Μηχανική Μάθηση

Εργαστήριο Τεχνητής Νοημοσύνης και Συστημάτων Μάθησης

AI|LS
ARTIFICIAL INTELLIGENCE
& LEARNING SYSTEMS LABORATORY

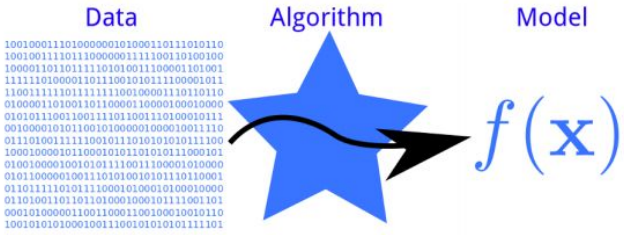
Τί είναι μια γάτα;

Τί είναι μια γάτα;

Πώς μαθαίνουμε τί είναι μια γάτα;

Ορισμοί της Μηχανικής Μάθησης

The subfield of computer science that "gives computers the ability to learn without being explicitly programmed".
(Arthur Samuel, 1959)

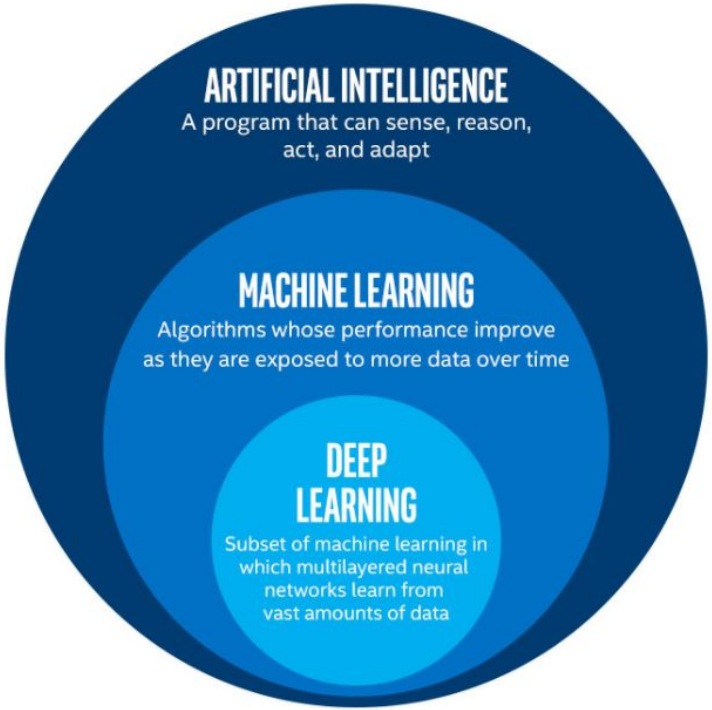


Using data for answering questions

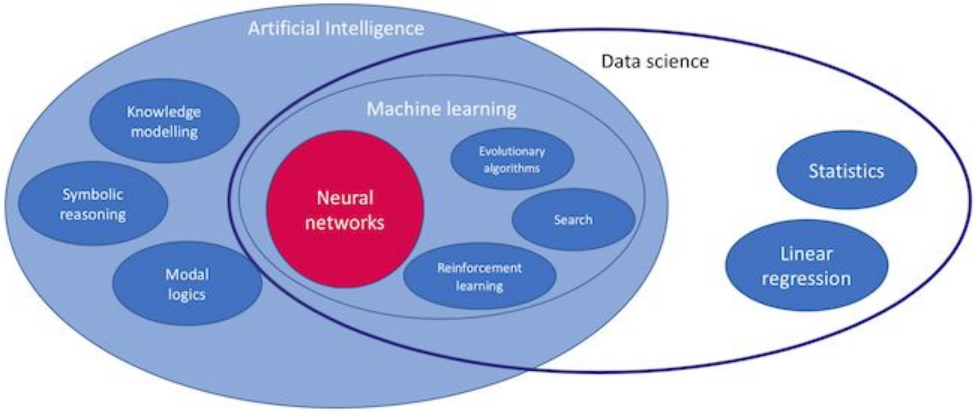
Training Predicting

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E."
(Tom Mitchell, 1997)

