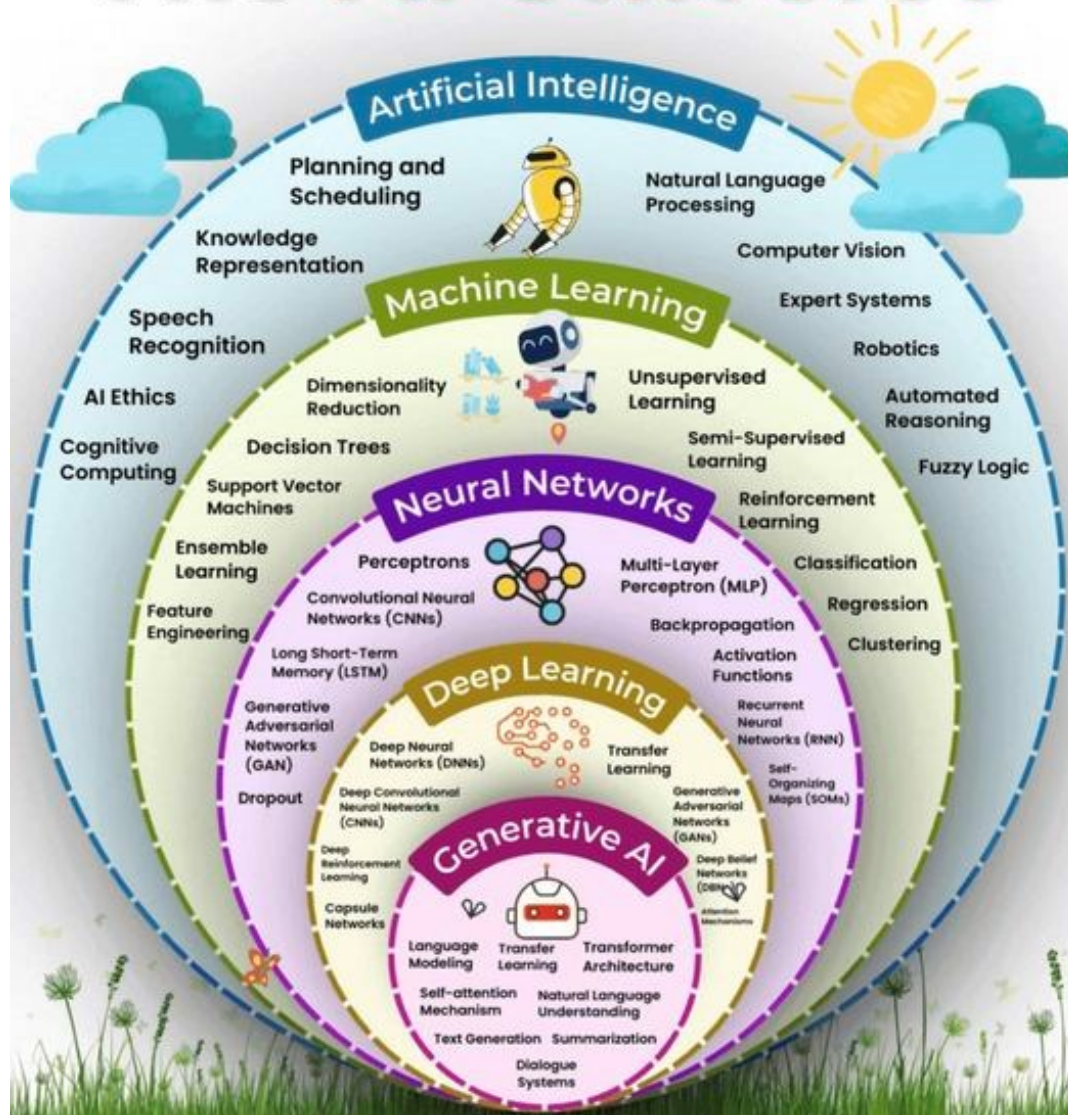


The AI Universe



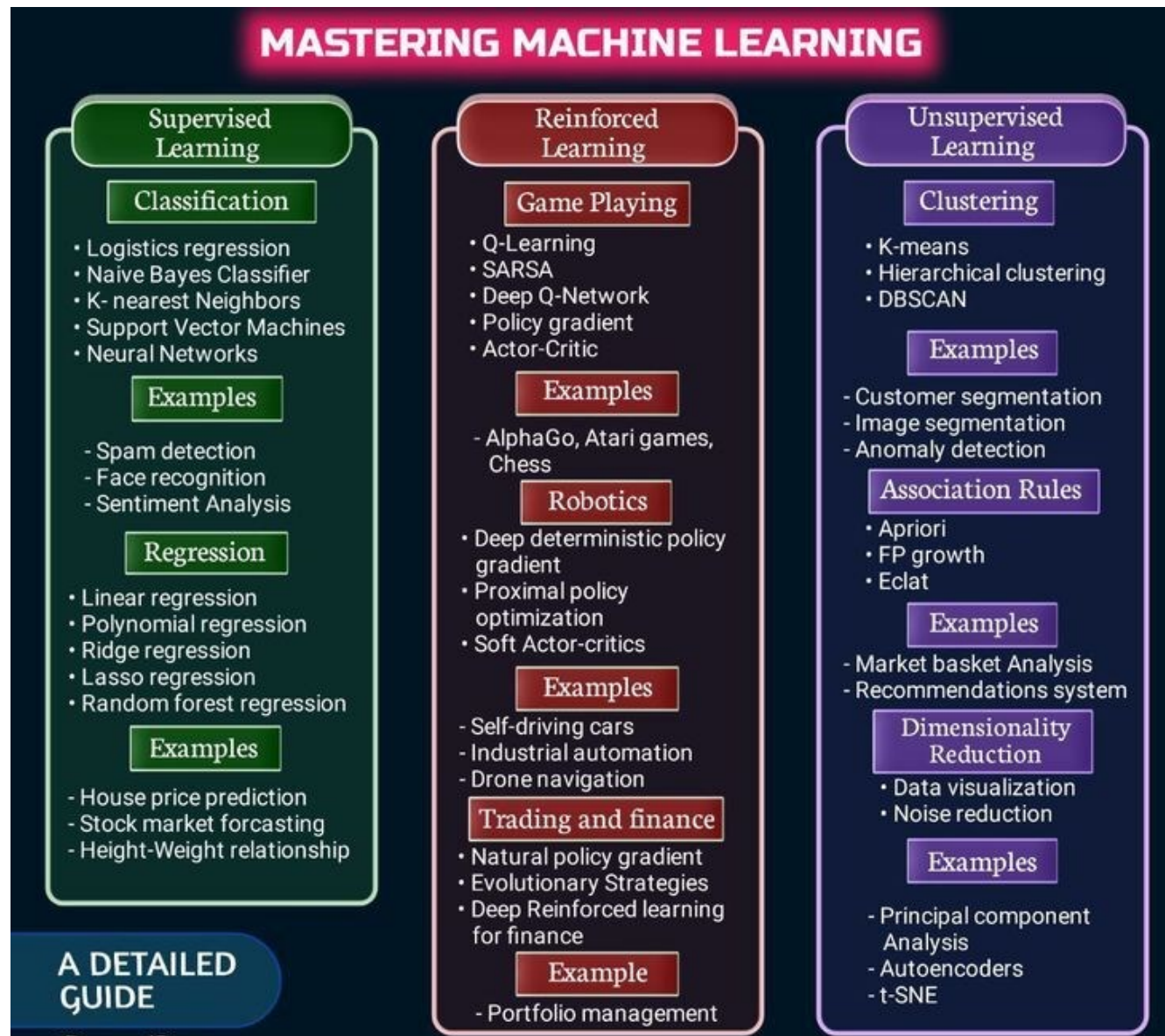
Μηχανική Μάθηση – Machine Learning

DWS101 - ΠΜΣ "ΕΠΙΣΤΗΜΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΠΑΓΚΟΣΜΙΟΥ ΙΣΤΟΥ"

AI101 - ΠΜΣ "ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ"



Κεφάλαιο 1



Εισαγωγή



Εισαγωγή

- ❖ Οι περισσότερες εφαρμογές ΤΝ του παρελθόντος αναπτύχθηκαν κωδικοποιώντας τη γνώση του προβλήματος σε μια τυπική γλώσσα αναπαράστασης γνώσης και χρησιμοποιώντας μηχανισμούς συλλογιστικής που αναπαρήγαγαν την ανθρώπινη συλλογιστική (**rule/knowledge-based approach**).
 - ❑ Συνεπώς η επίλυση των προβλημάτων βασιζόταν σε μεγάλο βαθμό στη ανθρώπινη γνώση και διαίσθηση.
- ❖ Σε πολύπλοκα προβλήματα όμως, με πολλές παραμέτρους, πολλά δεδομένα και πολύπλοκες ή άγνωστες σχέσεις μεταξύ αυτών, η παραπάνω "κλασική" προσέγγιση αδυνατούσε να δώσει λύση.
 - ❑ Πολλά projects που υλοποιήθηκαν με κωδικοποίηση της γνώσης σε μια τυπική γλώσσα αναπαράστασης, εγκαταλείφθηκαν.
- ❖ Για να ξεπεραστούν αυτές οι δυσκολίες αναπτύχθηκαν συστήματα ΤΝ τα οποία αποκτούν την ικανότητα να εκτελούν μια συγκεκριμένη λειτουργία επειδή καταφέρνουν να μάθουν μόνα τους τον τρόπο, μέσα από παραδείγματα (δεδομένα)
 - ❑ Αυτή η ικανότητα μάθησης από το σύστημα ονομάζεται **Μηχανική Μάθηση**.

[DataRobot Presents Automated Machine Learning](#)



Γνωστικό σύστημα

- ❖ **Γνωστικό σύστημα (cognitive system)** ονομάζεται ένα φυσικό ή τεχνητό σύστημα επεξεργασίας πληροφορίας, συμπεριλαμβανομένων εκείνων με δυνατότητες αντίληψης, μάθησης, συλλογισμού, λήψης απόφασης, επικοινωνίας και δράσης.
- ❖ Η έννοια της μάθησης σε ένα γνωστικό σύστημα, όπως γίνεται αντιληπτή στην καθημερινή ζωή, μπορεί να συνδεθεί με δύο βασικές ικανότητές του:
 - ☐ την πρόσκτηση γνώσης κατά την αλληλεπίδρασή του με το περιβάλλον μέσα στο οποίο δραστηριοποιείται, και
 - ☐ τη βελτίωση μέσω επανάληψης του τρόπου με τον οποίο εκτελεί μία ενέργεια, και συνεπώς και την απόδοσή του ως προς κάποια μετρική,
 - ✓ όπως για παράδειγμα τον απαιτούμενο χρόνο ή την ποιότητα του αποτελέσματος.
- ❖ Ένα σύστημα με δυνατότητα μάθησης μεταβάλλεται διαρκώς προς το "καλύτερο", όπως και αν αυτό ορίζεται, αναφορικά με τις λειτουργίες που είναι σε θέση να εκτελέσει, είτε
 - ☐ με μεταβολές στη βάση γνώσης του,
 - ☐ μετασχηματίζοντας την εσωτερική τους δομή (όπως τα βάρη στα τεχνητά NN)



Διάφοροι ορισμοί μάθησης

❖ Μάθηση

- ❑ "Η μάθηση σηματοδοτεί αλλαγές σε ένα σύστημα, που είναι προσαρμοστικές με την έννοια ότι επιτρέπουν στο σύστημα να κάνει την ίδια εργασία, ή εργασίες της ίδιας κατηγορίας, πιο αποδοτικά και αποτελεσματικά την επόμενη φορά" (Simon, 1983).
- ❑ "Μάθηση είναι να κάνουμε χρήσιμες αλλαγές στο μυαλό μας " (Minsky, 1985).
- ❑ "Μάθηση είναι η δημιουργία ή η αλλαγή της αναπαράστασης των εμπειριών" (Michalski, 1986).

❖ Μηχανική Μάθηση

- ❑ "...η μελέτη υπολογιστικών μεθόδων για την απόκτηση νέας γνώσης, νέων δεξιοτήτων και νέων τρόπων οργάνωσης της υπάρχουσας γνώσης" (Carbonell, 1987).
- ❑ "Ένα πρόγραμμα υπολογιστή θεωρείται ότι **μαθαίνει** από την εμπειρία ***E*** σε σχέση με μια κατηγορία εργασιών ***T*** και μια μετρική απόδοσης ***P***, αν η απόδοση του σε εργασίες της ***T***, όπως μετριοούνται από την ***P***, βελτιώνονται με την εμπειρία ***E***" (Mitchell, 1997).
- ❑ "Κάτι μαθαίνει όταν αλλάζει την συμπεριφορά του κατά τέτοιο τρόπο ώστε να αποδίδει καλύτερα στο μέλλον" (Witten & Frank, 2000).



