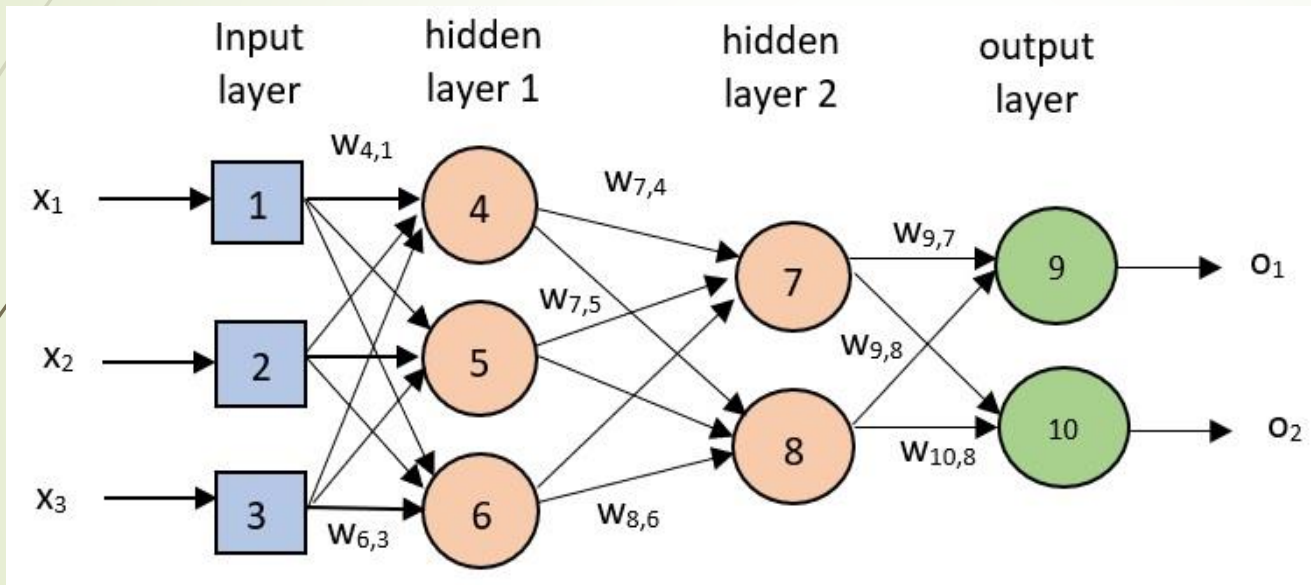
Γραμμική Διαχωριστικότητα ΣΔ



Ένας νευρώνας
δεν μπορεί να
μοντελοποιήσει
μη διαχωρίσιμα
ΣΔ.

Πολυεπίπεδα Perceptron (Multi-Layer Perceptron-MLP)

Δίκτυο πρόσθιας τροφοδότησης (Feed forward network)



1. Μη υπολογιστικό επίπεδο εισόδου.
2. Ένα ή περισσότερα κρυφά επίπεδα.
3. Πλήρης διασύνδεση επιπέδων.
4. Μη γραμμικές συναρτήσεις ενεργοποίησης στα κρυφά επίπεδα.
5. Γραμμική ή Λογιστική συνάρτηση ενεργοποίησης στην έξοδο.
6. Παριστάνει μια συνάρτηση μεταξύ εισόδου και εξόδου.

Νευρωνικά Δίκτυα-Εκπαίδευση

- ✓ Η εκπαίδευση είναι πιο πολύπλοκη από αυτή του απλού νευρώνα και συνίσταται και δω στον **προσδιορισμό των βαρών**, που είναι αρκετά περισσότερα τώρα, ώστε να δημιουργήσουμε ένα μοντέλο που να ταξινομεί σωστά τα παραδείγματα του ΣΕΚ. Τα βάρη ενός ΝΔ καλούνται **παράμετροι**.
- ✓ Η διαδικασία είναι παρόμοια με αυτή του νευρώνα. Περνάμε τα παραδείγματα του ΣΕΚ ένα-ένα και διορθώνουμε τα βάρη με βάση τα αποτελέσματα στην έξοδο. Το πέρασμα όλων των παραδειγμάτων του ΣΕΚ μια φορά ονομάζεται μια **εποχή (epoch)**. Χρειάζονται πολλές εποχές για να φτάσουμε στις τελικές τιμές των παραμέτρων.
- ✓ Πριν ξεκινήσουμε την εκπαίδευση πρέπει να καθορίσουμε κάποια στοιχεία που λέγονται **υπερπαράμετροι**, όπως ο αριθμός κρυφών επιπέδων, ο αριθμός νευρώνων ανά επίπεδο, ο ρυθμός μάθησης, ο αριθμός των εποχών.