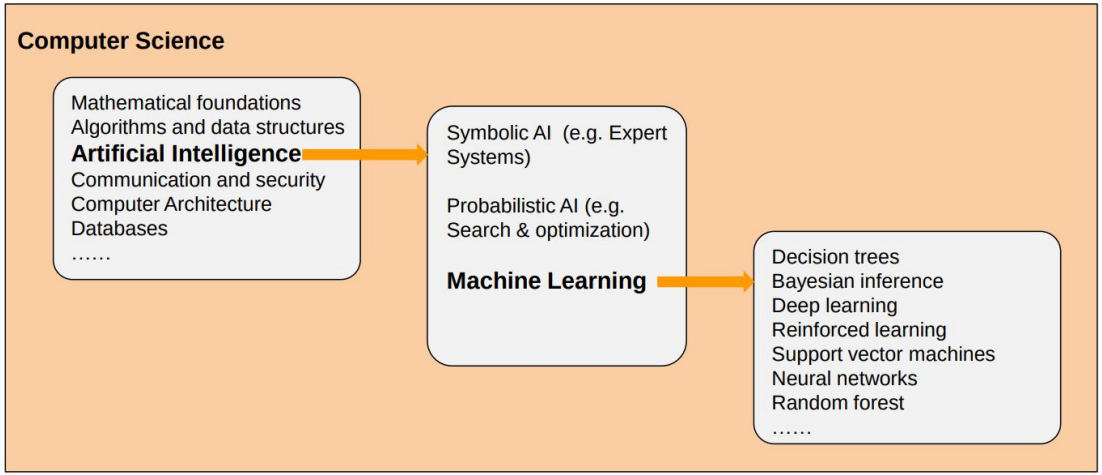
Ευφυή υπολογιστικά συστήματα



Ευφυή υπολογιστικά συστήματα



Επιστήμη Υπολογιστών, AI και Μηχανική Μάθηση



Artificial Intelligence (AI) is a branch or Computer Science that uses algorithms and techniques to mimic human intelligence

Machine Learning (ML) is one of several AI techniques for sophisticated cognitive tasks

A Paradigm Shift from Traditional AI

Traditional AI techniques

- **Static** – hard-coded set of steps and scenarios
- **Rule Based** – expert knowledge
- **No generalization** – handling special cases is difficult

---

Machine Learning

- **Dynamic** – evolves with data, finds new patterns
- **Data driven** – discovers knowledge
- **Generalization** – adapts to new situations and special cases

Στατιστική - AI - Μηχανικής Μάθηση

Symbolic AI

"Let us sit down with the world's best chess player, Ekpe Okorafor, and put his knowledge into a computer program"

Mathematical/Statistical AI

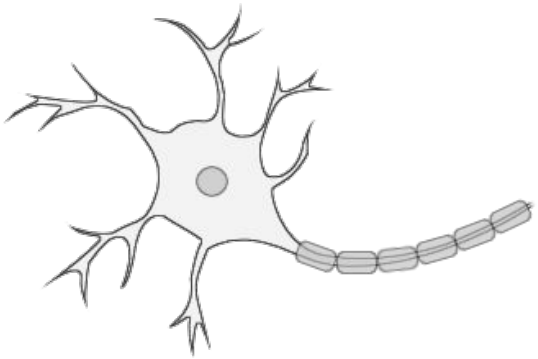
"Let us simulate all the different possible moves and the associated outcomes at each single step and go with the most likely to win"