Ορισμός Μηχανικής Μάθησης

- ❖ Ο άνθρωπος προσπαθεί να κατανοήσει το περιβάλλον του παρατηρώντας το και δημιουργώντας μια απλοποιημένη (αφαιρετική) εκδοχή του που ονομάζεται νοητικό μοντέλο (mental model) ή απλά μοντέλο (**model**).
- ❑ Η διαδικασία γενικότερα ονομάζεται επαγωγή (induction) και η δημιουργία ενός τέτοιου μοντέλου ονομάζεται επαγωγική μάθηση (**inductive learning**).
- ❖ Επιπλέον ο άνθρωπος έχει τη δυνατότητα να οργανώνει και να συσχετίζει τις εμπειρίες και τις παρατηρήσεις του δημιουργώντας νέες δομές που ονομάζονται νοητικά πρότυπα (mental patterns) ή απλά πρότυπα (**patterns**).
- ❖ **Μηχανική μάθηση** (machine learning) είναι η δημιουργία μοντέλων ή προτύπων με βάση ένα σύνολο δεδομένων, από ένα υπολογιστικό σύστημα.



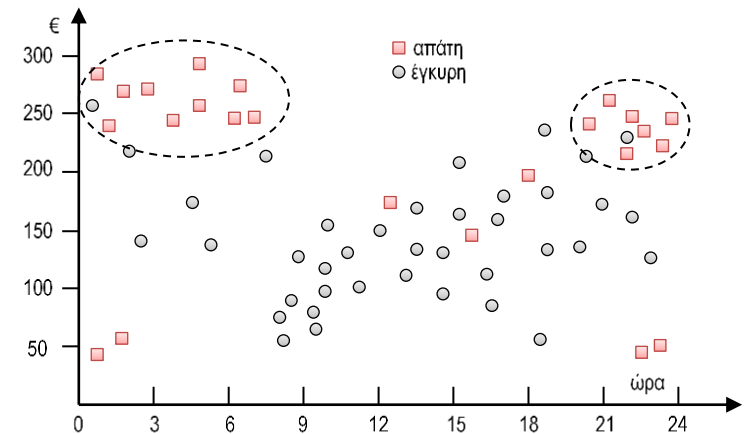
Γιατί να μελετήσουμε Μηχανική Μάθηση;

- ❖ Όπως ήδη αναφέρθηκε, είναι πολύ δύσκολο έως αδύνατο να επιλυθούν πολύπλοκα προβλήματα με πολλά δεδομένα και πολύπλοκες σχέσεις μεταξύ των δεδομένων, αναπαριστώντας τη γνώση του προβλήματος σε μια τυπική γλώσσα.
 - ❑ Τέτοια προβλήματα όμως έχουν μεγάλο αντίκτυπο στην κοινωνία και επιπλέον παρουσιάζουν μεγάλο επιστημονικό και επιχειρηματικό ενδιαφέρον
 - ✓ επομένως η αντιμετώπισή τους με τεχνικές μηχανικής μάθησης ήρθε σαν φυσικό επακόλουθο.
- ❖ Παραδείγματα τέτοιων προβλημάτων είναι τα ακόλουθα:
 - ❑ Χρήση ιστορικών δεδομένων για λήψη αποφάσεων:
 - ✓ Ιατρικά αρχεία ασθενών → Ιατρική γνώση
 - ❑ Δημιουργία σύνθετων προγραμμάτων:
 - ✓ Ρομπότ που μαθαίνει ένα χάρτη του περιβάλλοντος με το να περιπλανηθεί σε αυτό
 - ✓ Προγράμματα που μαθαίνουν να παίζουν παιχνίδια παίζοντας με τον εαυτό τους
 - ❑ Κατανόηση της ανθρώπινης νοημοσύνης:
 - ✓ Πως μαθαίνουμε να μιλάμε; Πως αναγνωρίζουμε πρόσωπα;
 - ❑ Δημιουργία προγραμμάτων που προσαρμόζονται στο χρήστη:
 - ✓ Αναγνώριση ομιλίας και γραφικού χαρακτήρα, Ευφυείς διεπαφές χρήστη
 - ❑ Δημιουργία πραγματικής Τεχνητής Νοημοσύνης
 - ✓ "Αν ένα ευφυές σύστημα – που έχει σχεδιαστεί και κατασκευαστεί ευφυέστατα – δεν μπορεί να μάθει να μην επαναλαμβάνει τα λάθη του, τότε δεν είναι ούτε τόσο ευφυές όσο ένα σκουλήκι ή μια θαλάσσια ανεμώνη ή μια γάτα." (Oliver Selfridge)



Παράδειγμα προβλήματος (1)

- ❖ Έστω το διάγραμμα στο διπλανό Σχήμα όπου αναπαρίστανται οι συναλλαγές μέσω πιστωτικών καρτών σε διάγραμμα χρόνου, δηλαδή πότε έγινε η συναλλαγή, και ποσού, δηλαδή τι ποσό χρεώθηκε στην κάρτα.
- ❖ Οι κύκλοι είναι γνωστές νόμιμες συναλλαγές, ενώ τα τετράγωνα αποδεδειγμένες περιπτώσεις απάτης (π.χ. συναλλαγές με κλεμμένη πιστωτική κάρτα).
- ❖ Πώς είναι δυνατό να αξιοποιηθούν αυτά τα ιστορικά (παλιά) δεδομένα ώστε να κατασκευαστεί ένα σύστημα που θα εντοπίζει μια συναλλαγή με κλεμμένη κάρτα;
- ❖ Παρατηρώντας το διάγραμμα, θα μπορούσαμε να πούμε ότι συναλλαγές που γίνονται πολύ αργά το βράδυ ή τις πρώτες πρωινές ώρες και αφορούν σε σχετικά μεγάλο ποσό (>200€) είναι κατά βάση ύποπτες περιπτώσεις και να γράψουμε μια απλή λογική έκφραση τύπου if...then.
- ❖ Τι γίνεται όμως όταν οι παράμετροι του προβλήματος δεν είναι μόνο δύο αλλά αρκετές δεκάδες ή και εκατοντάδες;



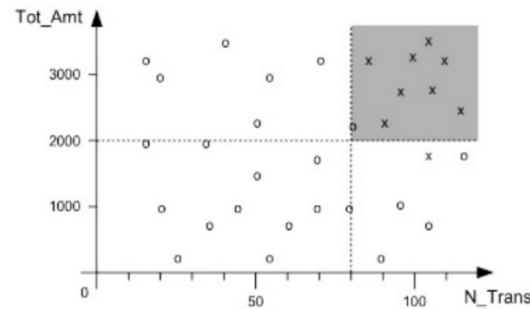


Παράδειγμα προβλήματος (1) (cont...)

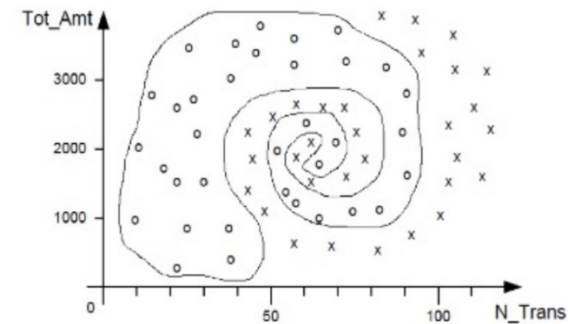
- ❖ Αυτό που καλείται να κάνει μηχανική μάθηση στο παραπάνω πρόβλημα είναι να δοκιμάσει διάφορες υποθέσεις (τρόπους) μοντελοποίησης της σχέσης που ίσως υπάρχει μεταξύ των παρατηρήσεων (στο παράδειγμα, η ώρα συναλλαγής και το ποσό) και του συμπεράσματος (απάτη ή όχι).
- ❖ Αν καταφέρει και το κάνει, εμείς θα γενικεύσουμε και θα θεωρήσουμε ότι εφόσον ισχύει για τα συγκεκριμένα δεδομένα θα ισχύει και γενικότερα σε άγνωστα δεδομένα (**υπόθεση επαγωγικής μάθησης**).



Παράδειγμα 2: Fraud detection - Expert Rules approach

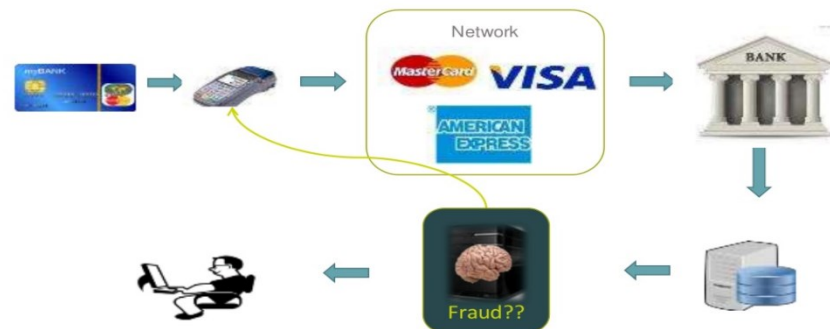


Rule: IF N_Trans > 80 AND
Tot_Amt > 2000 THEN fraud



Rule: ?? We can learn this by
means of **Machine Learning**

- ❖ Machine Learning can learn automatically rules able to find fraudulent patterns
IF COUNTRY=USA & LANGUAGE=EN & HAD_TEST=TRUE & NB_TX>10 & GENDER=MALE & AGE> 50
& ONLINE=TRUE & AMOUNT>1,000 & BANK=XXX THEN class = fraud





Ενδεικτικές Εφαρμογές

- ✓ Παραγωγή ιατρικής γνώσης από τα ιατρικά αρχεία ασθενών και γενικότερα εξαγωγή γνώσης από δεδομένα.
- ✓ Ιατρική διάγνωση από εξετάσεις απεικόνισης (αξονική/μαγνητική τομογραφία, ακτινογραφία, υπέρηχο, κτλ.).
- ✓ Αντίληψη και κατανόηση του περιβάλλοντος χώρου με αισθητήρες όρασης, κτλ.
- ✓ Κατανόηση και παραγωγή γραπτού/χειρόγραφου λόγου και ομιλίας.
- ✓ Αυτόματη μετάφραση.
- ✓ Αυτόματος υποτιτλισμός σε video και παραγωγή λεζάντας σε εικόνες.
- ✓ Εντοπισμός απάτης σε συναλλαγές.
- ✓ Δημιουργία αυτοκινούμενων οχημάτων.
- ✓ Δημιουργία αυτόνομων συσκευών-βοηθών.
- ✓ Παροχή εξατομικευμένων συστάσεων – recommender systems (π.χ. για αγορές).
- ✓ Μάθηση παιχνιδιών σε επίπεδο ανώτερο των καλύτερων ανθρώπων-παικτών.
- ✓ Κατανόηση της ανθρώπινης νοημοσύνης μέσω πειραματισμού με τεχνητή νοημοσύνη και αντίστροφα (πώς μαθαίνουμε να μιλάμε, πώς αναγνωρίζουμε πρόσωπα, κλπ.).
- ✓ Δημιουργία προγραμμάτων που προσαρμόζονται στο χρήστη.
- ✓ Βιοπληροφορική: στοίχιση ακολουθιών, ανάλυση δεδομένων από μικροσυστοιχίες
- ✓ Μηχανική όραση: αναγνώριση αντικειμένων, τμηματοποίηση εικόνας
- ✓ Ρομποτική: υπολογισμός κατάστασης, δημιουργία χάρτη, λήψη απόφασης
- ✓ Γραφικά: Δημιουργία ρεαλιστικών προσομοιώσεων
- ✓ Ομιλία: αναγνώριση, εξακρίβωση ομιλητή
- ✓ Οικονομική ανάλυση: πρόβλεψη τιμών μετοχών
- ✓ Ηλεκτρονικό εμπόριο: πράκτορες για αυτόματο εμπόριο, διατήρηση πελατών
- ✓ Φαρμακευτική: διάγνωση, θεραπεία, σχεδίαση φαρμάκων
- ✓ Ηλεκτρονικά παιχνίδια: Σχεδιασμός ικανών αντιπάλων
- ✓ Πολυμέσα: Ανάκτηση με βάση το περιεχόμενο