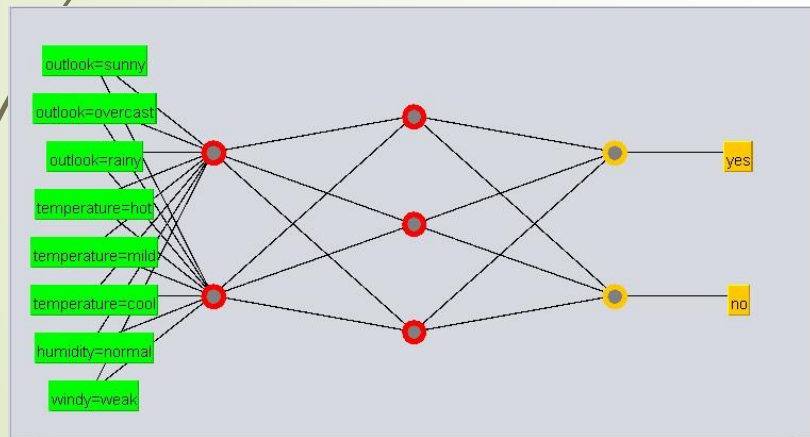
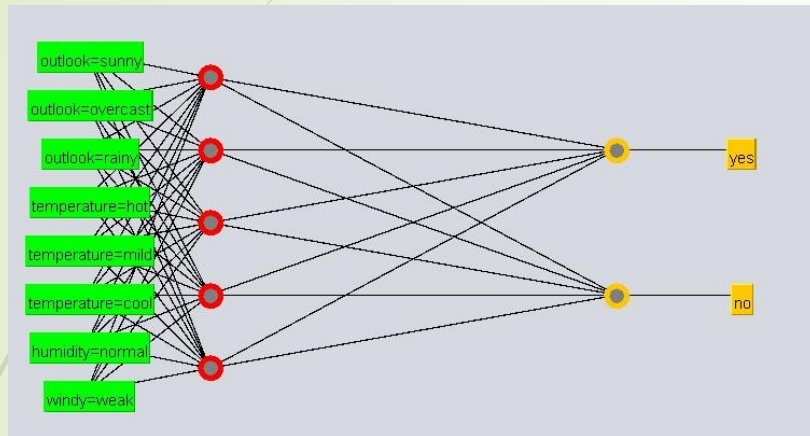
Νευρωνικά Δίκτυα-Εκπαίδευση

- ✓ Χρησιμοποιούνται πιο πολύπλοκοι **αλγόριθμοι εκπαίδευσης**:
 - Οπισθοδιάδοσης σφάλματος (Error back propagation)
 - Βαθμιαίας καθόδου (gradient descent)
- ✓ Στηρίζονται στην **ελαχιστοποίηση** μιας συνάρτησης του σφάλματος μεταξύ πραγματικής και επιθυμητής εξόδου, που αναφέρεται ως **συνάρτηση κόστους**, και έχει διάφορες εκδοχές:
 - Μέσο τετραγωνικό σφάλμα (Mean Squared Error: MSE)
 - Μέσο απόλυτο σφάλμα (Mean Absolute Error: MAE)

Νευρωνικά Δίκτυα-Γενίκευση

- ✓ Απώτερος στόχος της εκπαίδευσης είναι η **γενίκευση**.
- ✓ Η γενίκευση αναφέρεται στην ικανότητα ενός ΝΔ να μπορεί να ταξινομεί με επιτυχία παραδείγματα/περιπτώσεις για τις οποίες δεν έχει εκπαιδευτεί.
- ✓ Η εκτίμηση της ικανότητας γενίκευσης ενός ΝΔ γίνεται με εμπειρικό τρόπο: με τη χρήση ΣΕΚ και ΣΕΛ και τις μετρικές που προαναφέραμε.
- ✓ Ο χωρισμός σε ΣΕΚ και ΣΕΛ του ΣΔ και ο τρόπος χρήσης τους παίζει ρόλο στο αποτέλεσμα.
- ✓ Μια «αντικειμενική» μέθοδος είναι η **διασταυρωμένη επικύρωση k-τμημάτων** (k-fold cross validation).