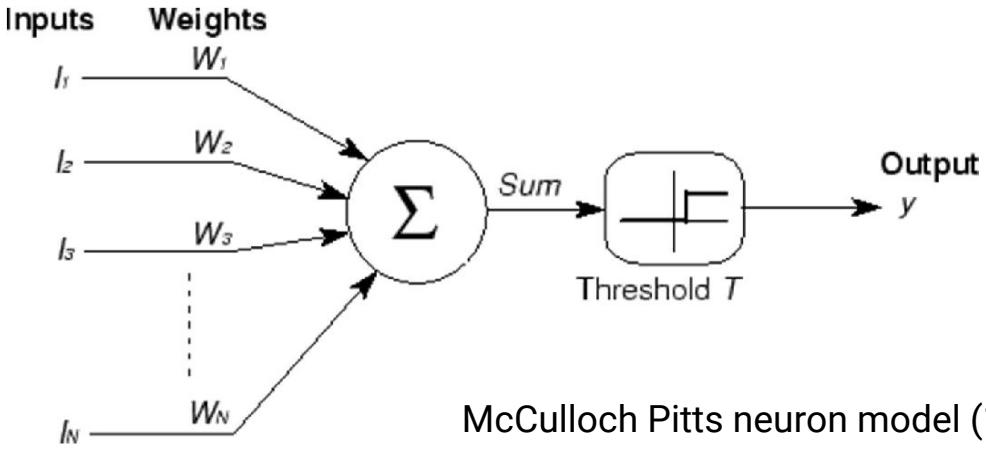
Machine Learning Approach

"Let us show millions of examples or real life and simulated games (won and lost) to the program, and let it learn from experience"

Βιομιμητισμός (Biomimicry)

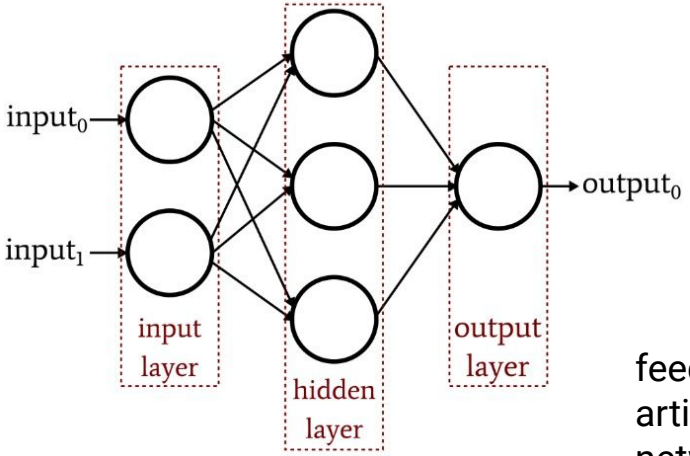
Βιομιμητισμός και μάθηση ή πώς η Μηχανική Μάθηση διαφοροποιείται από τη Στατιστική





McCulloch Pitts neuron model (1943)

Perceptron (1958) - Multilayer Perceptron (80s)



feedforward artificial neural network (ANN).

Perceptron Training Rule

- Weights modified for each example
- Update Rule:

$$w_i \leftarrow w_i + \Delta w_i$$

where

$$\Delta w_i = \eta (t - o) x_i$$

learning rate — target value — perceptron output — input value

Πότε χρησιμοποιούμε Μηχανική Μάθηση

Tasks programmers can’t describe

Complex multidimensional problems that can’t be solved by numerical reasoning

Hand writing

Cognitive Reasoning

Weather Forecasting

Network Intrusion

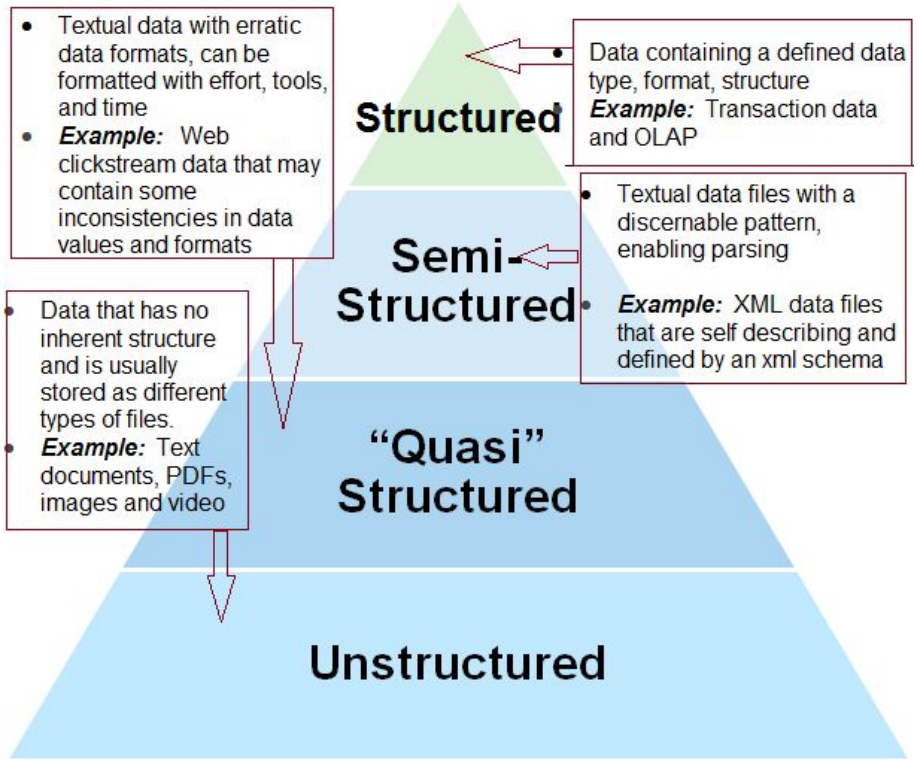
Health Care Outcomes

Movie Recommendation

Δομημένα και αδόμητα δεδομένα

	Structured Data	Unstructured Data
Characteristics	• Pre-defined data models • Usually text only • Easy to search	• No pre-defined data model • May be text, images, sound, video or other formats • Difficult to search
Resides in	• Relational databases • Data warehouses	• Applications • NoSQL databases • Data warehouses • Data lakes
Generated by	Humans or machines	Humans or machines
Typical applications	• Airline reservation systems • Inventory control • CRM systems • ERP systems	• Word processing • Presentation software • Email clients • Tools for viewing or editing media
Examples	• Dates • Phone numbers • Social security numbers • Credit card numbers • Customer names • Addresses • Product names and numbers • Transaction information	• Text files • Reports • Email messages • Audio files • Video files • Images • Surveillance imagery

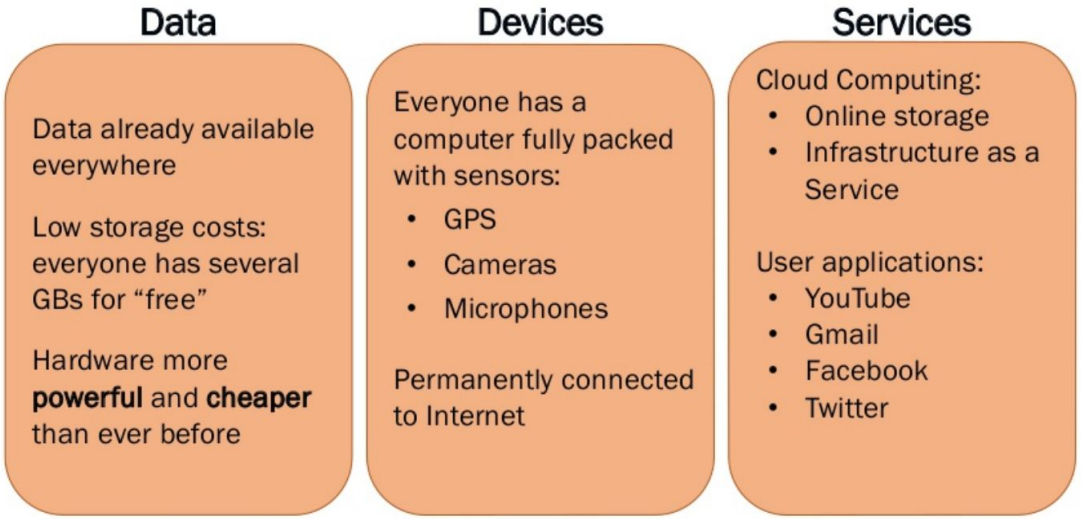
Δομημένα και αδόμητα δεδομένα



Τα δομημένα δεδομένα είναι ποιοτικά, "ακριβά" και λίγα. Κατάλληλα για κλασική ΤΝ, άμεση χρήση πχ σε κανόνες.