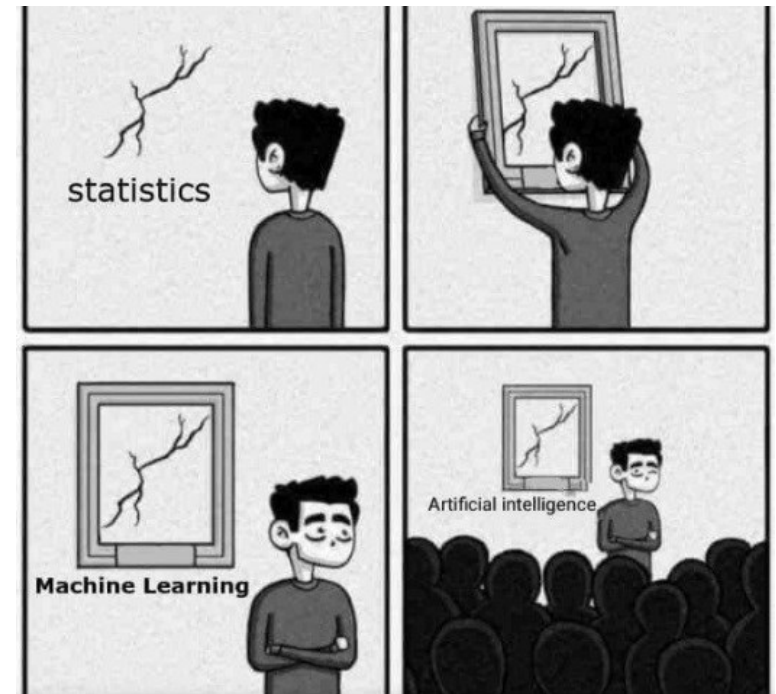
Συγγενικά αντικείμενα

- ❖ Ανακάλυψη Γνώσης από Βάσεις Δεδομένων (Knowledge Discovery in Databases-KDD)
 - ❑ Τυπικά, η μηχανική μάθηση σχετίζεται περισσότερο με το βήμα "εξόρυξη δεδομένων" (data mining) στη διαδικασία της ανακάλυψης γνώσης.
 - ✓ Κυρίως παρέχει τους αλγορίθμους μάθησης.
 - ✓ Έχει περισσότερες επιρροές από την Τεχνητή Νοημοσύνη.
 - ❑ Η ανακάλυψη γνώσης αφορά στην αποδοτική εκτέλεση των αλγορίθμων αυτών για δεδομένα που δεν χωράνε στην κύρια μνήμη του υπολογιστή καθώς και στη γενικότερη διαχείριση και ανάλυση δεδομένων που βρίσκονται αποθηκευμένα σε συστήματα βάσεων δεδομένων.
 - ✓ Τεχνολογία χαρακτηριστικών (Feature engineering)
 - ✓ Έχει περισσότερες επιρροές από τον τομέα των Βάσεων Δεδομένων
- ❖ Αναγνώριση Προτύπων (Pattern Recognition)
 - ❑ Η αναγνώριση προτύπων ασχολείται σχεδόν αποκλειστικά με αριθμητικά δεδομένα που κυρίως προέρχονται από σήματα, δηλαδή ήχους, εικόνες και βίντεο.
 - ❑ Αντίθετα, η μηχανική μάθηση ασχολείται και με συμβολικά δεδομένα που μπορεί να έχουν καταγραφεί με οποιοδήποτε τρόπο.



Σύγκριση με την Στατιστική

- ❖ Αν και οι δύο περιοχές έχουν πάρα πολλά κοινά, εντούτοις η **Μηχανική Μάθηση**:
 - ❑ Δεν απαιτεί πρότερη γνώση για την πιθανή υποκείμενη σχέση μεταξύ των μεταβλητών. Το μόνο που χρειάζεται είναι να εισαχθούν όλα τα δεδομένα στον αλγόριθμο ο οποίος τα επεξεργάζεται ανακαλύπτοντας μοντέλα/πρότυπα που μπορεί να χρησιμοποιηθούν για προβλέψεις σε μελλοντικά δεδομένα.
 - ❑ Παράγει τις υποθέσεις (μοντέλα/πρότυπα) και ελέγχει την ποιότητά τους σε σχέση με τα δεδομένα.
 - ❑ Αντιμετωπίζει έναν αλγόριθμο ως μαύρο κουτί, δηλαδή δεν την ενδιαφέρει η λειτουργία του αλγορίθμου αλλά μόνο η απαιτούμενη είσοδος και τα παραγόμενα αποτελέσματα του.
 - ❑ Εφαρμόζεται συνήθως σε πολυδιάστατα δεδομένα.
 - ❑ Συνήθως όσο περισσότερα δεδομένα χρησιμοποιούνται τόσο ποιοτικότερο είναι το αποτέλεσμα (π.χ. υπολογίζεται ακριβέστερη πρόβλεψη).
 - ❑ Απαιτεί κατάλληλη επαλήθευση για το ποιος αλγόριθμος εφαρμόζεται καλύτερα στα δεδομένα.
 - ❑ Είναι υποπεριοχή της επιστήμης των υπολογιστών και της ΤΝ.





❖ Αντίθετα η Στατιστική:

- ☐ Πρέπει να γνωρίζει τη μέθοδο συλλογής των δεδομένων, την κατανομή του πληθυσμού και τις αναμενόμενες ιδιότητες των αποτελεσμάτων αν εκτελεστεί η πειραματική εργασία πολλές φορές.
- ☐ Επαληθεύει ή απορρίπτει με τη βοήθεια των δεδομένων, υποθέσεις που διατυπώνει ο άνθρωπος, όχι το σύστημα.
- ☐ Πρέπει να γνωρίζει επακριβώς τι αναζητείται και να γίνει η επιλογή των κατάλληλων παραμέτρων που θα βοηθήσουν σε αυτό.
- ☐ Συνήθως εφαρμόζεται σε δεδομένα λίγων διαστάσεων.
- ☐ Απαιτεί επαλήθευση για να επιβεβαιωθεί η ορθότητα των αποτελεσμάτων.
- ☐ Είναι υποπεριοχή των Μαθηματικών.



Σύντομη ιστορία

- ❖ Μελετήθηκε από τότε που ξεκίνησε και η Τεχνητή Νοημοσύνη.
 - ❑ Το ανάλογο ενός νευρώνα (Rosenblatt, 1957)
 - ❑ Μάθηση συνάρτησης αξιολόγησης για αναζήτηση στο παιχνίδι ντάμα (Samuel, 1959)
- ❖ Ονομάστηκε "Μηχανική Μάθηση" στις αρχές του 1980
- ❖ Γνώρισε μεγάλη ανάπτυξη κυρίως στη δεκαετία του '90 και αργότερα του '10. Η πρόσφατη μεγάλη ανάπτυξη οφείλεται κύρια στους ακόλουθους παράγοντες:
 - ❑ Συνεχόμενη αύξηση των δεδομένων
 - ✓ Επιχειρήσεις: Τραπεζικές συναλλαγές, αγορές
 - ✓ Επιστήμες: Πειράματα φυσικής και χημείας
 - ✓ Προσωπικά: Κείμενα, φωτογραφίες, ήχοι, βίντεο
 - ❑ Βελτίωση της διαθέσιμης υπολογιστικής ισχύος
 - ❑ Πρόοδος σε αλγορίθμους και θεωρία
 - ❑ Αυξανόμενη ζήτηση για βιομηχανικές εφαρμογές



Timeline of Machine Learning

Decade ↕	Summary ↕
pre-1950	Statistical methods are discovered and refined.
1950s	Pioneering machine learning research is conducted using simple algorithms.
1960s	Bayesian methods are introduced for probabilistic inference in machine learning. ^[1]
1970s	' AI winter ' caused by pessimism about machine learning effectiveness.
1980s	Rediscovery of backpropagation causes a resurgence in machine learning research.
1990s	Work on Machine learning shifts from a knowledge-driven approach to a data-driven approach. Scientists begin creating programs for computers to analyze large amounts of data and draw conclusions – or "learn" – from the results. ^[2] Support-vector machines (SVMs) and recurrent neural networks (RNNs) become popular. ^[3] The fields of computational complexity via neural networks and super-Turing computation started. ^[4]
2000s	Support-Vector Clustering ^[5] and other kernel methods ^[6] and unsupervised machine learning methods become widespread. ^[7]
2010s	Deep learning becomes feasible, which leads to machine learning becoming integral to many widely used software services and applications. Deep learning spurs huge advances in vision and text processing.
2020s	Generative AI leads to revolutionary models, creating a proliferation of foundation models both proprietary and open source, notably enabling products such as ChatGPT (text-based) and Stable Diffusion (image based). Machine learning and AI enter the wider public consciousness. The commercial potential of AI based on machine learning causes large increases in valuations of companies linked to AI.

https://en.wikipedia.org/wiki/Timeline_of_machine_learning



Δεδομένα

THE INTERNET IN **2023** EVERY MINUTE



Created by: eDiscovery Today & LTMG



Huge Amounts of Data

- ❖ We are drowning in the deluge of data that are collected world-wide, while starving for knowledge at the same time
- ❖ Big Data
 - ❑ Σε έρευνα του 2012 επιβεβαιώθηκε ότι καθένας από εμάς παράγει κατά μέσο όρο 512 MB δεδομένων την ημέρα!
 - ❑ Αυτό σημαίνει ότι εάν, για παράδειγμα, θέλαμε να γράψουμε σε χαρτί όλα αυτά τα δεδομένα που συλλέγονται σε καθημερινό επίπεδο, τότε ο όγκος αυτών των χαρτιών θα ήταν τετραπλάσιος της απόστασης που έχει η Γη από τον ήλιο.
 - ❑ Τα στοιχεία δείχνουν ότι μέχρι το 2025 ο κάθε άνθρωπος θα παράγει πάνω από 62 GB ημερησίως.
Πηγή: www.lifo.gr



**“Mining needle in a haystack.
So much hay and so little time”**



Big Data Characteristics - The 4 V's

- ❖ Volume
 - ☐ The size of the data determines the value and potential insight
- ❖ Variety
 - ☐ The type and nature of the data
 - ☐ Big data draws from text, images, audio, video
- ❖ Velocity
 - ☐ The speed at which the data is generated and processed
 - ☐ Big data is often available in real-time
- ❖ Veracity
 - ☐ The data quality of captured data can vary greatly, affecting the accurate analysis

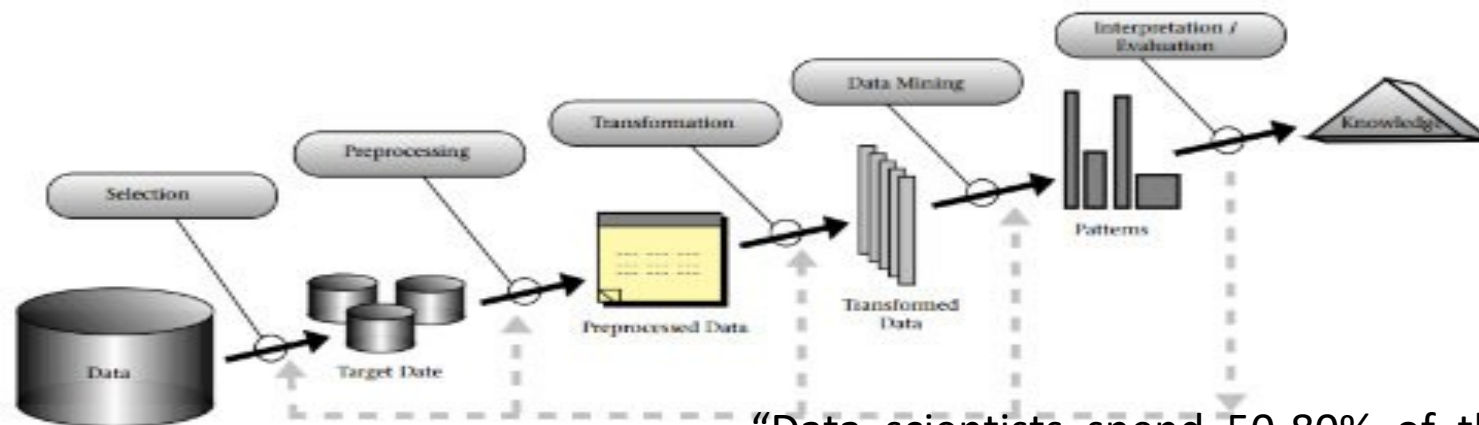
Traditional computing techniques are not able to handle such large datasets.

