



ΤΕΧΝΟΛΟΓΙΚΟ
ΕΚΠΑΙΔΕΥΤΙΚΟ
ΙΔΡΥΜΑ
ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ

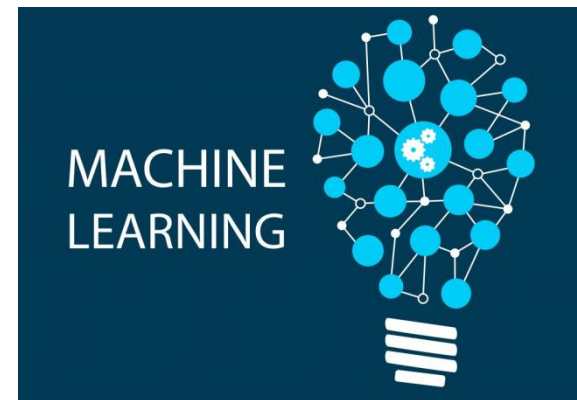
ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ ΛΗΨΗΣ ΑΠΟΦΑΣΗΣ ΣΤΙΣ ΕΠΙΣΤΗΜΕΣ ΥΓΕΙΑΣ

ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ
(Machine Learning)

Δρ. Ισίδωρος Περίκος

Μηχανική Μάθηση

- Η μηχανική μάθηση είναι μια περιοχή της τεχνητής νοημοσύνης η οποία αφορά αλγορίθμους και μεθόδους που επιτρέπουν στους Η/Υ να «μαθαίνουν»
- Σκοπός = η εξόρυξη δεδομένων



Μηχανική Μάθηση

- Τι είναι Μηχανική Μάθηση?