Τα αδόμητα δεδομένα είναι χαμηλής ποιότητας, "φθηνα" και πάρα πολλά. Χρήση Μηχανικής Μάθησης, έμμεση χρήση: πρέπει να μάθουμε από τα δεδομένα (Learning from Data)

Μηχανική Μάθηση, Big Data, Hardware, Software



Είδη της Μηχανικής Μάθησης

Είδη της Μηχανικής Μάθησης

Supervised

Unsupervised

Reinforcement

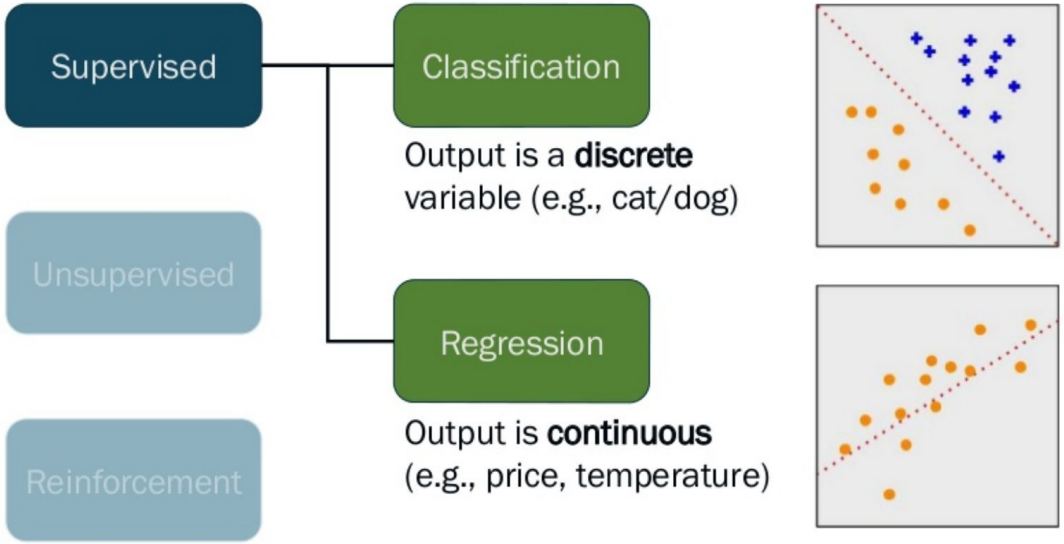
Learn through **examples** of which we know the desired output (what we want to predict).

Is this a cat or a dog?

Are these emails spam or not?

Predict the market value of houses, given the square meters, number of rooms, neighborhood, etc.

Είδη της Μηχανικής Μάθησης



Είδη της Μηχανικής Μάθησης

Supervised

Unsupervised

Reinforcement

There is **no *desired* output**. Learn something about the data. *Latent* relationships.

I have photos and want to put them in 20 groups.

I want to find anomalies in the credit card usage patterns of my customers.

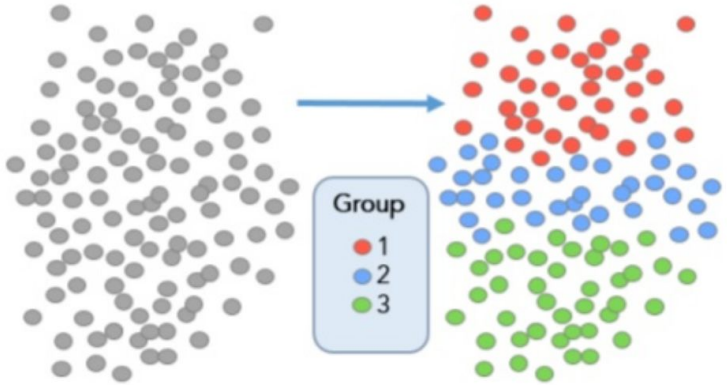
Είδη της Μηχανικής Μάθησης

Supervised

Unsupervised

Reinforcement

Useful for learning structure in the data **(clustering)**, hidden correlations, reduce dimensionality, etc.



