

Μηχανική Μάθηση – Μάθημα 1

Βασικές έννοιες

Κώστας Διαμαντάρας
Τμήμα Πληροφορικής
ΤΕΙ Θεσσαλονίκης
2011

Στοιχεία επικοινωνίας

- Κώστας Διαμαντάρας
- Τηλ. 2310 013592
- Email: kdiamant@it.teithe.gr
- <http://www.it.teithe.gr/~kdiamant/>
- Ανακοινώσεις, ύλη, διαφάνειες :
<http://blackboard.teithe.gr>

Τι είναι η «Μάθηση» (Learning) ;

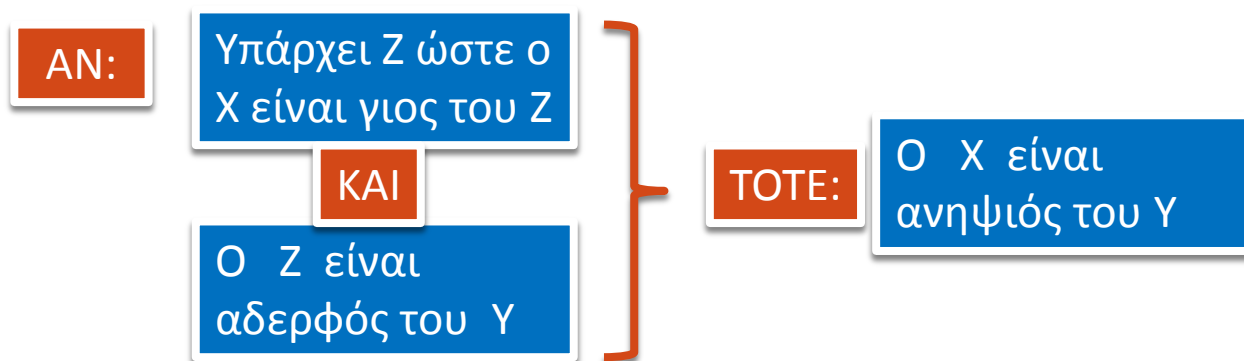
- **Μάθηση** : **χρήση δεδομένων** που προσλαμβάνει κάποιος από το περιβάλλον (πχ. ηχητικά, οπτικά σήματα, κλπ) έτσι ώστε αυτός να **βελτιώνεται σταδιακά** κατά την εκτέλεση μιας λειτουργίας.
- Η μάθηση είναι βασικό συστατικό της **νοημοσύνης**.
- Παραδείγματα λειτουργιών:
 - κατανόηση της γλώσσας
 - παραγωγή ομιλίας
 - αναγνώριση προσώπων και αντικειμένων
 - ανάπτυξη στρατηγικής σε διάφορες καταστάσεις (πχ. παιχνίδια, κοινωνική συμπεριφορά, κλπ), και πολλά άλλα...

Μηχανική Μάθηση (Machine Learning)

- **Μηχανική Μάθηση**: χρήση *δεδομένων* από έναν **αλγόριθμο** ο οποίος εκτελείται σε μια υπολογιστική μηχανή έτσι ώστε αυτή να **βελτιώνεται σταδιακά** κατά την εκτέλεση μιας λειτουργίας.
- Οι λειτουργίες είναι αντίστοιχες αυτών που εμπίπτουν στην ανθρώπινη νοημοσύνη, όπως
 - Η μηχανική κατανόηση της γλώσσας και η παραγωγή ομιλίας (natural language processing /understanding)
 - Η μηχανική αναγνώριση προτύπων (pattern recognition)
 - Η ανάπτυξη στρατηγικής σε διάφορες καταστάσεις (πχ. παιχνίδια, πλοήγηση στο χώρο, κλπ), και πολλά άλλα...

Μηχανική μάθηση και Τεχνητή Νοημοσύνη

- Η Μάθηση θεωρείται βασικό συστατικό της Ανθρώπινης Νοημοσύνης. Η Μηχανική Μάθηση θεωρείται βασικό συστατικό της **Τεχνητής Νοημοσύνης (ΤΝ)**.
- Η **παραδοσιακή ΤΝ** δεν βασίζεται στη μάθηση αλλά σε κανόνες που παρέχονται από κάποιον «ειδικό». Για παράδειγμα, η λειτουργία αναγνώρισης της σχέσης «ο X είναι ανηψιός του Y » βασίζεται σε κανόνες του τύπου:



Παραδοσιακή Τεχνητή Νοημοσύνη (TN)

- **Πλεονεκτήματα:** Η παραδοσιακή TN δίνει εξαιρετικά αποτελέσματα σε καλώς ορισμένα προβλήματα με σταθερούς κανόνες (πχ. το σκάκι, η διάγνωση βλαβών σε βιομηχανικά εξαρτήματα ή μηχανές, η διάγνωση ασθενειών, κλπ) κι εφόσον είναι δυνατή η χρήση ειδικών (*experts*) για την εξαγωγή των απαραίτητων κανόνων.
- Βασικά εργαλεία της παραδοσιακής Τεχνητής Νοημοσύνης (TN) είναι οι γλώσσες λογικού προγραμματισμού όπως `Prolog`, `LISP`, κα.

Παραδοσιακή ΤΝ

- Τα βασικά **μειονεκτήματα** της παραδοσιακής προσέγγισης είναι:
 - Η έλλειψη ευελιξίας
 - Η αδυναμία προσαρμογής σε νέα δεδομένα
 - Η δυσκολία εξαγωγής του συνόλου των κανόνων. Σε πολλές περιπτώσεις, είτε δεν είναι εύκολη η εύρεση του κατάλληλου «ειδικού» που θα παρέχει τους κανόνες, είτε οι κανόνες είναι πολύπλοκοι, είτε είναι εξαιρετικά δύσκολο να διατυπωθούν.