Adopted solution: Artificial Intelligence

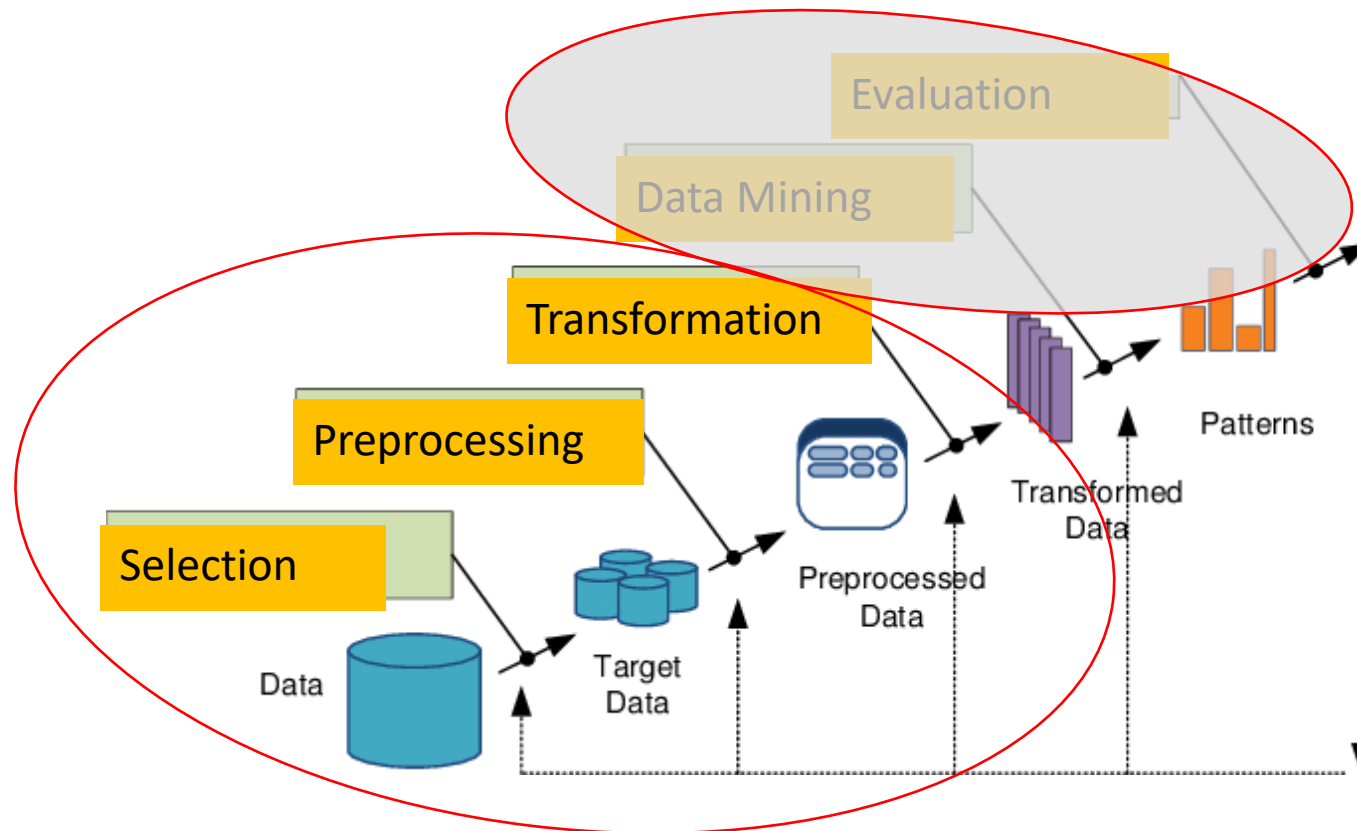


Knowledge Discovery in Databases (KDD)

- ❖ Is a multi-faceted process comprising 7 steps.
- ❖ Encompasses a variety of tasks that are not statistical in nature
 - ❑ Covers the entire process of data analysis, including data cleaning and preparation and visualization of the results, and how to produce predictions in real-time, etc.
 - ❑ Tasks are of central importance and when combined can account for more than 60% of all KDD time spent
- ❖ Includes a single step called “Data Mining” which is the application of Machine Learning methods used in the mining of data
 - ❑ Helping to distinguish between random noise and significant findings, and providing a theory for estimating probabilities of predictions, etc.



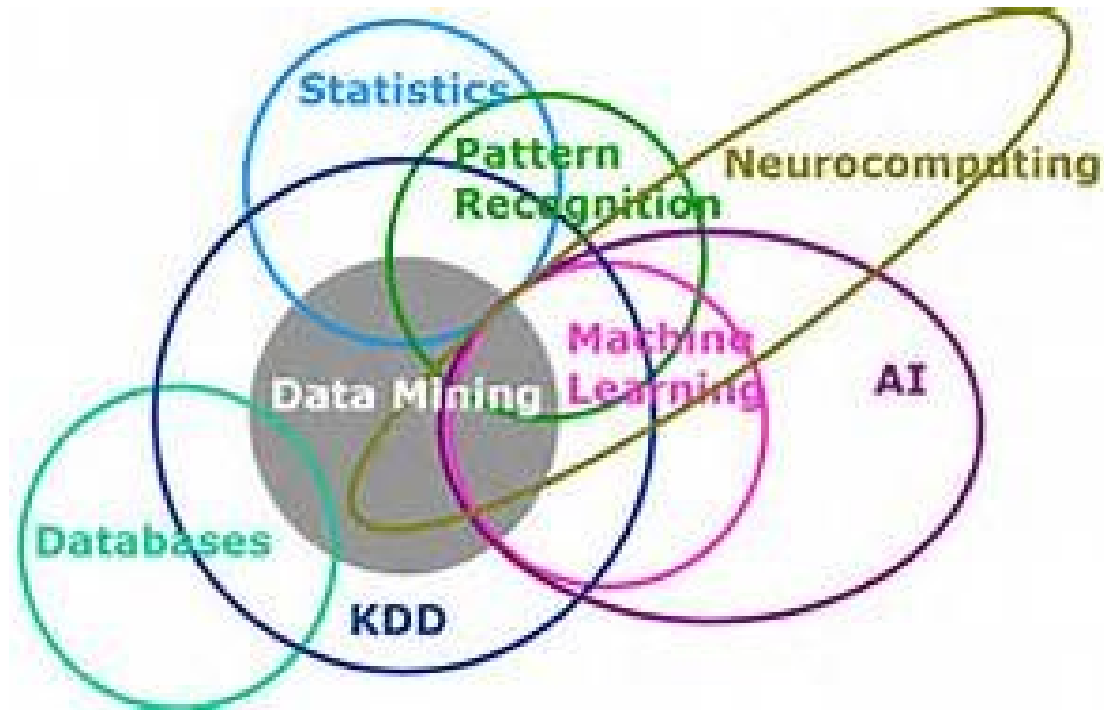
“Data scientists spend 50-80% of their time in data collection and preparation”





Terminology

- ❖ The concepts and terminology are overlapping and seemingly repetitive at times



- ❖ The KDD Process was actually intended to describe a Knowledge Discovery in Databases process,
 - ☐ which kind of eventually became **Data Mining**
 - ☐ which kind of eventually became **Data Science**



Παράδειγμα 3: Λήψη απόφασης

❖ Να γίνει μάθηση:

- ☐ Κατηγοριών υψηλού κινδύνου για καισαρική τομή για μελλοντικούς ασθενείς

❖ Δίνονται:

- ☐ 9714 εγγραφές ασθενών που περιγράφουν μια εγκυμοσύνη και γέννα

- ☐ Κάθε εγγραφή αποτελείται από 215 χαρακτηριστικά, π.χ.:

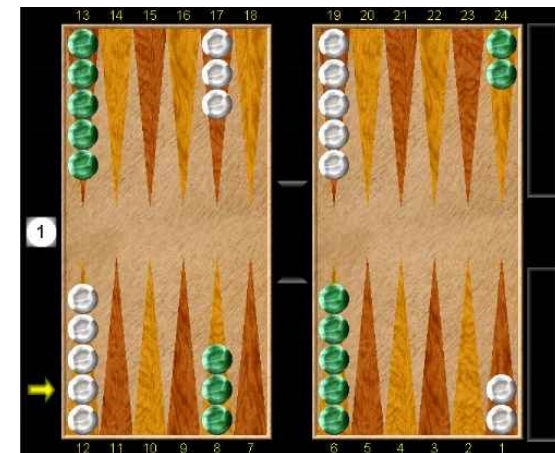
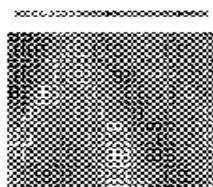
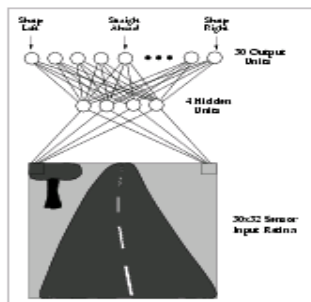
- ✓ Ηλικία: 32
- ✓ Πρώτη εγκυμοσύνη: Όχι
- ✓ Αναιμία: Όχι
- ✓ Διαβήτης: Όχι
- ✓ Προηγούμενη πρόωρη γέννα: Όχι
- ✓ Υπέρηχος: μη κανονικός
- ✓ ...
- ✓ Επείγουσα κατάσταση καισαρικής: Ναί
- ✓

❖ Παράδειγμα κανόνα:

- ☐ **Εάν** ο υπέρηχος είναι μη κανονικός **και** δεν υπάρχει προϊστορία κανονικής γέννας **τότε** η πιθανότητα επείγουσας κατάστασης καισαρικής είναι 0.6



Παράδειγμα 4: Σύνθετα προγράμματα





Παράδειγμα 5: Outlier (or Anomaly) Detection. Examples of Real World

Anomalies translate to significant (often critical) real life entities, e.g.

❖ Network /Cyber Intrusions

- ❑ A web server involved in ftp traffic



❖ Credit Card Fraud

- ❑ An abnormally high purchase made on a credit card



❖ Healthcare Informatics / Medical diagnostics

- ❑ Detect anomalous patient records



❖ Industrial Damage Detection

- ❑ Detection of faults and failures in complex industrial systems, structural damages, intrusions in electronic security systems, abnormal energy consumption, etc.

❖ Image Processing / Video surveillance

❖ Novel Topic Detection in Text Mining

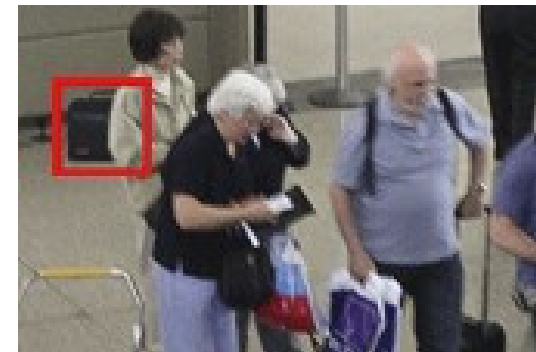
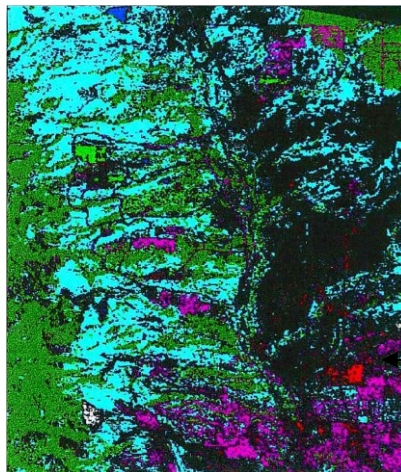
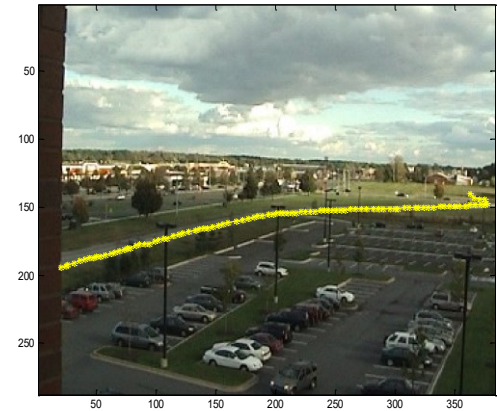
❖ ...





Παράδειγμα 6: Image Processing

- ❖ Detecting anomalous regions within an image
- ❖ Used in
 - ☐ mammogram image analysis
 - ☐ video surveillance
 - ☐ satellite image analysis
- ❖ Key Challenges
 - ☐ Detecting collective anomalies
 - ☐ Data sets are very large



Anomaly



Παράδειγμα 7: Games

- ❖ AlphaGo (της εταιρείας DeepMind του ομίλου Alphabet στον οποίο ανήκει και η Google)
 - ❑ Το **AlphaGo** χρησιμοποιούσε ένα νευρωνικό δίκτυο το οποίο εκπαιδεύτηκε παίζοντας με εκατοντάδες επαγγελματίες αλλά και ερασιτέχνες του συγκεκριμένου παιχνιδιού και τελικά κατάφερε το 2017 να νικήσει τον Ke Jie, νούμερο 1 στο συγκεκριμένο παίγνιο.
 - ❑ Στη συνέχεια ανέπτυξε το **AlphaGo Zero** το οποίο εκπαιδεύτηκε παίζοντας με τον εαυτό του για 8 ώρες και αργότερα το **AlphaZero** το οποίο εκπαιδεύτηκε για να παίζει εκτός του παιχνιδιού Go, και σκάκι και Shōgi υπερνικώντας τα μέχρι τώρα αντίστοιχα λογισμικά (Stockfish και Elmo).