Παραδείγματα αλγορίθμων ΜΜ

Supervised

- Linear classifier
- Naive Bayes
- Support Vector Machines (SVM)
- Decision Tree
- Random Forests
- k-Nearest Neighbors
- **Neural Networks (Deep learning)**

Unsupervised

- PCA
- t-SNE
- k-means
- DBSCAN

Reinforcement

- SARSA–λ
- Q-Learning

Είδη της Μηχανικής Μάθησης

Supervised

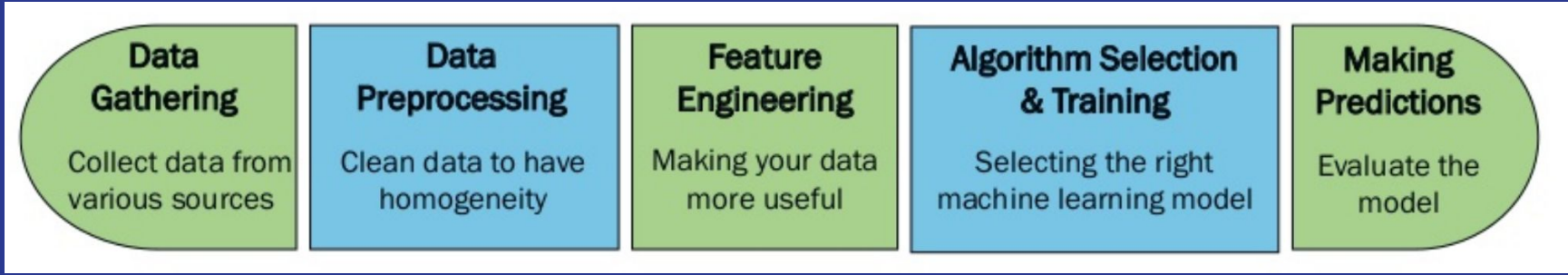
Unsupervised

Reinforcement

An agent **interacts** with an **environment** and watches the result of the interaction.

Environment gives feedback via a positive or negative **reward signal**.

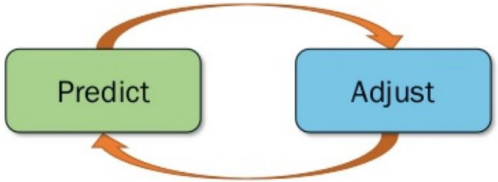
Επίλυση ενός προβλήματος ΜΜ



Επιλογή Αλγορίθμου και Εκπαίδευση

Goal of training: making the correct prediction as often as possible

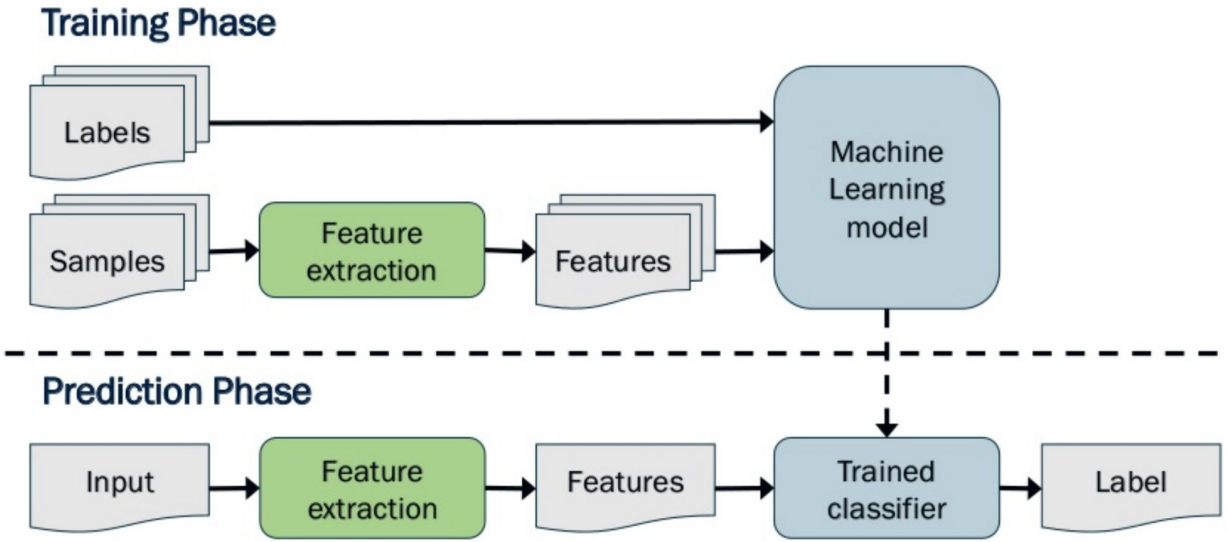
- Incremental improvement:



- Use of metrics for **evaluating** performance and comparing solutions
- **Hyperparameter tuning**: more an art than a science

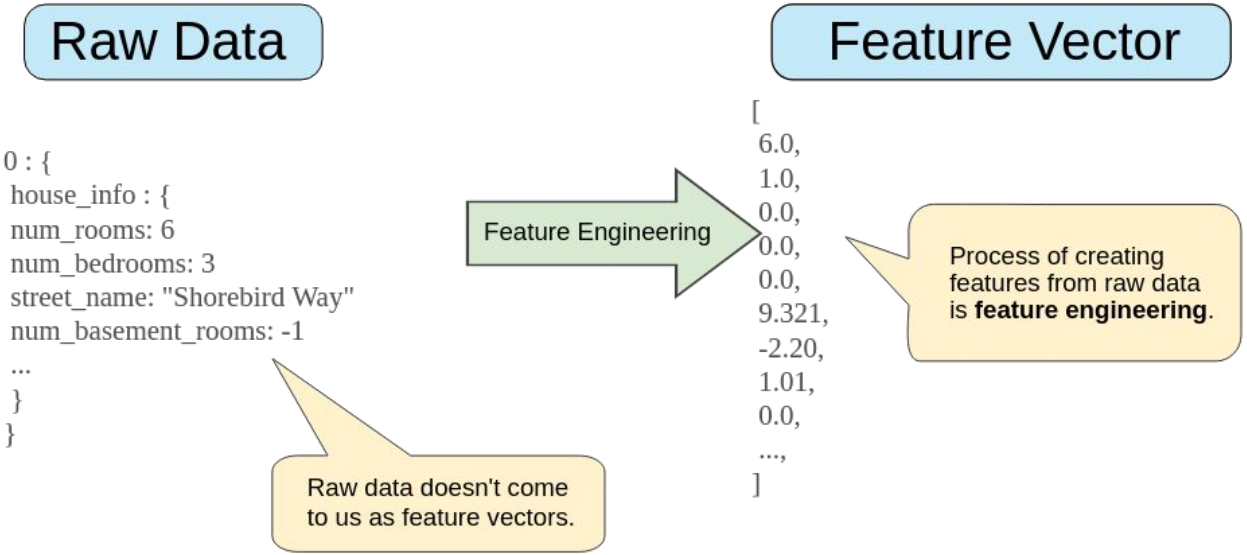
- Every ML algorithm has three components:
  - **Representation**
  - **Optimization**
  - **Evaluation**

Πρόβλεψη - Inference



Representation & Features

(αναπαράσταση και χαρακτηριστικά)



Τα δεδομένα

- Για να εφαρμόσουμε μεθόδους Μηχανικής Μάθησης χρειαζόμαστε έναν ελάχιστο όγκο δεδομένων, που αποκαλούμε σύνολο δεδομένων (data set)
- Το σύνολο δεδομένων αποτελείται από έναν αριθμό δειγμάτων (samples)
- Κάθε δείγμα είναι ένα διάνυσμα που αποτελείται από χαρακτηριστικά (features)

Τα δεδομένα

- Συνήθως κάθε δείγμα αναπαριστά αριθμητικά μια μέτρηση ή μια οντότητα από τον φυσικό κόσμο
- Παράδειγμα: μετρήσαμε σήμερα τη μέση θερμοκρασία (25°C, νεφοκάλυψη (10%) και τη βαρομετρική πίεση (1013,2 mbar). Το διάνυσμα του σημερινού δείγματος θα είναι [25 10 1013,2]. Το αυριανό δείγμα θα έχει κάποιες άλλες τιμές κοκ.