Τι προσφέρει η μάθηση;

- Το βασικά πλεονεκτήματα της μάθησης είναι η **ευελιξία** και η **προσαρμοστικότητα**.
- Ας δούμε τη διαφορά σε ένα παράδειγμα ευφυούς λειτουργίας **αναγνώρισης προτύπων**.
- Έστω ότι θέλουμε να ξεχωρίζουμε ζώα μεταξύ τους και συγκεκριμένα θέλουμε να διακρίνουμε ενήλικα άλογα από ενήλικα πρόβατα.
- Προφανώς η είσοδος στο αλγόριθμο αναγνώρισης δεν θα είναι το ίδιο το ζώο αλλά κάποια **χαρακτηριστικά** του που μπορούν να ποσοτικοποιηθούν...

Παράδειγμα: Αναγνώριση ζώων

- ...ένα χρήσιμο χαρακτηριστικό για την διάκριση μεταξύ αλόγων και προβάτων είναι, για παράδειγμα, το ύψος, x , του ζώου.
- Στην παραδοσιακή ΤΝ η αναγνώριση του ζώου θα γινόταν με κανόνες της μορφής



- Σ' αυτό το πρόβλημα αρκεί μόνο το χαρακτηριστικό x για επιτυχή αναγνώριση. Σε πολλά προβλήματα, ωστόσο, απαιτείται η χρήση ενός διανύσματος χαρακτηριστικών που καλείται **πρότυπο**.

Παράδειγμα: Αναγνώριση ζώων

- Τι θα γίνει όμως αν κληθούμε να διακρίνουμε άλογα από ελέφαντες;
- Αν συνεχίσουμε να χρησιμοποιούμε το ύψος ως χαρακτηριστικό τότε θα πρέπει να δημιουργήσουμε νέους κανόνες, πχ.

ΑΝ:	$x \geq 3$ μέτρα	ΤΟΤΕ:	Το ζώο είναι ελέφαντας
ΑΝ:	$x < 3$ μέτρα	ΤΟΤΕ:	Το ζώο είναι άλογο

!! Κάθε φορά που θα αντιμετωπίζουμε ένα νέο πρόβλημα αναγνώρισης θα πρέπει να δημιουργούμε ένα νέο σετ κανόνων, άσχετων με το προηγούμενο πρόβλημα...

Παράδειγμα:

Αναγνώριση ζώων με μάθηση

- Οι κανόνες είναι κατάλληλοι για ένα συγκεκριμένο πρόβλημα αλλά δεν προσαρμόζονται σε άλλα προβλήματα
- Ένας αλγόριθμος αναγνώρισης που χρησιμοποιεί **μάθηση** δέχεται στην είσοδό του πολλά **πρότυπα** από κάθε κλάση ζώου. Π.χ. αν κληθεί να διακρίνει άλογα από πρόβατα τότε θα «δει» πολλά πρότυπα αλόγων και προβάτων. Στην συγκεκριμένη περίπτωση κάθε πρότυπο, \mathbf{p} , είναι ένα διάνυσμα με ένα μόνο στοιχείο: το ύψος x του ζώου:

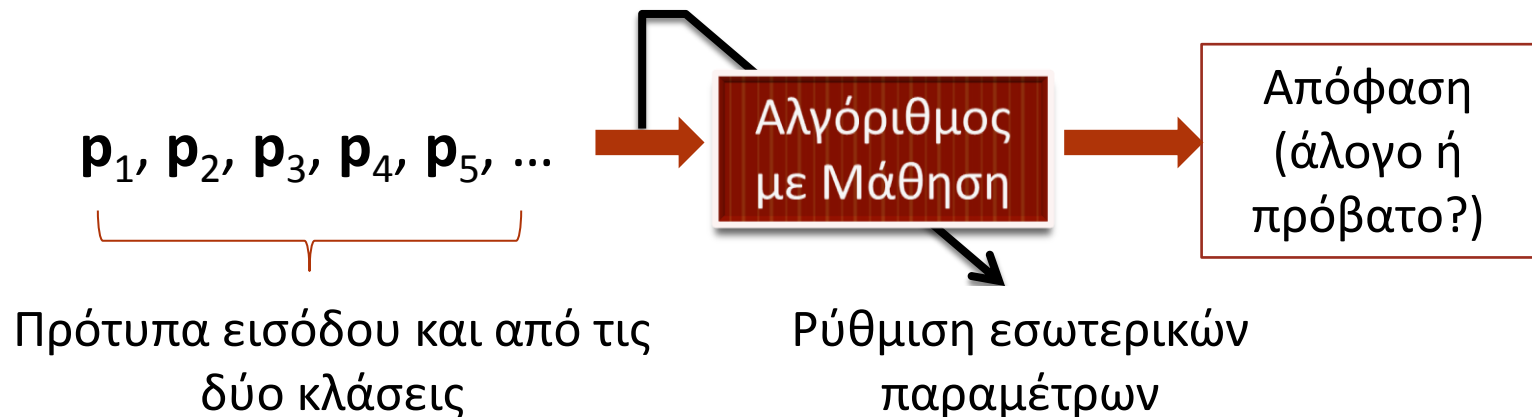
$$\mathbf{p} = [x]$$

- Στη γενική περίπτωση το πρότυπο είναι ένα διάνυσμα με n χαρακτηριστικά: $\mathbf{p} = [x_1, x_2, \dots, x_n]$