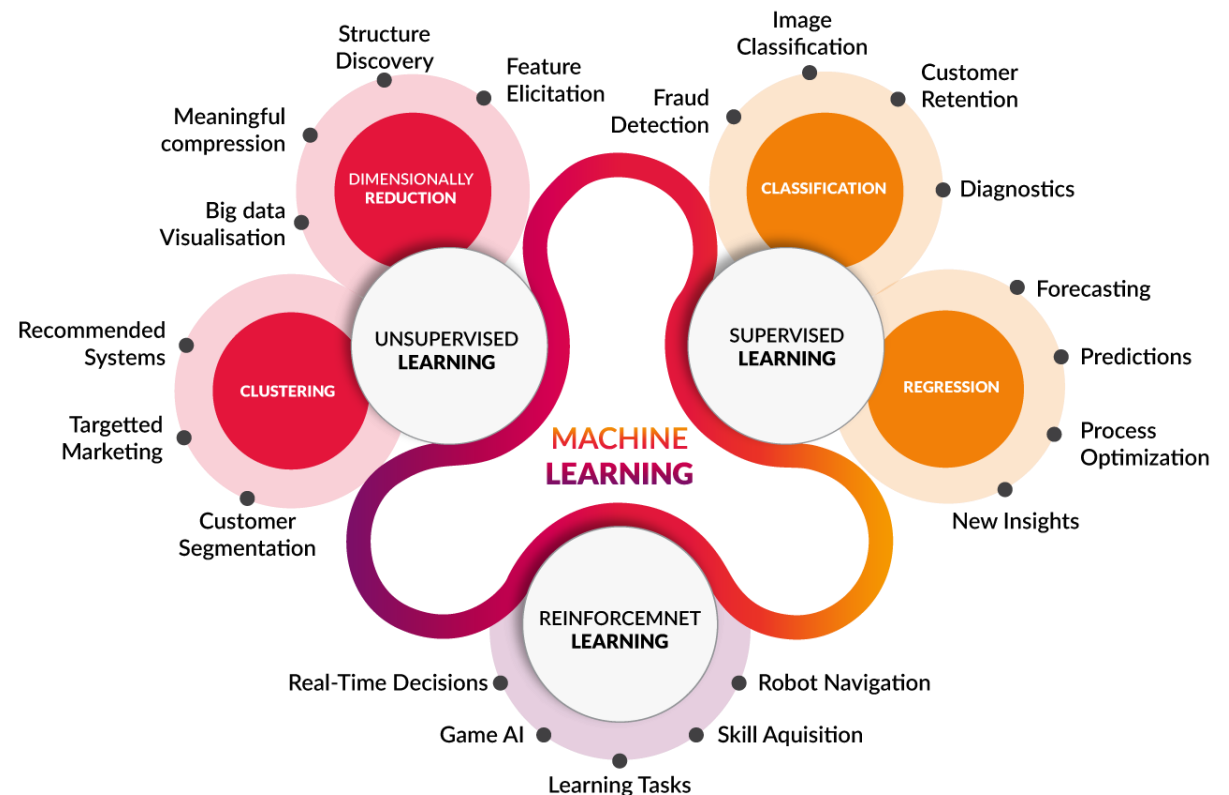
Καλά προβλήματα για Μηχανική Μάθηση

- ❖ Δεν υπάρχει άνθρωπος ειδικός
 - ❑ π.χ. Ανάλυση DNA
- ❖ Οι άνθρωποι μπορούν να εκτελέσουν την εργασία αλλά όχι να εξηγήσουν πως, ώστε να μπορέσουμε να χρησιμοποιήσουμε τη γνώση τους για να γράψουμε προγράμματα.
 - ❑ π.χ. Αναγνώριση χαρακτήρων
- ❖ Συχνή αλλαγή συμπεριφοράς
 - ❑ π.χ. Πρόβλεψη τιμών μετοχών με βάση πρόσφατα δεδομένα
- ❖ Κάθε χρήστης χρειάζεται μια εξειδικευμένη λειτουργία
 - ❑ π.χ. Φιλτράρισμα ειδήσεων



Κατηγορίες μεθόδων Μηχανικής Μάθησης

- ❖ Έχουν αναπτυχθεί πολλές τεχνικές μηχανικής μάθησης οι οποίες χρησιμοποιούνται ανάλογα με τη φύση του προβλήματος και εμπίπτουν σε ένα από τα παρακάτω τρία είδη:
 - ❑ **Μάθηση με επίβλεψη** (Supervised Learning) ή μάθηση με παραδείγματα (learning from examples)
 - ❑ **Μάθηση χωρίς επίβλεψη** (Unsupervised Learning) ή μάθηση από παρατήρηση (learning from observation)
 - ❑ **Ενισχυτική μάθηση** (Reinforcement Learning)



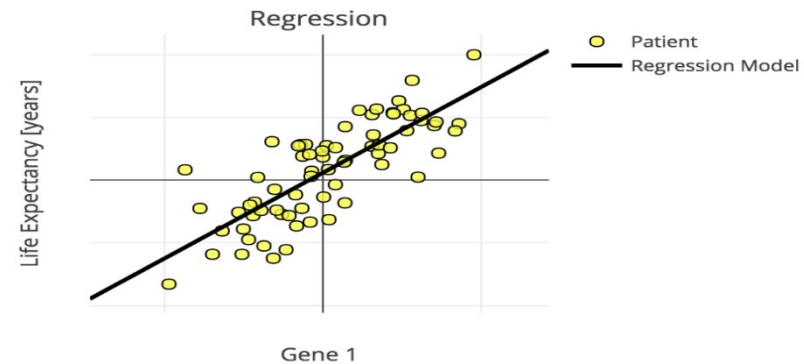
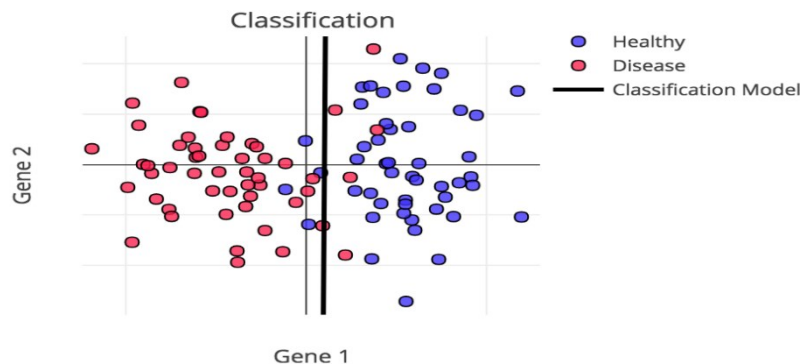


Κατηγορίες μεθόδων Μηχανικής Μάθησης



Supervised

- ❑ The computer is presented with example inputs and their desired outputs, given by a "teacher", and the goal is to learn a general rule (model) that maps inputs to outputs.
- ❑ Inputs are divided into two (binary) or more classes (multiclass), and the learner must produce a **model** that assigns unseen inputs to one or more (multi-label classification) of these classes
- ❑ Model is used for **prediction**
- ❑ In commercial use, this is known as **predictive analytics**
- ❑ According to the desired output we distinguish two types of supervised learning:
 - ✓ **Classification**, when the outputs are discrete
 - ✓ **Regression**, when the outputs are continuous

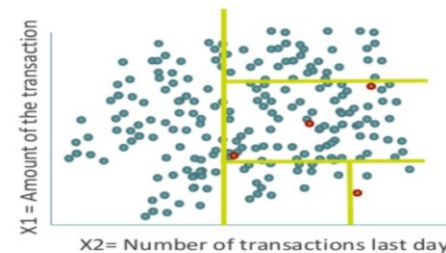
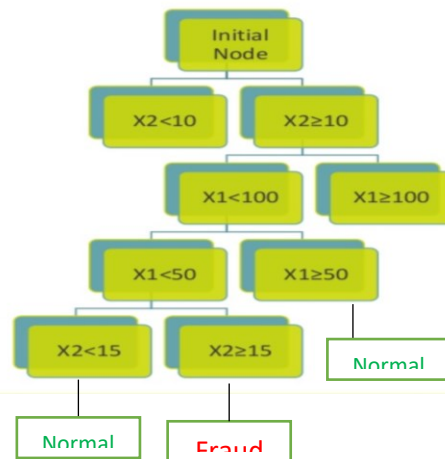




❖ Example applications include

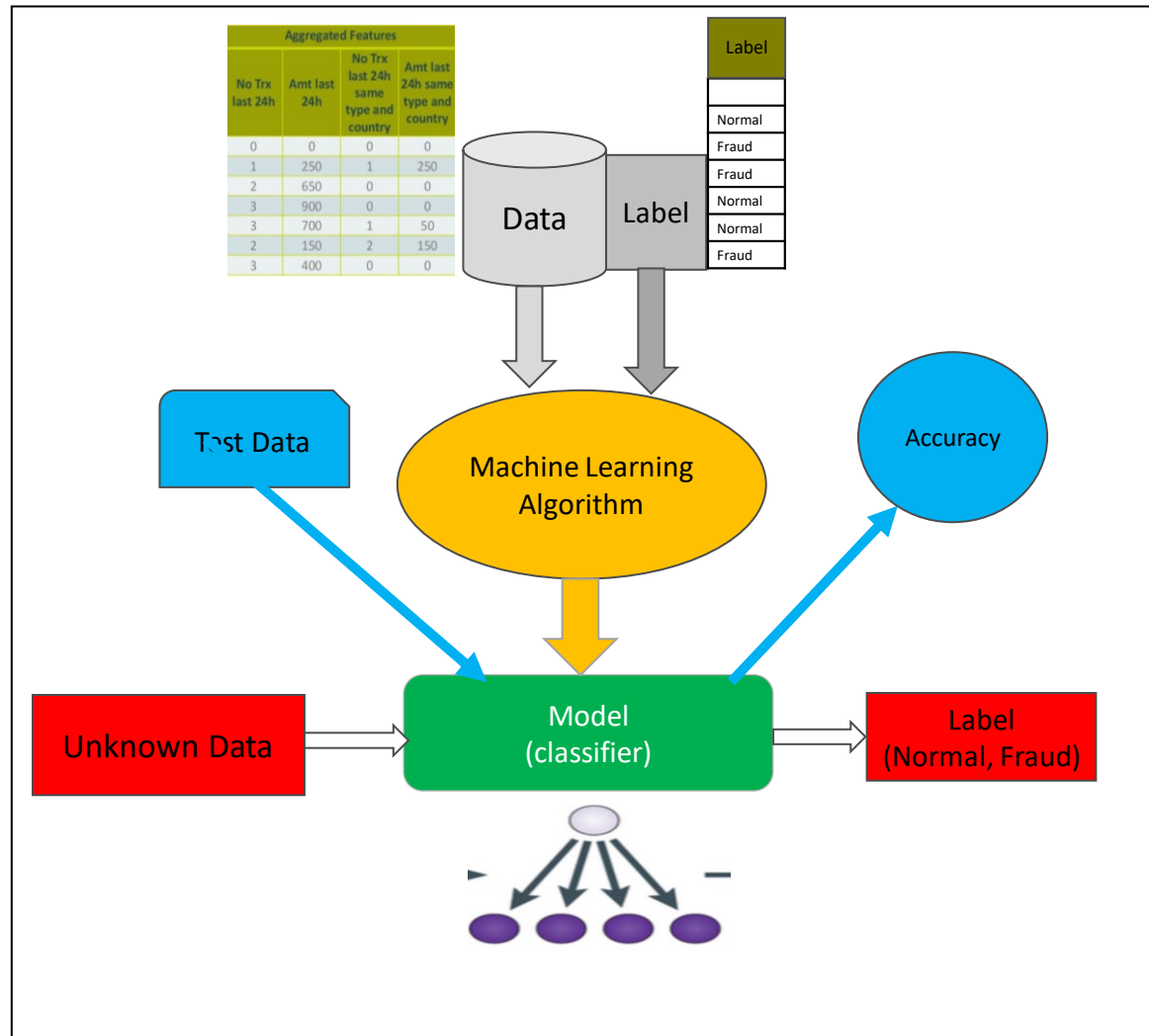
- ☐ email filtering
- ☐ detection of network intruders or malicious insiders
- ☐ optical character recognition (OCR)
- ☐ recommender systems
- ☐ sentiment analysis
- ☐ online advertising
- ☐ computer vision