Learn from experience



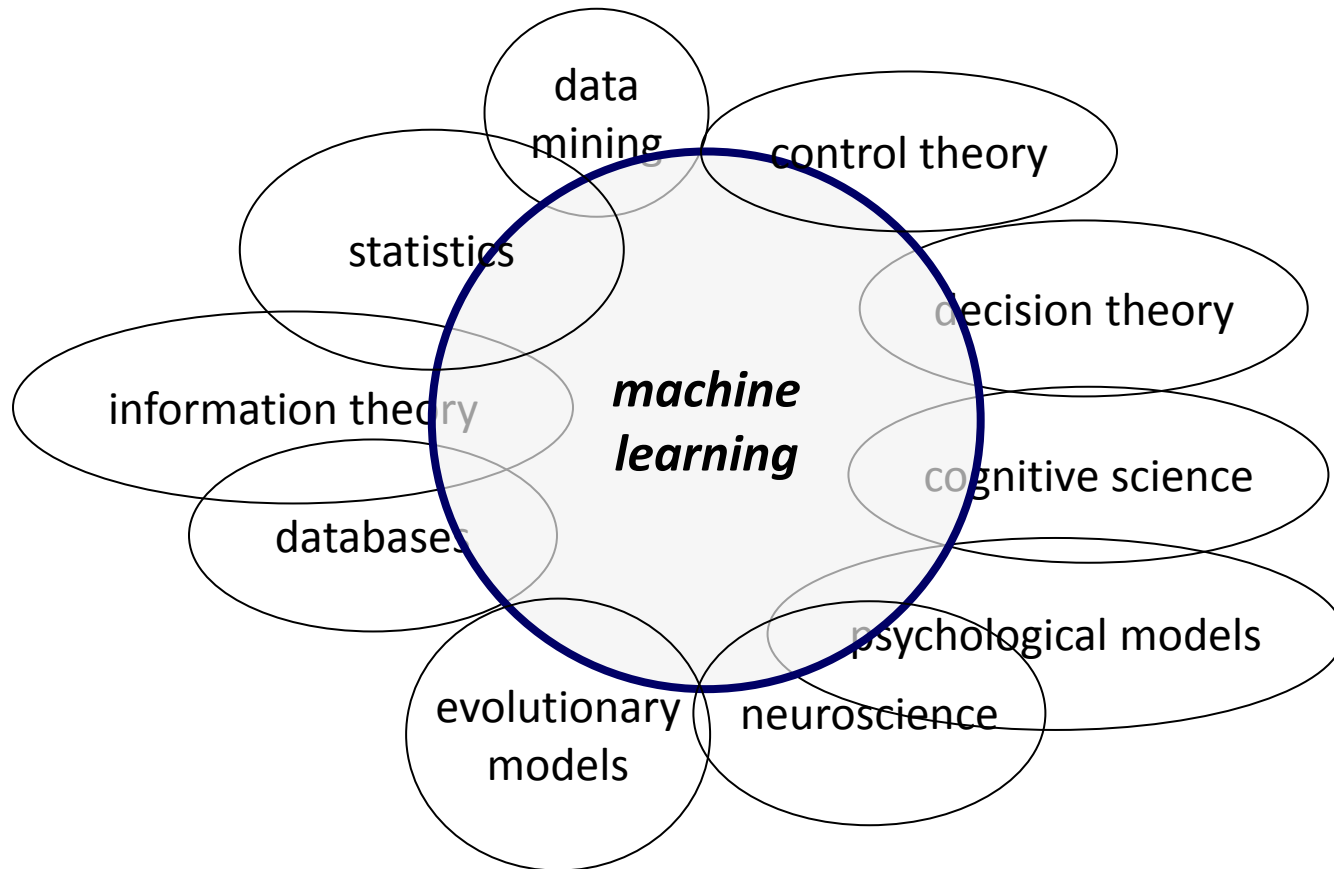
Learn from experience



Follow instructions



Σχετικά Πεδία της Μηχανικής Μάθησης



Εφαρμογές Μηχανικής Μάθησης

- Τεχνητές της μηχανικής μάθησης βρίσκουν εφαρμογή σε πολλά πεδία όπως:
 - Επεξεργασία Φυσικής γλώσσας
 - Αναγνώριση ομιλίας
 - Ανάλυση εικόνας
 - Ανάλυση ιατρικών δεδομένων
 - Ρομποτική
 - ...

Είδη μηχανικής μάθησης

- ❑ **Μάθηση με επίβλεψη (supervised learning) ή μάθηση με παραδείγματα (learning from examples)**
 - Το σύστημα καλείται να «μάθει» μια έννοια ή συνάρτηση από ένα σύνολο δεδομένων, η οποία αποτελεί περιγραφή ενός μοντέλου.
- ❑ **Μάθηση χωρίς επίβλεψη (unsupervised learning) ή μάθηση απο παρατήρηση (learning from observation)**
 - Το σύστημα καλείται μόνο του να ανακαλύψει συσχετίσεις ή ομάδες σε ένα σύνολο δεδομένων, δημιουργώντας πρότυπα χωρίς να είναι γνωστό αν υπάρχουν, πόσα είναι και ποια.
- **Ενισχυτική μάθηση (reinforcement learning)**
 - Μάθηση μέσω ενίσχυσης (επιβράβευσης)

Μάθηση με Επίβλεψη

- Μάθηση με επίβλεψη το σύστημα «μαθαίνει» επαγωγικά μια συνάρτηση που ονομάζεται **συνάρτηση στόχος (target function)**
- Η συνάρτηση στόχος χρησιμοποιείται για την πρόβλεψη της τιμής μιας μεταβλητής που ονομάζεται εξαρτημένη μεταβλητή ή μεταβλητή εξόδου,
- Βάσει των τιμών ενός συνόλου μεταβλητών που ονομάζονται ανεξάρτητες μεταβλητές ή μεταβλητές εισόδου.

Μπορούμε να δούμε τη διαδικασία μάθησης γενικά ως τη διαδικασία μάθησης της αναπαράστασης μιας συνάρτησης

Επαγωγική Μάθηση

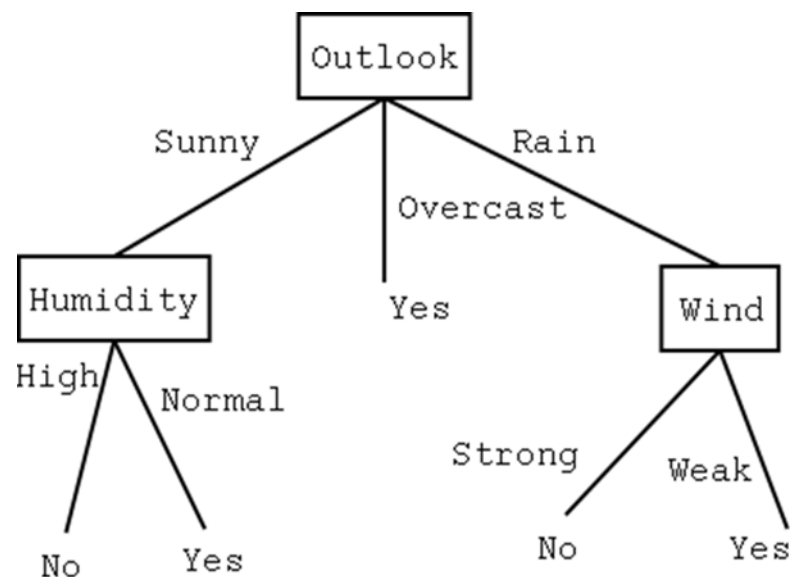
- **Επαγωγική Μάθηση** (inductive learning) στηρίζεται στην υπόθεση της επαγωγικής μάθησης (inductive learning hypothesis) δηλ. είναι η μάθηση μιας έννοιας μέσω ενός συνόλου παραδειγμάτων
- Σύμφωνα με την οποία:

Κάθε υπόθεση h που προσεγγίζει καλά τη συνάρτηση στόχο για ένα αρκετά μεγάλο σύνολο παραδειγμάτων, θα προσεγγίζει το ίδιο καλά τη συνάρτηση στόχο και για τις περιπτώσεις που δεν έχει εξετάσει

Δέντρα Απόφασης

- **Δέντρο Απόφασης** ή Δέντρο Κατηγοριοποίησης είναι ένα δέντρο με τις ακόλουθες ιδιότητες:

- Κάθε εσωτερικός κόμβος και η ρίζα ονοματίζεται με το όνομα ενός χαρακτηριστικού.
- Κάθε κλάδος ονοματίζεται με ένα κατηγορημα διάσπασης του χαρακτηριστικού που αποτελεί το όνομα του κόμβου-πατέρα.
- Κάθε φύλλο ονοματίζεται με το όνομα μιας κλάσης



Αλγόριθμος ID3

- Χρησιμοποιεί σαν κριτήριο για τον προσδιορισμό του «καλύτερου χαρακτηριστικού διάσπασης» το «κέρδος πληροφορίας» (information gain).
- Το «κέρδος πληροφορίας» μετριέται ποσοτικά με την εντροπία (entropy).
 - Η εντροπία εν γένει εκφράζει το μέγεθος της ανομοιογένειας σε ένα σύνολο δεδομένων. Π.χ. αν όλα τα δεδομένα ανήκουν σε μια κλάση, τότε δεν υπάρχει ανομοιογένεια: η εντροπία είναι μηδέν.
 - Το ζητούμενο σ' ένα ΔΑ είναι ο διαχωρισμός του συνόλου εκπαίδευσης, μ' ένα επαναληπτικό τρόπο, σε υποσύνολα μηδενικής εντροπίας.
 - Αν p η πιθανότητα να συμβεί ένα γεγονός, τότε $\log(1/p)$ παριστάνει το ποσό της τυχαιότητας με βάση την πιθανότητα.
 - Η αναμενόμενη πληροφορία με βάση την p ορίζεται: $p \log(1/p)$
 - Αν έχω δύο συμπληρωματικά γεγονότα e , e' με p , p' , τότε η αναμενόμενη πληροφορία είναι: $p \log(1/p) + p' \log(1/p')$

Αλγόριθμος ID3

1. Έστω S ένα σύνολο παραδειγμάτων.
2. Βρες την 'καλύτερη' παράμετρο/μεταβλητή a .
3. Χώρισε το S σε υποσύνολα S_i , $i=1, k$ έτσι ώστε όλα τα παραδείγματα στο S_i να έχουν την ίδια τιμή για την a . Κάθε S_i συνιστά ένα κόμβο του δέντρου.
4. Για κάθε S_i

Αν όλα τα παραδείγματα ανήκουν στην ίδια κλάση

4.1 τότε δημιούργησε ένα φύλλο με όνομα την κλάση

4.2 αλλιώς, πήγαινε στο 1 για $S = S_i$.

Αλγόριθμος ID3

- Η εύρεση της καλύτερης a στηρίζεται στο
- κέρδος πληροφορίας $G(S, a) = E(S) - E(S, a)$

όπου

- $E(S) = -p^+ \log_2(p^+) - p^- \log_2(p^-)$
- όπου $p^+ = n^+/N$ και $p^- = n^-/N$ ($N = n^+ + n^-$)
- και
- $E(S, a) = \sum P(S_i) * E(S_i) \quad i = 1, k$ ($k = \text{ο αριθμός τιμών της } a$)
- με $P(S_i) = |S_i|/|S|$

Παράδειγμα-Αλγόριθμος ID3 (1)

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	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Παράδειγμα-Αλγόριθμος ID3 (2)

- Επιλογή ρίζας

$$G(S, Outlook) = E(S) - (|S_{\text{sunny}}|/|S|) E(S_{\text{sunny}}) \\ - (|S_{\text{overcast}}|/|S|) E(S_{\text{overcast}}) \\ - (|S_{\text{rain}}|/|S|) E(S_{\text{rain}})$$

Χαρακτηριστικό-στόχος: PlayTennis:PT (yes, no)

$$p_1 = p(\text{PT=yes}) = 9/14, p_2 = p(\text{PT=no}) = 5/14$$

$$E(S) = p_1 \log(1/p_1) + p_2 \log(1/p_2) = -p_1 \log(p_1) - p_2 \log(p_2) = \\ -(9/14) \log(9/14) - (5/14) \log(5/14) = 0,283$$