Παράδειγμα:

Αναγνώριση ζώων με μάθηση

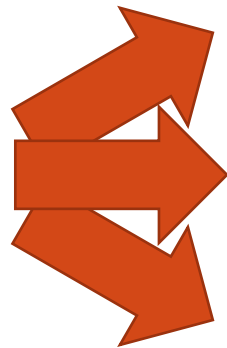
- Ο αλγόριθμος δέχεται στην είσοδο πολλά πρότυπα και από τις δύο κλάσεις και ρυθμίζει τις εσωτερικές του παραμέτρους έτσι ώστε να κάνει επιτυχώς τη διάκριση μεταξύ των δύο κλάσεων. Με άλλα λόγια, θα βρει μόνος του το κατώφλι των 2 μέτρων που απαιτείται για να ξεχωρίζει τα άλογα από τα πρόβατα.



Παράδειγμα 2:

Αναγνώριση προσώπου

- Ο άνθρωπος διαθέτει ένα εσωτερικό (βιολογικό) σύστημα αναγνώρισης μέσω του οποίου μαθαίνει από βρέφος να αναγνωρίζει τα οικεία του πρόσωπα (πχ., τη μητέρα του). Χρησιμοποιεί κυρίως οπτικά δεδομένα, αλλά και επίσης ακουστικά δεδομένα, κ.α.



Παράδειγμα 2:

Αναγνώριση προσώπου

- Το πρόβλημα της **Αναγνώρισης Προσώπου (απλούστερη εκδοχή)**:
 - Είσοδος: Μια εικόνα E η οποία περιέχει μόνο ένα πρόσωπο με γωνία λήψης από μπροστά.
 - Ζητούμενη απόκριση: Απόφαση αν η E περιέχει ένα συγκεκριμένο πρόσωπο X



Αναγνώριση προσώπου: Ωμά δεδομένα vs. Χαρακτηριστικά

- Τα «ωμά δεδομένα» του προβλήματος είναι η εικόνα εισόδου E που λαμβάνεται από την κάμερα.
- Προφανώς ένα μηχανικό σύστημα αναγνώρισης προσώπου δεν μπορεί να υλοποιηθεί σαν ένα απλό ερώτημα (query) σε μια βάση εικόνων διότι, ακόμα και στην απλή αυτή εκδοχή, μια τέτοια βάση θα έπρεπε να περιέχει όλες τις πιθανές όψεις του X για κάθε κλιμάκωση μεγέθους, αλλαγή φωτεινότητας, γωνία περιστροφής, κλπ. Οι πιθανές όψεις είναι άπειρες!

Αναγνώριση προσώπου:

Ωμά δεδομένα vs. Χαρακτηριστικά

- Ανάγκη χρήσης **χαρακτηριστικών** του προσώπου ώστε να μπορεί να διακρίνεται το X από τα άλλα πρόσωπα. Τα χαρακτηριστικά αυτά θα πρέπει να είναι αναλοίωτα στην περιστροφή, στην κλιμάκωση μεγέθους, στην αλλαγή της φωτεινότητας, κλπ.
- Ένα τέτοιο χαρακτηριστικό θα μπορούσε να είναι, για παράδειγμα, η απόσταση των ματιών μεταξύ τους σε σχέση με την απόστασή τους από το στόμα:

$$\chi = \frac{[\text{απόσταση ματιών μεταξύ τους}]}{[\text{απόσταση ματιών από το στόμα}]}$$

- Μπορεί να σκεφτεί κανείς και άλλα παρόμοια χαρακτηριστικά...

Αναγνώριση προσώπου: Γενική αρχιτεκτονική συστήματος

- Με βάση τα παραπάνω ένα σύστημα αναγνώρισης προσώπου πρέπει να εμπεριέχει ένα **υποσύστημα εξαγωγής χαρακτηριστικών (feature extraction)** που λαμβάνει ως είσοδο τα ωμά δεδομένα και παράγει ως έξοδο ένα **διάνυσμα χαρακτηριστικών (feature vector)** το οποίο θα τροφοδοτεί το υποσύστημα ταξινόμησης προτύπων το οποίο καλείται **Ταξινομητής (Classifier)**