❖ **Algorithms:** Decision Trees, Nearest Neighbor, Rule Learning, Support Vector Machines, Neural Networks,





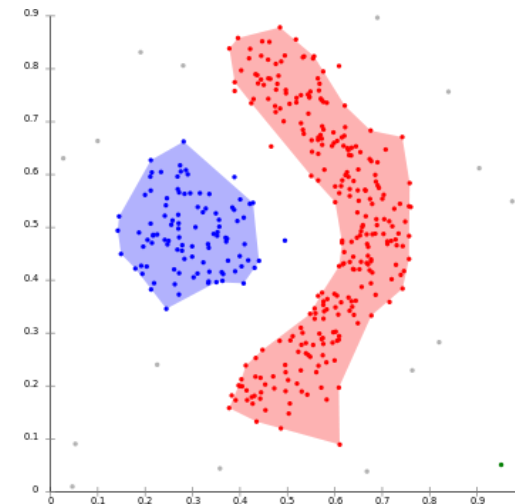
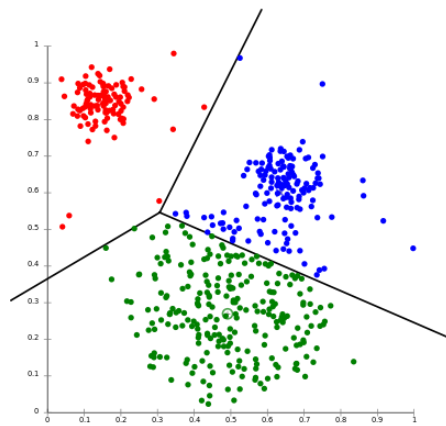
How supervised learning works?





❖ Unsupervised learning

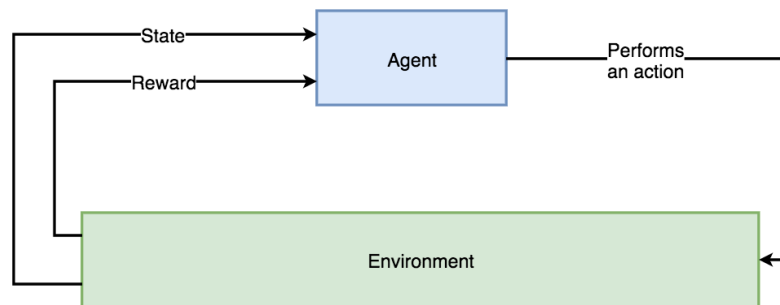
- ☐ No labels are given to the learning algorithm, leaving it on its own to find hidden patterns in data
- ☐ In clustering, a set of inputs is to be divided into groups.
- ☐ Unlike in classification, the groups are not known beforehand, making this typically an unsupervised task
- ☐ **Algorithms:** K-means, EM, DBSCAN, CURE, ...



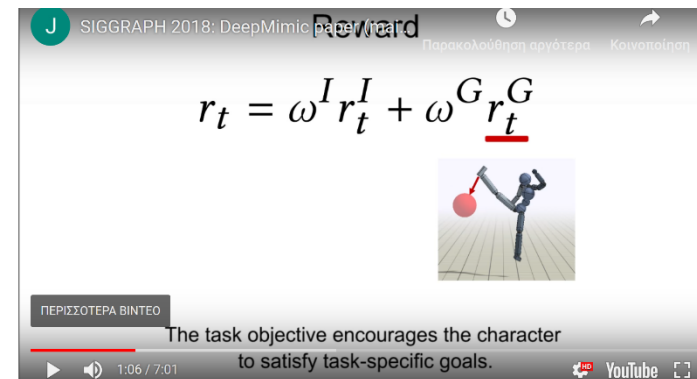


❖ Reinforcement Learning

- ❑ A computer program interacts with a dynamic environment in which it must perform a certain goal (such as driving a vehicle or playing a game against an opponent).
- ❑ The program is provided feedback in terms of rewards and punishments as it navigates its problem space.
- ❑ **Algorithms:** Q-Learning, SARSA, R-Learning

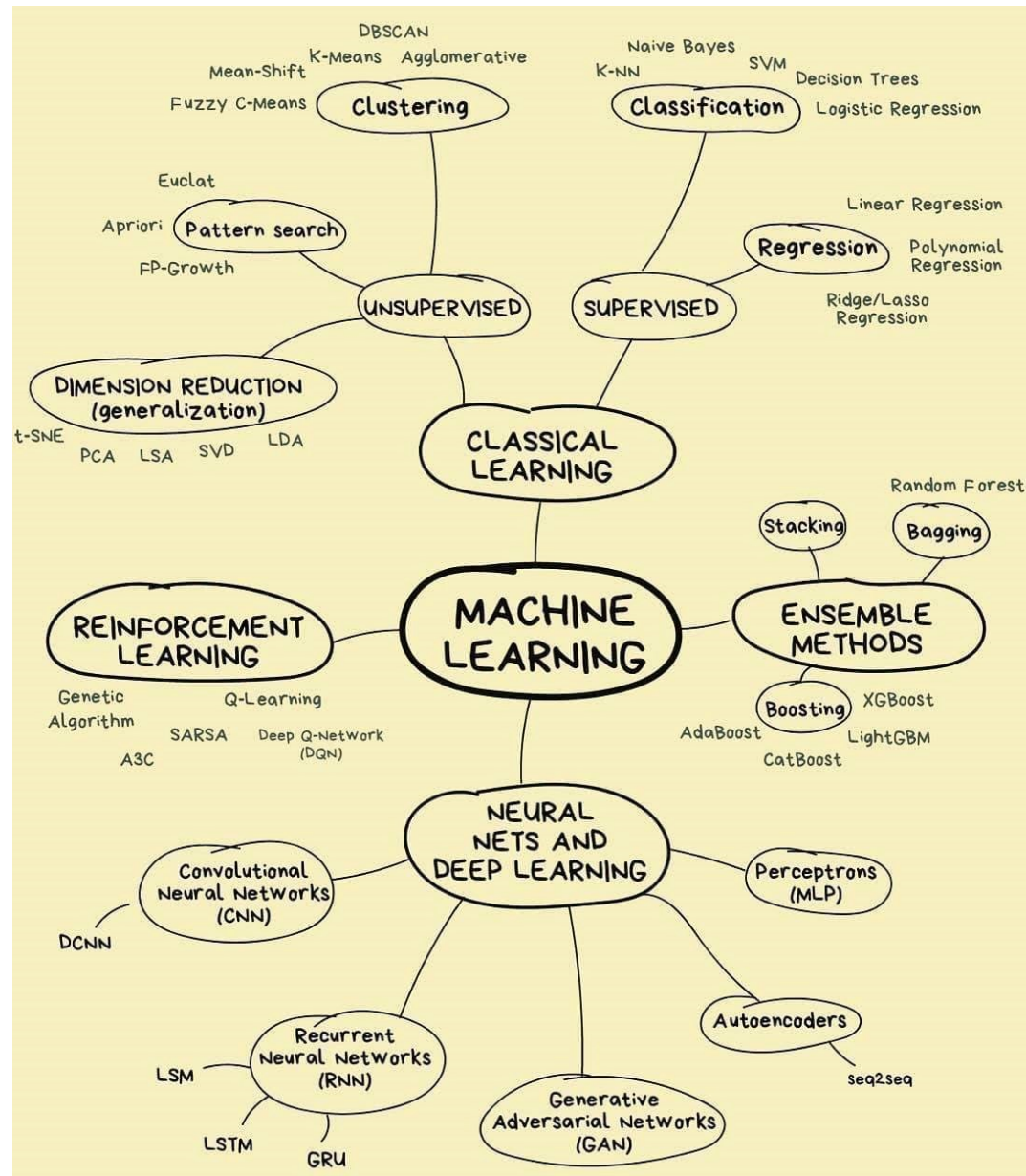


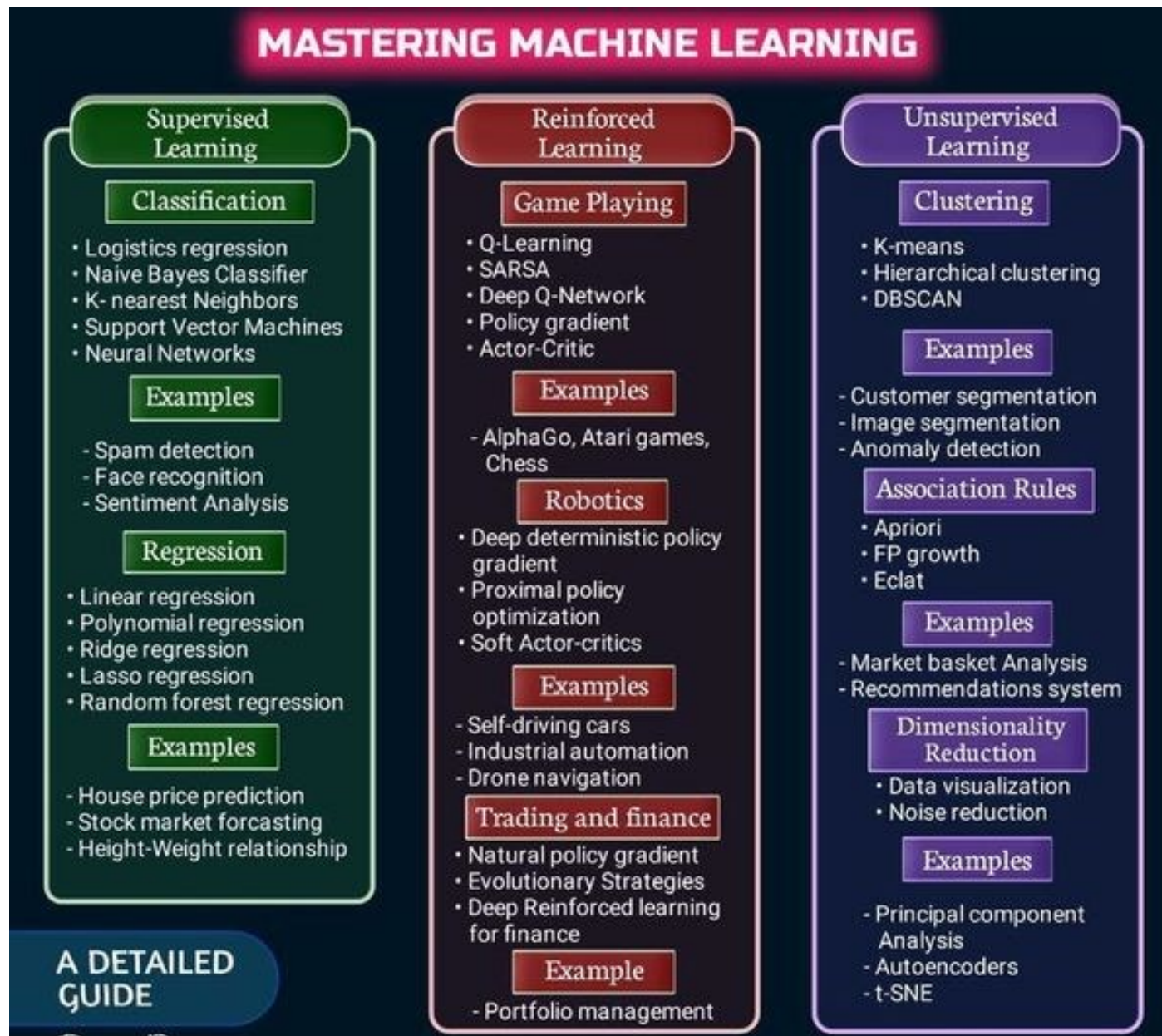
1





Κατηγορίες μεθόδων Μηχανικής Μάθησης







Τεχνητά Νευρωνικά Δίκτυα

Artificial Neural Networks

❑ Έχουν προταθεί από το 1943

- ✓ Το όνομα και η δομή τους είναι εμπνευσμένα από τον ανθρώπινο εγκέφαλο
- ✓ Μιμούνται τον τρόπο που οι βιολογικοί νευρώνες δίνουν σήμα ο ένας στον άλλο

❑ Αποτελούνται από κόμβους (νευρώνες) οργανωμένους σε επίπεδα (Layers)