Νευρωνικά Δίκτυα-WEKA



=== Summary ===

Correctly Classified Instances	10	71.4286 %
Incorrectly Classified Instances	4	28.5714 %
Kappa statistic	0.4286	
Mean absolute error	0.2871	
Root mean squared error	0.4788	
Relative absolute error	61.2579 %	
Root relative squared error	99.0052 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
yes	0,667	0,200	0,857	0,667	0,750	0,447	0,800	0,911
no	0,800	0,333	0,571	0,800	0,667	0,447	0,800	0,693
Weighted Avg.	0,714	0,248	0,755	0,714	0,720	0,447	0,800	0,833

=== Confusion Matrix ===

```
a b <-- classified as
6 3 | a = yes
1 4 | b = no
```

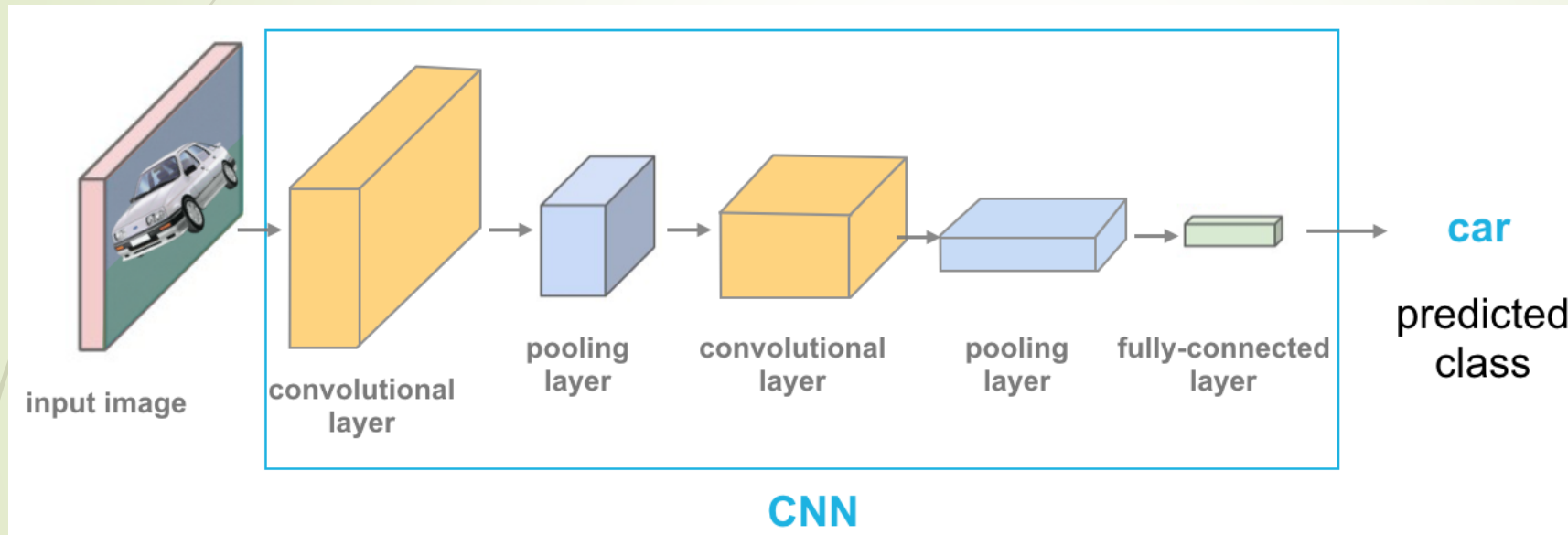
Περιγραμμά Παρουσίασης

- Τι είναι Τεχνητή Νοημοσύνη (ΤΝ)
- Βασικές Προσεγγίσεις και Συστήματα ΤΝ
- Τι είναι Μηχανική Μάθηση (ΜΜ)-Βασικές Μέθοδοι
- Επιβλεπόμενη Μάθηση
- Νευρωνικά Δίκτυα
- **Βαθιά Μάθηση-Εφαρμογές ΜΜ**

Βαθιά Μάθηση (Deep Learning)