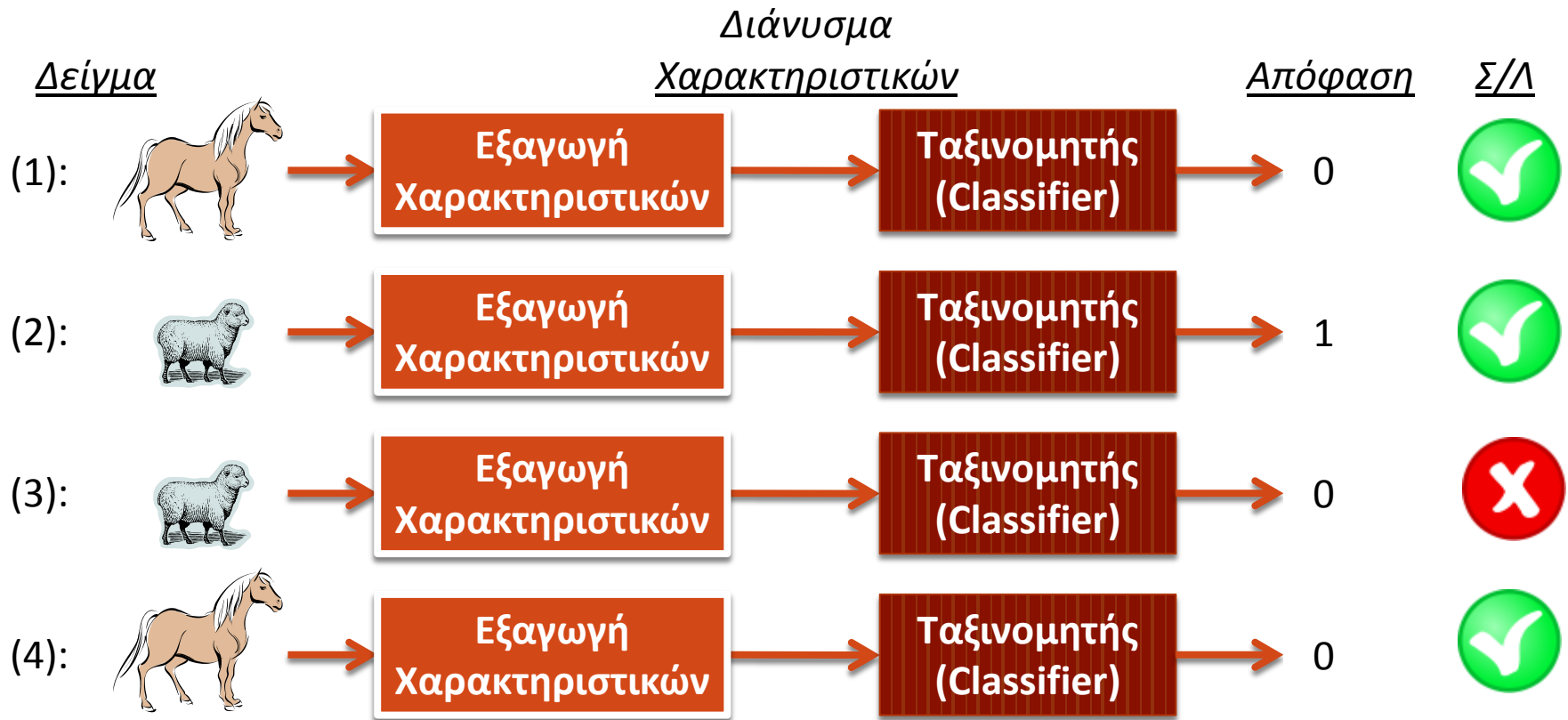


Αναγνώριση προτύπων

- Τα προβλήματα της αναγνώρισης ζώων ή προσώπων που είδαμε παραπάνω εντάσσονται στη γενικότερη κατηγορία προβλημάτων **Αναγνώρισης Προτύπων**:
- **Αναγνώριση προτύπων**: το επιστημονικό πεδίο που έχει σκοπό την κατάταξη αντικειμένων σε κατηγορίες που λέγονται **κλάσεις**.
- Οι κλάσεις δεν είναι αναγκαστικά δύο (όπως ήταν πχ στην περίπτωση της διάκρισης μεταξύ αλόγων και προβάτων). Γενικά ένα πρόβλημα αναγνώρισης μπορεί να αφορά $N > 2$ κλάσεις, όπως, πχ., η αναγνώριση των ψηφίων 0, 1, ..., 9 ή των γραμμάτων: Α, Β, Γ, ..., Ω, ή αναγνώριση προσώπων από ένα πλήθος ατόμων, κλπ.

Λειτουργία Μηχανής Αναγνώρισης

- Παράδειγμα: Δύο κλάσεις 0=άλογο, 1=πρόβατο



Εξαγωγή Χαρακτηριστικών (Feature Extraction)

- Όπως είδαμε στα προηγούμενα παραδείγματα βασικό προαπαιτούμενο για την αναγνώριση είναι η χρήση χαρακτηριστικών (features).
- Πολλές φορές τα «ωμά δεδομένα» περιέχουν μεγάλο όγκο πληροφορίας η οποία όχι δεν είναι απαραίτητη για την αναγνώριση αλλά δυσχαιραίνει το έργο του αλγορίθμου αναγνώρισης διότι καταναλίσκει μεγάλο χώρο μνήμης και κάνει την εκτέλεση του αλγορίθμου από απλώς υπολογιστικά πολύπλοκη έως απαγορευτική.

Πού εφαρμόζεται η αναγνώριση προτύπων;

- Μηχανική όραση:
 - Αναγνώριση προσώπων
 - Οπτική αναγνώριση χαρακτήρων (χειρόγραφων ή τυπωμένων)
- Ακουστική
 - Αναγνώριση ομιλίας
 - Αναγνώριση φωνής ομιλητή
 - Αναγνώριση μουσικής
- Ιατρική / Βιολογία
 - Διαγνωστική με βοήθεια υπολογιστή: επεξεργασία ηλεκτροκαρδιογραφήματος (ECG) ηλεκτρο-εγκεφαλογραφήματος (EEG), μαγνητικής τομογραφίας (MPI), κλπ
 - Αναγνώριση γονιδίων
- Κλπ...

Άλλες εφαρμογές της μάθησης

- **Παλινδρόμηση (Regression)**

- Το πρόβλημα κατά το οποίο η μηχανή καλείται να εκτιμήσει την τιμή εξόδου που αντιστοιχεί σε ένα πρότυπο εισόδου. Η τιμή εξόδου αναζητάται μέσα από ένα συνεχές σύνολο τιμών (πχ. το σύνολο των πραγματικών αριθμών, \mathcal{R})
- Η αναγνώριση προτύπων με N κλάσεις μπορεί να θεωρηθεί μια υποπερίπτωση της παλινδρόμησης στην οποία το σύνολο τιμών εξόδου είναι διακριτό, πχ, το σύνολο $\{1, 2, \dots, N\}$.