- ✓ Ένα επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα και ένα εξόδου

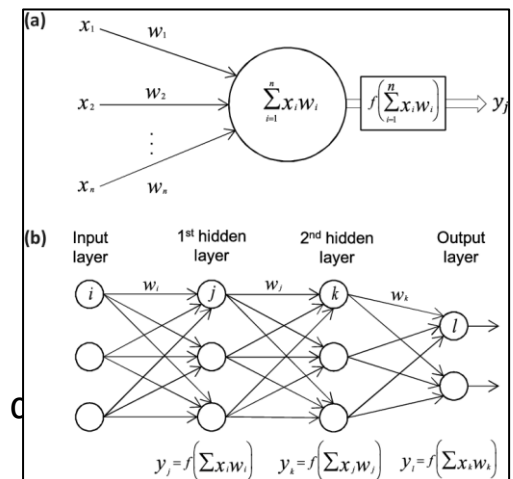
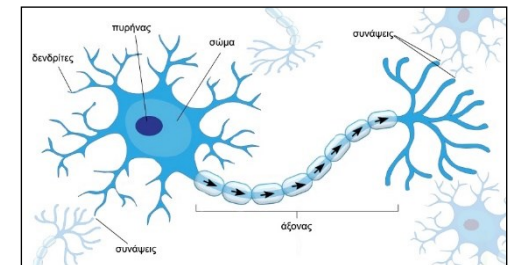
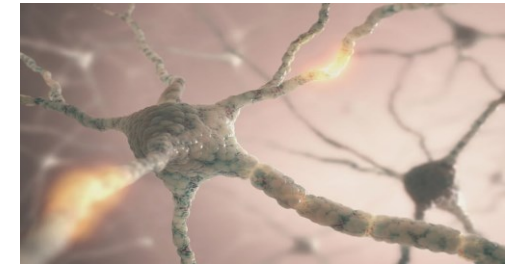
❑ Λειτουργία:

- ✓ Ένας νευρώνας (ή Perceptron) είναι μια μαθηματική συνάρτηση που απλώς
- ✓ λαμβάνει εισόδους από τους «δενδρίτες» (δηλ. από άλλους νευρώνες),
- ✓ τις πολλαπλασιάζει με μια τιμή (βάρος≈συνάψεις),
- ✓ τις προσθέτει,
- ✓ το άθροισμα τροφοδοτεί μια άλλη συνάρτηση (ενεργοποίησης) και
- ✓ τελικά παράγει έξοδο στον άξονα του που τροφοδοτεί άλλους νευρώνες

❑ Διαφορετικές αρχιτεκτονικές:

- ✓ Ο αριθμός των νευρώνων, των επιπέδων, οι συναρτήσεις ενεργοποίησης, κλπ

-
- ✓ Ο εγκέφαλος μας περιέχει περίπου 86 δισεκατομμύρια νευρώνες και περισσότερα από 1000 τρισεκατομμύρια συνάψεις (συνδέσεις)





Βαθιά Μάθηση - Deep Learning

- ❑ Πιθανώς ένα από τα πιο καυτά θέματα στη Μηχανική Μάθηση σήμερα

- ✓ Επιτυγχάνει αξιοσημείωτα αποτελέσματα

- ❑ Χρησιμοποιεί μεγάλα πολυεπίπεδα νευρωνικά δίκτυα

- ❑ Διαφορετικά είδη:

- ✓ Deep Feedforward Networks or Multilayer Perceptrons (MLPs)

- ✓ CNN (Convolutional NN) [1995]

- TCN (Temporal Convolutional Networks) [2016]

- ✓ RNN (Recurrent NN) [1986]

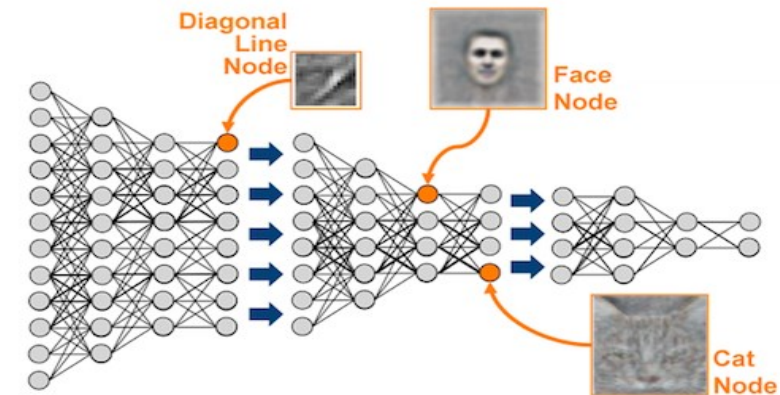
- LSTMs (Long Short-Term Memory Units) [1997]

- Gated RNN (GRU) [2014]

- Encoder and Decoder through an attention mechanism [2015]

- ✓ GAN (Generative Adversarial Networks) [2014]

- ✓ Transformer [2017]

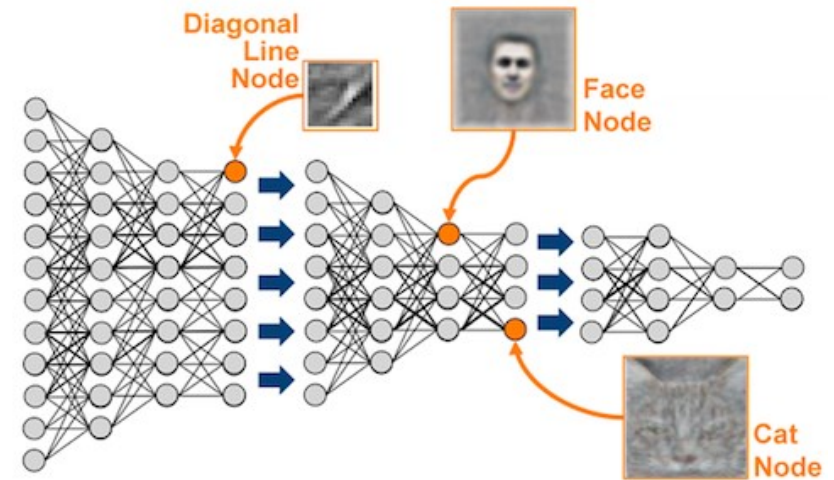


- ❑ Generative AI – Παραγωγική ή Γεννητική ΤΝ



Deep Learning

- ❖ Probably one of the hottest topics in Machine learning today
 - ❑ achieves remarkable outcomes
- ❖ Uses large multi-layered neural networks
- ❖ It falls under a class of unsupervised learning algorithms
- ❖ Applications
 - ❑ Image recognition (e.g. Tagging faces in photos)
 - ❑ Voice recognition (e.g. Voice based search, Siri)
 - ❑ Pattern detection (e.g. Handwriting recognition)
- ❖ Why DNN (and DL) is now in the foreground?
 - ❑ We now have fast enough computers and enough data to actually train large neural networks
 - ❑ Keras (Tensorflow, Theano) are python libraries that make writing deep learning models easy, and give the option of training them on a GPU
- ❖ While Deep Learning (DL) had many impressive successes, it is only a small part of ML, which is a small part of AI
- ❖ We argue that future AI should explore other ways beyond DL





Machine Learning (ML) Algorithms

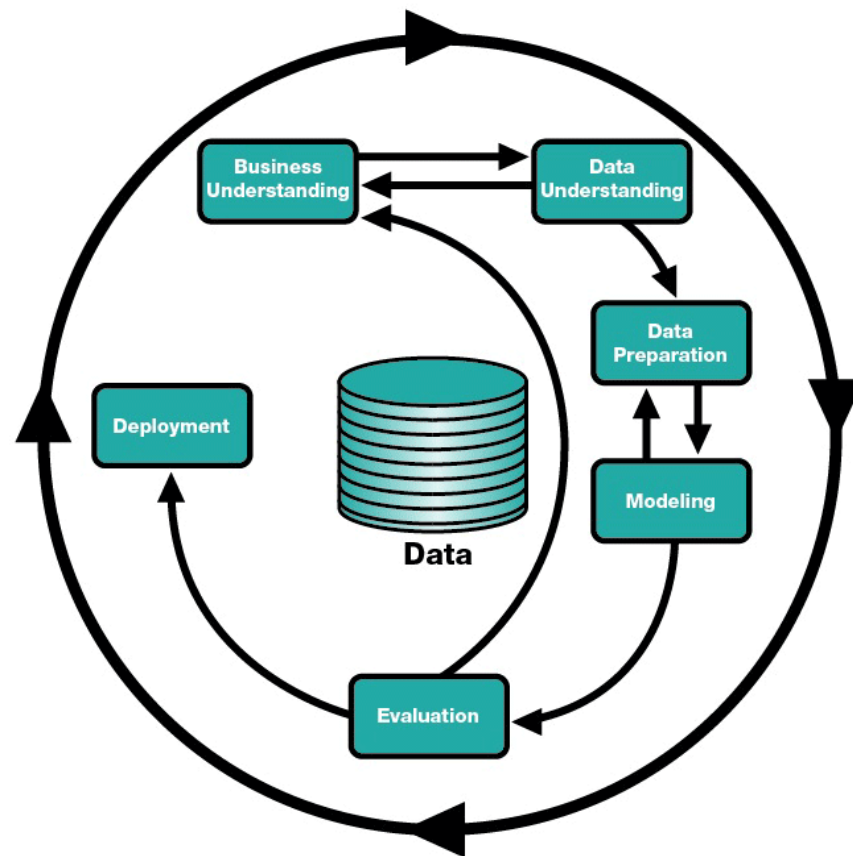
- ❖ Hundreds of ML algorithms available
 - ☐ The best one does not exist (No-free lunch theorem)
- ❖ Some have better performances under certain conditions
 - ☐ e.g. Several studies have reported that Random Forest is the most accurate for fraud detection
 - ☐ XGBoost (Netflix)
- ❖ Which Software should we use?
 - ☐ Weka, [RapidMiner](#), Keras, TensorFlow, Theano, PyTorch?
 - ☐ [Java](#), [Python](#), [Julia](#), [Scala](#)?
 - ☐ [R](#) used to be the standard between data scientists
 - ✓ Open source (Free) software and developed by academics
 - ☐ Python
 - ✓ combines the power of general-purpose programming languages with the ease of use of domain-specific scripting languages like MATLAB or R
 - ✓ creation of complex GUIs and web services, and for integration into existing systems
- ❖ Because finding patterns is hard, often not enough training data is available, and also because of the high expectations it often fails to deliver





Methodology

- ❖ CRISP-DM ([Cross Industry Standard Process for Data Mining](#)), is a data mining process model that describes commonly used approaches that data mining experts use to tackle problems.
- ❖ CRISP-DM remains the most popular methodology for analytics, data mining, and data science projects.





Limitations of ML

Limitations include:

- ❖ The need for massive data sets
 - ☐ in some cases, is extremely difficult (e.g clinical-trial data)
- ❖ Labeling training data
 - ☐ which often must be done manually and is necessary for supervised learning
- ❖ Difficulties in explaining results
 - ☐ Why was a certain decision reached?
- ❖ Generalizing learning
 - ☐ to handle similar use cases
- ❖ Potential (social/cultural) bias in data and algorithms
 - ☐ training data are not (always) representative of the larger population to which an AI model is applied



KNOWING YOUR TASK AND YOUR DATA

- ❖ The most important part in the machine learning process is understanding the data you are working with and how it relates to the task you want to solve
- ❖ Each algorithm is different in terms of what kind of data and what problem setting it works best for
- ❖ It will not be effective to randomly choose an algorithm and throw your data at it
 - ❑ It is necessary to understand what is going on in your dataset before you begin building a model
- ❖ The algorithms and methods in machine learning are only one part of a greater process to solve a particular problem, and it is good to keep the big picture in mind at all times
- ❖ Is machine learning all about choosing the best algorithm?
 - ❑ The fact that many people think that machine learning is just about choosing the right algorithm is why so many machine learning projects fail and so many organizations are wondering if it's the right thing to be investing in
 - ✓ More details [here](#)



How well do I need to know machine learning algorithms to be a data scientist in a tech company (1/2)

The 7 most common things required by big companies for data science related positions

☐ Basic Programming Languages

- ✓ You should know a statistical programming language, like R or Python (and Numpy, Pandas Libraries), and a database querying language like SQL

☐ Statistics

- ✓ You should be able to explain phrases like null hypothesis, P-value, maximum likelihood estimators and confidence intervals
- ✓ Statistics is important to crunch data and to pick out the most important figures out of a huge dataset

☐ Machine Learning

- ✓ You should be able to explain K-nearest neighbors, random forests, and ensemble methods (are implemented in R or Python)
- ✓ These algorithms show to employers that you have exposure to how data science can be used in more practical manners

☐ Data Wrangling

- ✓ You should be able to clean up data. i.e. understanding that "California" and "CA" are the same thing - a negative number cannot exist in a dataset with population.
- ✓ It is all about identifying corrupt (or impure) data and correcting/deleting them



How well do I need to know machine learning algorithms to be a data scientist in a tech company (2/2)

☐ Data Visualization

- ✓ Data scientist is useless on his or her own. They need to communicate their findings to Product Managers.
- ✓ Familiarity with data visualization tools like ggplot is very important (so you can SHOW data, not just talk about them)

☐ Software Engineering

- ✓ You should know algorithms and data structures (Queues, Arrays, Lists, Stacks, Trees, etc.), as they are often necessary in creating efficient algorithms for machine learning

☐ Product Management

- ✓ Those who understand the product are the ones who will know what metrics are the most important
- ✓ Product-oriented data scientist will pick the right metrics to experiment with
- ✓ Know what these terms mean: Usability Testing, Wireframing, Retention and Conversion Rates, Traffic Analysis, Customer Feedback, Internal Logs, [A/B Testing](#)



How do I get started with learning Machine Learning by myself?