- ✓ Είναι μια πρόσφατη υποπεριοχή της μηχανικής μάθησης.
- ✓ Αναφέρεται σε πολύπολκα ΝΔ, δηλ. δίκτυα με πάνω από δύο κρυφά επίπεδα.
- ✓ Η ύπαρξη πολλών κρυφών επιπέδων συντελεί στη δυνατότητα αναπαράστασης πιο πολύπλοκων ΣΔ.
- ✓ Απαιτούν μεγάλο όγκο δεδομένων και χρόνο εκπαίδευσης.
- ✓ Κάνουν «βαθύτερη», δηλ. πιο λεπτομερή, αναπαράσταση των χαρακτηριστικών των δεδομένων.
- ✓ Πήρε μεγάλη δημοσιότητα με την εμφάνιση του λεγόμενου **Συνελικτικού Νευρωνικού Δικτύου (Convolutional Neural Network-CNN)**, που επέτυχε εκπληκτικά αποτελέσματα στην **αναγνώριση εικόνων**.
- ✓ Έκτοτε εμφανίστηκαν πολλές παραλλαγές ΝΔ βαθιάς μάθησης.

Συνελικτικό Νευρωνικό Δίκτυο (CNN)

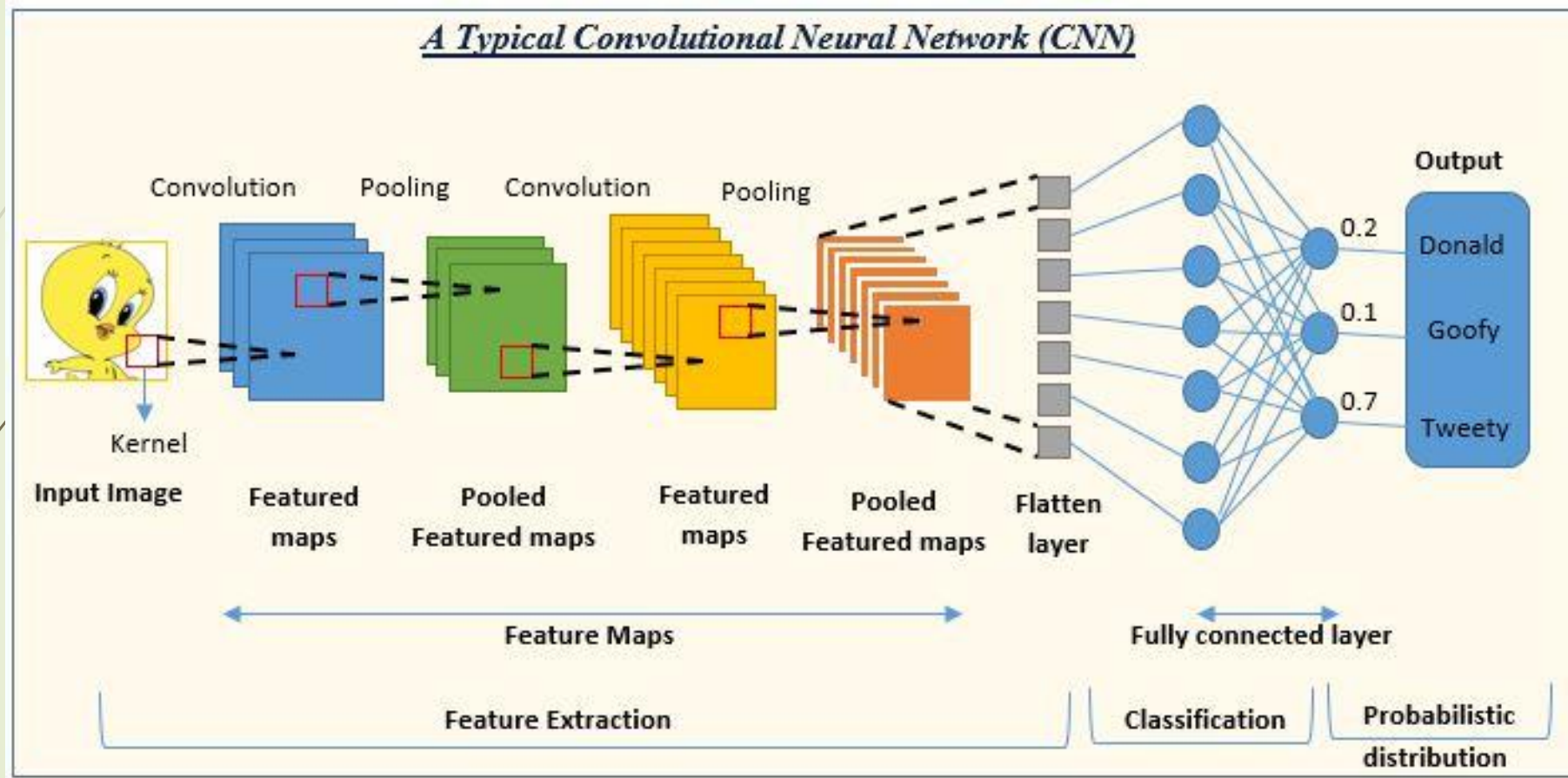


Κατηγορίες επιπέδων

- Input
- Convolutional
- Pooling
- Fully – Connected

Πηγή: https://cezannec.github.io/Convolutional_Neural_Networks/

Συνελικτικό Νευρωνικό Δίκτυο (CNN)

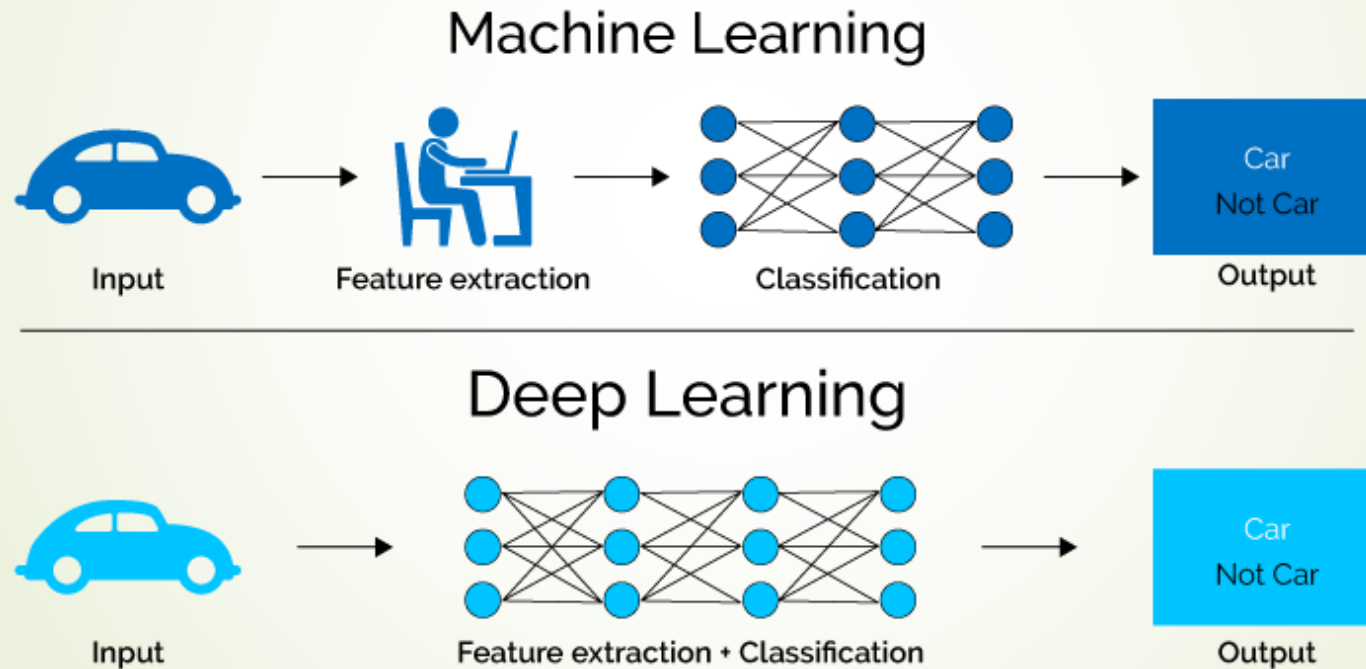


Πηγή: <https://www.analyticsvidhya.com/blog/2022/01/convolutional-neural-network-an-overview/>

Συνελικτικό Νευρωνικό Δίκτυο (CNN)

Το μεγάλο πλεονέκτημα: Δεν απαιτείται εξαγωγή χαρακτηριστικών

Το μεγάλο μειονέκτημα: Απαιτείται μεγάλος όγκος δεδομένων



Πηγή: <https://towardsdatascience.com/cnn-application-on-structured-data-automated-feature-extraction-8f2cd28d9a7e>