Παλινδρόμηση

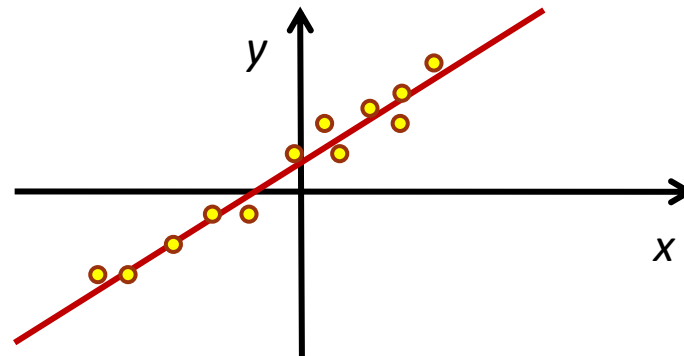
- **Παράδειγμα 1:** Έστω μια συνάρτηση $\hat{y} = f(x)$ η οποία ανήκει σε μια συγκεκριμένη κατηγορία συναρτήσεων, πχ. τις γραμμικές συναρτήσεις

$$\hat{y} = ax + b.$$

- Από την $f()$ παίρνουμε κάποια δείγματα, δηλ. ζευγάρια τιμών $[x_i, y_i]$, μόνο που στις μετρήσεις μας μπορεί να υπάρχει κάποιο σφάλμα: $y_i = \hat{y}_i + \varepsilon_i$

Δίνονται:

x_1, y_1
x_2, y_2
\vdots
x_p, y_p



- Σκοπός: η μάθηση των καλύτερων παραμέτρων a, b , ώστε για οποιοδήποτε άγνωστο x να γίνεται εκτίμηση του y με το μικρότερο δυνατό σφάλμα.

Ζητάμε: Για $x_{\text{νέο}}, y_{\text{νέο}} = ?$

Παλινδρόμηση

- **Παράδειγμα 2:** Εκτίμηση θερμοκρασίας ϑ με βάση την υγρασία u , το υψόμετρο ψ , και την πίεση του αέρα π : Τα χαρακτηριστικά εισόδου δημιουργούν το πρότυπο διάνυσμα $\mathbf{p} = [u, \psi, \pi]$.
- Στο σύστημα παλινδρόμησης δίνονται πολλά πρότυπα μαζί με τις αντίστοιχες θερμοκρασίες για κάθε πρότυπο

Δίνονται:

$$\begin{array}{l} \mathbf{p}_1 = [u_1, \psi_1, \pi_1], \quad \vartheta_1 \\ \mathbf{p}_2 = [u_2, \psi_2, \pi_2], \quad \vartheta_2 \\ \vdots \\ \mathbf{p}_p = [u_p, \psi_p, \pi_p], \quad \vartheta_p \end{array}$$

- Σκοπός: η μάθηση των καλύτερων παραμέτρων του συστήματος ώστε να γίνεται εκτίμηση της θερμοκρασίας με το μικρότερο δυνατό σφάλμα.

Ζητάμε: $\text{Για } \mathbf{p}_{\text{νέο}} = [u, \psi, \pi], \quad \vartheta = ?$