- ❖ First, forget what most people tell you and do your own research
- ❖ Do a search here: Job Search | Indeed (<https://gr.indeed.com/>)
- ❖ Type in: Machine learning Engineer

- ❖ Two things you'll need to know really well are SQL and Python.
 - ☐ SQL is the number 1 skill for an applied machine learning engineer.
 - ☐ You'll also need to know the core machine learning libraries.
 - ☐ Pandas, SciKit-Learn, matplotlib, Seaborn, Keras
 - ☐ Applied machine learning is programming and data wrangling. So, focus on that.
- ❖ Here's a free Python course to get you started. You'll need know everything in this course and the follow up course really well.
 - ☐ [The Complete Python Course for Machine Learning Engineers](#)

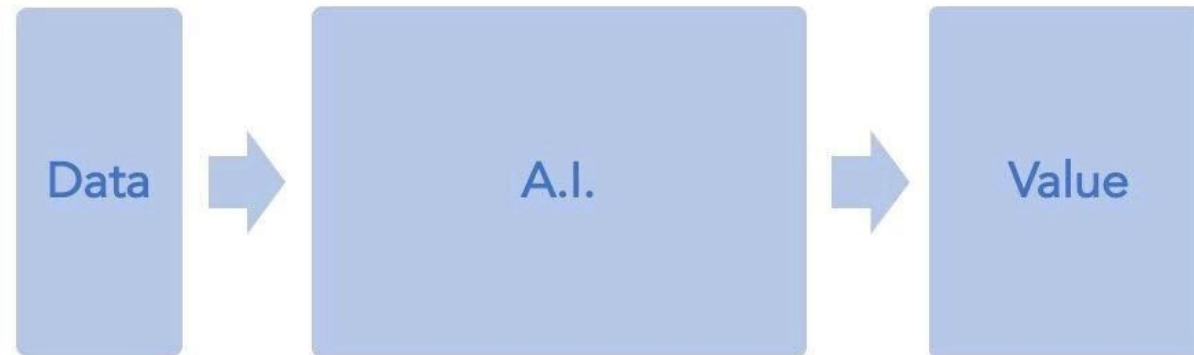


Brutal truth about machine learning

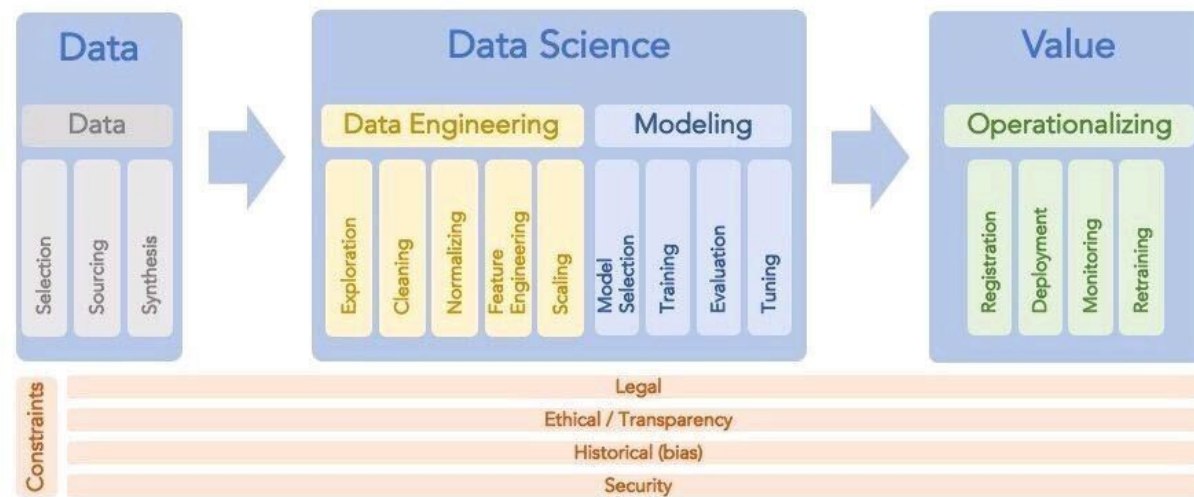
- ❖ Most business datasets are either too dirty or too sparse to make machine learning useful
- ❖ The actual use-cases for machine learning are few vs. more general business analytics
 - ❑ Many business problems do not require prediction
 - ❑ Inferential statistics and variance analysis will take care of 80–90% of insightful
- ❖ Businesses value interpretability over accuracy
 - ❑ This is why less efficient but more interpretable methods (e.g. logistic/linear regression) are still in use over the black box methods that can hit 95+%)
- ❖ Even with machine learning in place, business leaders remain skeptical of error rates
 - ❑ This is a cognitive bias; people value successes of human decisions and magnify the admittedly fewer failures of machine learning
 - ❑ Look at the media firestorm on the few self-driving car mishaps vs. the thousands to millions of people dying due to other causes
- ❖ The use-cases where machine learning does shine (e.g. object detection, text analysis, sound analysis, image analysis) require relatively more mature and advanced business culture and process. This rules out 90+% of typical businesses



What companies think A.I. looks like



What it actually is



© Andy Scherpenberg



AI in Big Companies

- ❖ Microsoft
 - ❑ Azure
 - ✓ It provides SaaS, PaaS and IaaS
 - ✓ Supports many programming languages, tools and frameworks, Microsoft and third-party
- ❖ Microsoft Cognitive Toolkit (CNTK)
 - ❑ is a deep learning framework ...to learn like the human brain
- ❖ Oracle AI: Ready to Go. Ready to Build. Ready to Work.
 - ❑ AI-powered cloud applications
 - ❑ makes it easy for enterprises to realize value from artificial intelligence and ML
- ❖ IBM, Watson
 - ❑ an IBM supercomputer that combines artificial intelligence (AI) and sophisticated analytical software for optimal performance as a "question answering" machine (posed in natural language)
 - ❑ the deep learning AI for business
- ❖ Amazon – aws
 - ❑investing deeply in artificial intelligence for over 20 years. Machine learning (ML) algorithms drive many of our internal systems



❖ Keras, Tensorflow

- ☐ an open source machine learning software library
- ☐ allows easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices
- ☐ TPU (Tensor Processing Unit) is an AI accelerator application-specific integrated circuit (ASIC) for machine learning algorithms (neural networks)



❖ Word2vec

- ☐ a group of related models (shallow, two-layer neural networks)
- ☐ vector representations of words
- ☐ takes as its input a large corpus of text and produces a vector space, typically of several hundred dimensions, with each unique word in the corpus being assigned a corresponding vector in the space.
- ☐ word vectors are positioned in the vector space such that words that share common contexts in the corpus are located in close proximity to one another in the space



Google AI and Games

❖ DeepMind (AI company, part of the Alphabet group which owns Google)



❖ AlphaGo

☐ computer program that plays the board game Go

✓ Monte Carlo tree search (MCTS), heuristic search algorithm improved by a neural network to predict next move and the eventual winner



☐ initially trained on thousands of human amateur and professional games to learn how to play

☐ at 2016 won Lee Sedol a 9-dan professional Go player

☐ at 2017 beat Ke Jie, the world No.1 ranked player at the time

❖ AlphaGo Zero

☐ using no data from human games but playing only games against itself for 8 hours

❖ AlphaZero

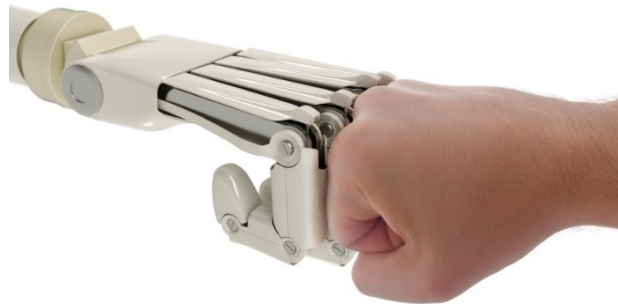
☐ could play chess and Shōgi in addition to Go

☐ Defeated a top chess program (Stockfish) and a top Shōgi program (Elmo)



Conclusion

- ❖ Η μηχανική μάθηση είναι ένα ισχυρό εργαλείο της ΤΝ που μας επιτρέπει να αυτοματοποιήσουμε, να επιταχύνουμε και να επεκτείνουμε τους τρόπους παραγωγής γνώσης, καθώς και να φτιάξουμε συστήματα που εκτελούν πολύπλοκες διεργασίες χωρίς να τα προγραμματίσουμε ρητά, ώστε να λύσουμε δύσκολα προβλήματα και να δώσουμε νέα ώθηση στην εξέλιξη, μετασχηματίζοντας για ακόμη μία φορά (4η) την κοινωνία.
- ❖ The solution is to combine machine-learning algorithms with data collected by human analysts
- ❖ People can still play a role validating the results





Περιεχόμενα μαθήματος και βιβλιογραφία (1/2)

https://docs.google.com/document/d/1l-qWx_ukk9JkKL5XaNrTFthkBUHZNCduZH-JqSZU5hE/edit?usp=sharing

1	Εισαγωγή
2	Μάθηση Συνάρτησης (Μάθηση με επίβλεψη)
3	Δένδρα Απόφασης
4	Μάθηση Κανόνων / Επαγωγικός Λογικός Προγραμματισμός
5	Μάθηση Βασισμένη σε Περιπτώσεις
6	Μάθηση κατά Bayes
7	Νευρωνικά Δίκτυα (NN)
8	Μηχανές Υποστήριξης Διανυσμάτων (SVM)
9	Μάθηση με Γενετικούς Αλγόριθμους (GA)
10	Αξιολόγηση Μοντέλων
11	Επιλογή χαρακτηριστικών, Διακριτοποίηση
12	Συνδυασμός Μοντέλων
13	Ομαδοποίηση
14	Κανόνες Συσχέτισης
15	Ενισχυτική Μάθηση



Πληροφορίες για το μάθημα

- ❖ Τα slides των παρουσιάσεων, πρόσθετη βιβλιογραφία, χρήσιμοι σύνδεσμοι, περιοδικά και συνέδρια της περιοχής, και άλλες πληροφορίες για το μάθημα βρίσκονται στην Ιστοσελίδα του μαθήματος στο **elearning**









Questions?





Top Machine Learning Algorithms for Predictions

Name	Type	Description	Advantages	Disadvantages
Linear Regression		-The best fit line through all data points	-Easy to understand -you can clearly see what the biggest drivers of the model are.	-sometimes too simple to capture complex relationships between variables, -Tendency for the model to overfit.
Logistic Regression		-The adoption for linear regression to problems of classification	-Easy to understand	-sometimes too simple to capture complex relationships between variables, -Tendency for the model to overfit.
Decision Tree		-A graph that uses branching method to match all possible outcomes of a decision	-Easy to understand and implement.	-Not often used for prediction because it's also often too simple and not powerful enough for complex data.
Random Forest		- Takes the average of many decision trees. Each tree is weaker than the full decision tree, but combining them we get better overall performance.	-A sort of „wisdom of the crowd“, Tend to result in very high quality results. -Fast to train	-Can be slow to output predictions relative to other algorithms. -Not easy to understand predictions.
Gradient Boosting		-Uses even weaker decision trees that increasingly focused on „hard examples“	-High-performing	-A small change in the future set or training set can create radical changes in the model. -Not easy to understand predictions.
Neural Networks		-Mimics the behaviour of the brain. NNs are interconnected Neurons that pass messages to each other. Deep Learning uses several layers of NNs to put one after the other.	-Can handle extremely complex tasks. No other algorithm comes close in image recognition.	-very very slow to train. Because they have so many layers. Require a lot of power. -Almost impossible to understand predictions.