Παραλλαγές CNN

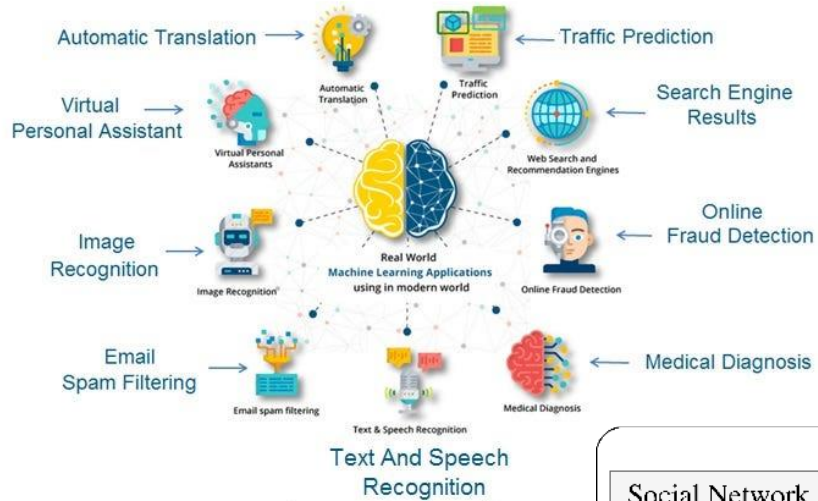
- AlexNet
- ZFNet
- VGGNet
- GoogLeNet
- ResNet

Αρχιτεκτονικές ΝΔ Βαθιάς Μάθησης

- Convolutional Neural Networks (CNNs)
- Recurrent Neural Networks (RNNs)
- Long Short Term Memory Networks (LSTMs)
- Generative Adversarial Networks (GANs)
- Deep Belief Networks (DBNs)
- Autoencoders
- Transformers (π.χ. BERT)

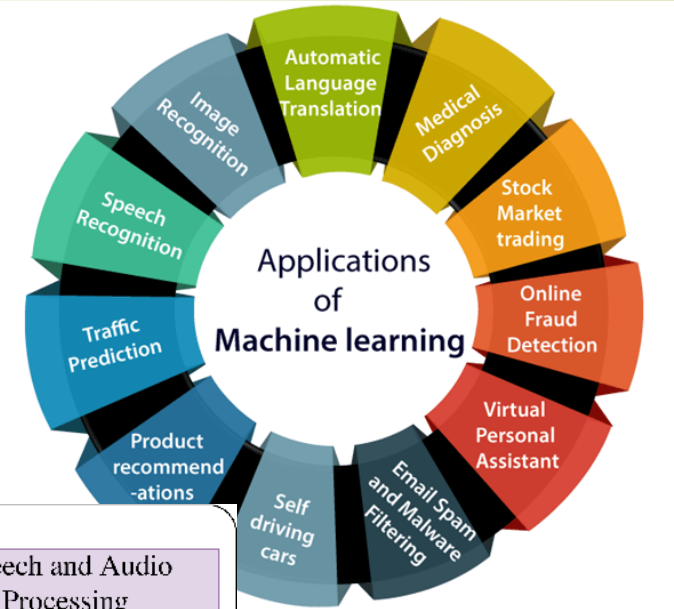
Εφαρμογές Μηχανικής Μάθησης

Real World Applications Of Machine Learning



Σημαντικές Εφαρμογές DL

- Επεξεργασία Φυσικής Γλώσσας (αναγνώριση και δημιουργία κειμένου και φωνής, συναισθηματική ανάλυση κειμένου και φωνής)
- Επεξεργασία και αναγνώριση εικόνων (προσώπων, αντικειμένων)
- Αναγνώριση κινούμενων αντικειμένων



Deep Learning-based Applications

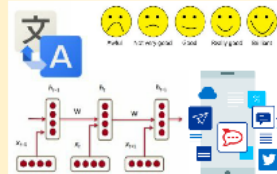
Social Network Analysis



Autonomous Driving



Natural Language Processing



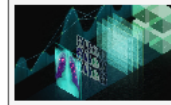
Sentiment Classification
Entity Extraction
Translation

Visual Data Processing



Computer Vision
Multimedia Data Analysis

Biomedicine



Disaster



Speech and Audio Processing



Speech Enhancement
Speech Recognition

Information
Retrieval



Ευχαριστώ για την υπομονή σας!