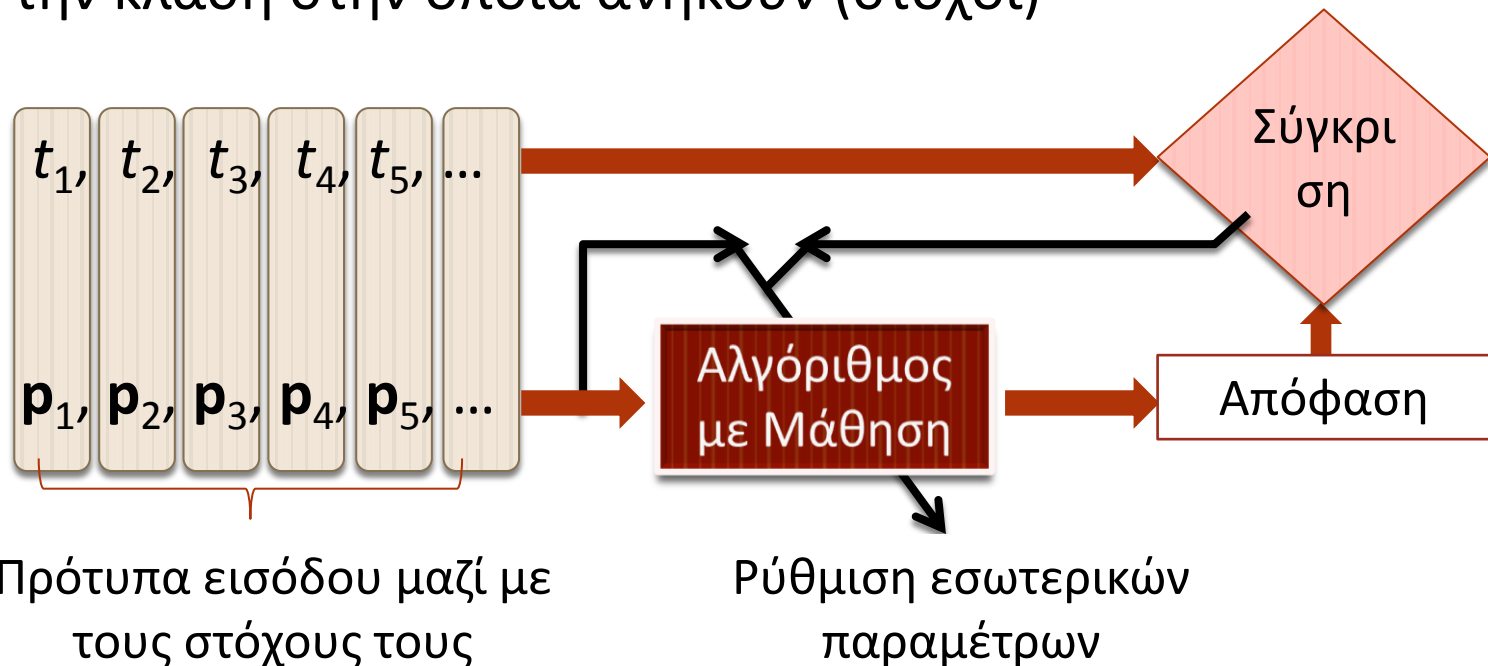
Άλλες εφαρμογές της μάθησης

- **Εξαγωγή Χαρακτηριστικών (Feature Extraction)**
- Η εξαγωγή χαρακτηριστικών μέσα από ένα μεγάλο όγκο δεδομένων μπορεί να γίνει χρησιμοποιώντας μεθοδολογία μάθησης, για παράδειγμα με χρήση αυτοπροσαρμοζόμενων μεθόδων στατιστικής ανάλυσης (πχ. Ανάλυση Κυρίων Συνιστωσών, Learning Vector Quantization, κα)
- **Συμπίεση Δεδομένων (Data Compression)**
- Η συμπίεση δεδομένων είναι συναφής με την εξαγωγή χαρακτηριστικών καθώς όγκος δεδομένων μεγάλων διαστάσεων αντικαθίσταται από δεδομένα μικρότερης διάστασης. Επιθυμητό είναι τα δεδομένα μικρής διάστασης να μπορούν να ανακατασκευάζουν τα αρχικά δεδομένα με όσο το δυνατό λιγότερο σφάλμα.

Οι δύο βασικοί τύποι μάθησης

ΜΑΘΗΣΗ ΜΕ ΕΠΙΒΛΕΨΗ (SUPERVISED LEARNING)

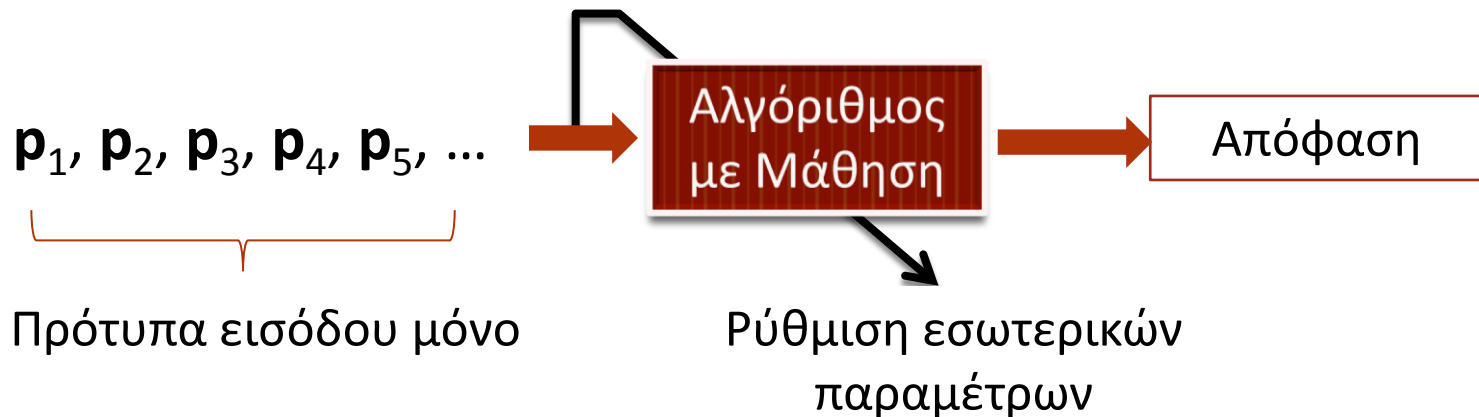
- Ο αλγόριθμος χρησιμοποιεί τα πρότυπα εισόδου μαζί με την κλάση στην οποία ανήκουν (στόχοι)



Οι δύο βασικοί τύποι μάθησης

ΜΑΘΗΣΗ ΧΩΡΙΣ ΕΠΙΒΛΕΨΗ (UNSUPERVISED LEARNING)

- Ο αλγόριθμος χρησιμοποιεί μόνο τα πρότυπα εισόδου χωρίς να γνωρίζει σε ποια κλάση ανήκει το καθένα.