Παράδειγμα-Αλγόριθμος ID3 (3)

$$E(S_{\text{sunny}}) = -(2/5) \log(2/5) - (3/5) \log(3/5) = 0,292$$

$$E(S_{\text{rain}}) = -(3/5) \log(3/5) - (2/5) \log(2/5) = 0,292$$

$$E(S_{\text{overcast}}) = -(4/4) \log(4/4) - (0/4) \log(0/4) = 0$$

$$|S| = 14, |S_{\text{sunny}}| = 5, |S_{\text{rain}}| = 5, |S_{\text{overcast}}| = 4$$

$$G(S, \text{Outlook}) = 0,074$$

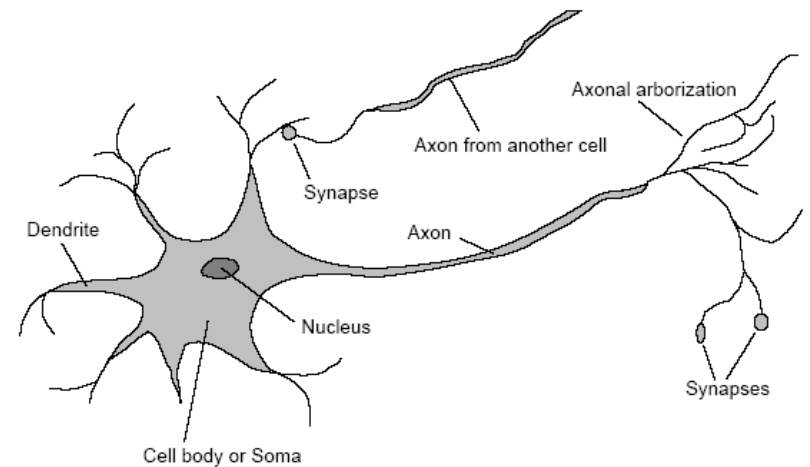
Ομοίως $G(S, \text{Humidity}) = 0,04565$, $G(S, \text{Wind}) = 0,0144$
και $G(S, \text{Temperature}) = 0,0087$

Μοντελα Μάθησης

- Στατιστικά Μοντέλα Bayes: Με βάση την στατιστική και τον νόμο του Bayes τα μοντέλα προσπαθούν να εκτιμήσουν την πιο πιθανή λύση στο πρόβλημα με βάση τα δεδομένα που έχουν διαθεσιμα.
- Τεχνητά Νευρωνικά Δίκτυα: Είναι εμπνευσμένα από την λειτουργία των νευρώνων του εγκεφάλου και χρησιμοποιούνται για την εκπαίδευση και την προσδιορισμός των εσωτερικών παραμέτρων (“συναπτικά βάρη”)

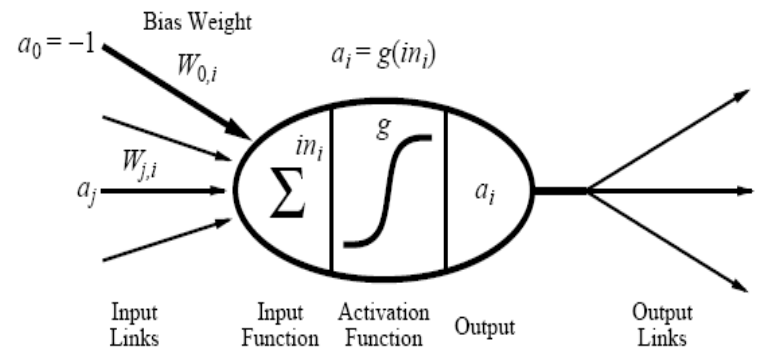
Τεχνητά Νευρωνικά Δίκτυα

- Ανθρώπινος νευρώνας ανθρώπινου εγκεφάλου



- Τεχνητός νευρώνας

$$a_i \leftarrow g(in_i) = g(\sum_j W_{j,i} a_j)$$



Συλλογιστική βασισμένη σε περιπτώσεις (Case-Based Reasoning)

- Η επίλυση προβλημάτων προκύπτει από την επιλογή και ανάκτηση των πλέον σχετικών περιπτώσεων από μια βάση δεδομένων και την προσαρμογή τους ώστε να χαρακτηριστούν κατάλληλες σε νέες καταστάσεις.

Συλλογιστική βασισμένη σε περιπτώσεις (Case-Based Reasoning)

- Η Συλλογιστική βασισμένη σε περιπτώσεις βασίζεται σε δυο αρχές σχετικές με τη φύση του κόσμου:
- Ο κόσμος είναι κανονικός: όμοια προβλήματα έχουν όμοια λύση.
- Προβλήματα που κάποιος αντιμετωπίζει τείνουν να ξανασυμβούν. Άρα μελλοντικά προβλήματα είναι πολύ πιθανόν να μοιάζουν με τρέχοντα προβλήματα.