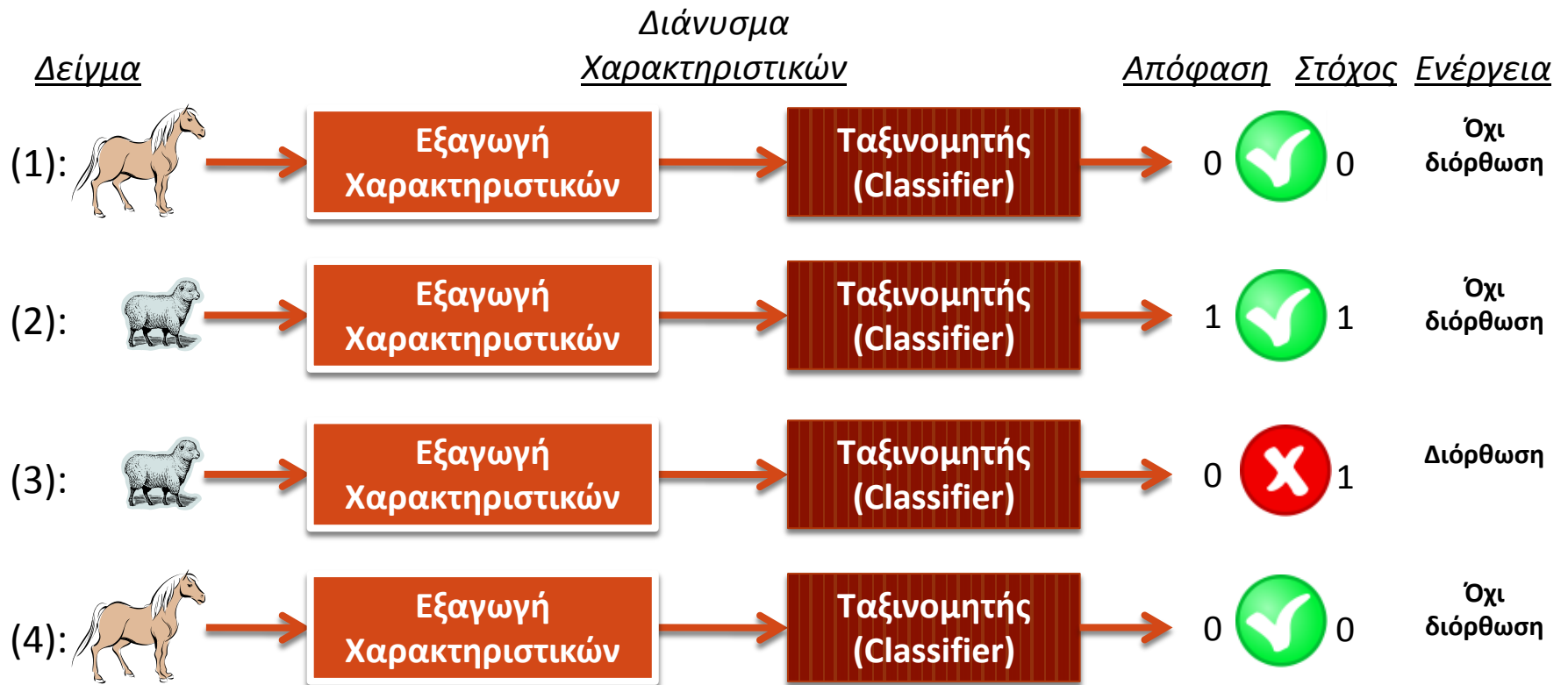


Βασικές έννοιες μάθησης: Εκπαίδευση και Ανάκληση

- **Ορισμός. Εκπαίδευση:** Η παρουσίαση πολλών προτύπων στο σύστημα (με ή χωρίς στόχους, ανάλογα με το τύπο μάθησης) με σκοπό την ρύθμιση των παραμέτρων του ώστε αυτό να βελτιώνεται στην λειτουργία αναγνώρισης ή σε όποια άλλη λειτουργία τάχθηκε.
- **Ορισμός. Ανάκληση:** Η εισαγωγή ενός ή περισσότερων προτύπων με σκοπό την εξαγωγή της απόκρισης του συστήματος χωρίς εκπαίδευση.

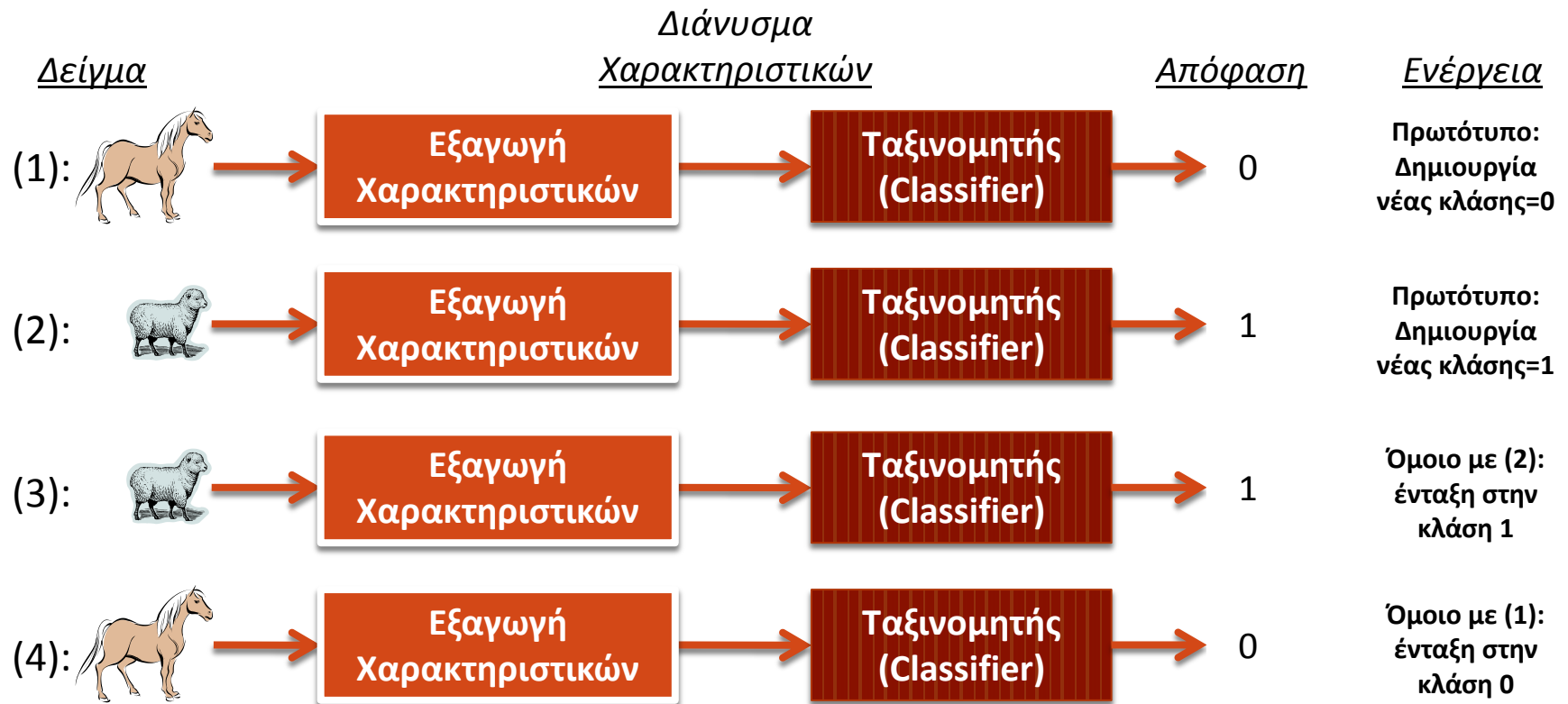
Εκπαίδευση Με Επίβλεψη

- Παράδειγμα:



Εκπαίδευση Χωρίς Επίβλεψη

- Παράδειγμα:



Μέτρηση επίδοσης συστήματος αναγνώρισης

- **Πίνακας σύγχυσης (Confusion matrix):** Ο πίνακας **C** με στοιχεία $C(i,j)$ = πλήθος δειγμάτων που ανήκουν στην κλάση i και ταξινομήθηκαν στην κλάση j . Παράδειγμα:

Ταξινομήθηκαν στην κλάση:

	1	2	...	N
1	37	3	...	14
2	6	82	...	6
⋮	⋮	⋮		⋮
N	0	2	...	152

Ανήκουν στην κλάση:

- Άθροισμα των στοιχείων της διαγωνίου = πόσα πρότυπα συνολικά από τις N κλάσεις ταξινομήθηκαν **σωστά**,
- Άθροισμα των υπόλοιπων στοιχείων πλην της διαγωνίου = πόσα πρότυπα συνολικά ταξινομήθηκαν **λάθος**.

Δυαδική ταξινόμηση: Πίνακας Σύγκυσης

- Ειδική περίπτωση: Όταν οι κλάσεις είναι δύο, τότε λέμε ότι έχουμε πρόβλημα **δυαδικής ταξινόμησης**. Συχνά χρησιμοποιούμε τους αριθμούς 0, 1 για την αρίθμηση των κλάσεων αντί για 1, 2. Πίνακας σύγκυσης:

	<i>negative</i>	<i>positive</i>
<i>0</i>	98	3
<i>1</i>	8	110

True negatives

False positives

False negatives

True positives

The diagram shows a 2x2 confusion matrix for binary classification. The columns are labeled 'negative' and 'positive'. The rows are labeled '0' and '1'. The matrix cells contain the counts: (0, negative) = 98, (0, positive) = 3, (1, negative) = 8, and (1, positive) = 110. Callouts with arrows point to specific cells: 'True negatives' points to the cell (0, negative) with a blue box; 'False positives' points to the cell (0, positive) with a red box; 'False negatives' points to the cell (1, negative) with a red box; and 'True positives' points to the cell (1, positive) with a blue box.

Δείκτες Επίδοσης: Ευστοχία (Precision), Επιστροφή (Recall)

- **Ευστοχία :** $Precision = \frac{tp}{tp + fp} = \frac{110}{110 + 3} = 97,3\%$

- **Επιστροφή:** $Recall = \frac{tp}{tp + fn} = \frac{110}{110 + 8} = 93,2\%$

- **Ακρίβεια:** $Accuracy = \frac{tp + tn}{tp + fp + tn + fn} = \frac{110 + 98}{110 + 98 + 3 + 8} = 94,9\%$

Ταξινομήθηκαν στην κλάση:

	0	1
0	98	3
1	8	110

Ανήκουν στην κλάση:

Labels: TN (True Negative), FN (False Negative), FP (False Positive), TP (True Positive)

Μοντέλα Μάθησης

- **Στατιστικά μοντέλα Bayes** (ή Μπείζιανά μοντέλα). Βασισμένα στην στατιστική και στο νόμο του Bayes, τα μοντέλα αυτά προσπαθούν να εκτιμήσουν την πιο «πιθανή» λύση στο πρόβλημα με βάση τα δεδομένα που έχουμε διαθέσιμα.
- **Τεχνητά Νευρωνικά Δίκτυα** (ΤΝΔ). Εμπνευσμένα από την λειτουργία των νευρώνων και του εγκεφάλου χρησιμοποιούν εκπαίδευση για την ρύθμιση των εσωτερικών τους παραμέτρων που καλούνται «συναπτικά βάρη». Υπάρχουν πολλές αρχιτεκτονικές ΤΝΔ που ακολουθούν τόσο το μοντέλο της μάθησης με επίβλεψη όσο και το μοντέλο της μάθησης χωρίς επίβλεψη.