Τα δεδομένα

- Πολλές φορές δεν μας δίνονται μετρήσεις αλλά πρέπει να δημιουργήσουμε εμείς τα χαρακτηριστικά, να κάνουμε δηλαδή εξαγωγή χαρακτηριστικών (feature extraction) από τα αντικείμενα της συλλογής. Παραδείγματα:
- Συλλογή εικόνων: μετατρέπουμε κάθε εικόνα σε ένα διάνυσμα όπου τα χαρακτηριστικά είναι τιμές ιστογραμμάτων χρωμάτων
- Συλλογή κειμένων: μετατρέπουμε κάθε κείμενο σε ένα διάνυσμα όπου κάθε χαρακτηριστικό είναι η συχνότητα εμφάνισης κάθε λέξης

Τα δεδομένα

- Ειδικά στην επιβλεπόμενη μάθηση, σε κάθε δείγμα αντιστοιχεί και μια ετικέτα κατηγορίας ή κλάσης (class label) ή μια τιμή εξόδου. Για παράδειγμα στο δείγμα καιρού [25 10 1013,2] της σημερινής μέρας μπορεί να αντιστοιχεί :
  - η ετικέτα "Έβρεξε" ή
  - η τιμή εξόδου 1,5 που αντιστοιχεί στα mm βροχής που παρατηρήθηκαν τη συγκεκριμένη μέρα
- Οι ετικέτες ή οι τιμές χρησιμοποιούνται για την εκπαίδευση του συστήματος στην επιβλεπόμενη μάθηση

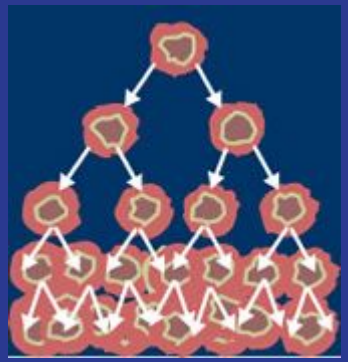
Παράδειγμα dataset: Wisconsin Breast Cancer

Attribute Information:

1) ID number
2) Diagnosis (M = malignant, B = benign)
3-32)

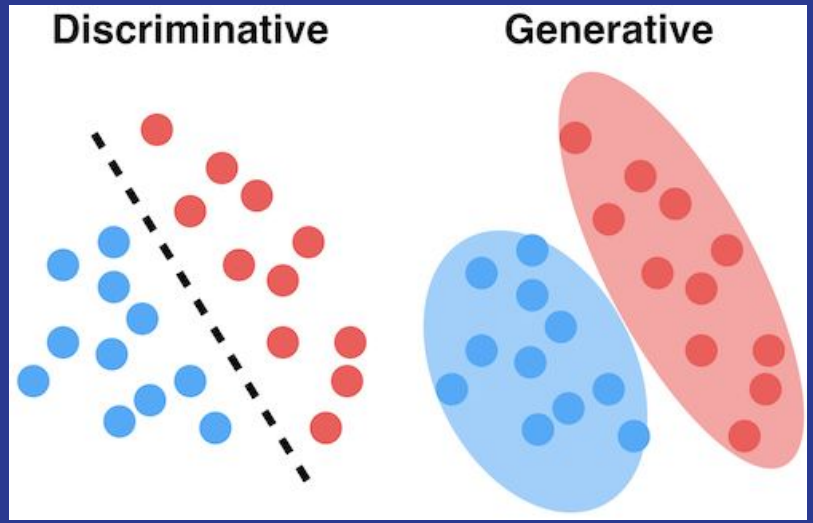
Ten real-valued features are computed for each cell nucleus:

a) radius (mean of distances from center to points on the perimeter)
b) texture (standard deviation of gray-scale values)
c) perimeter
d) area
e) smoothness (local variation in radius lengths)
f) compactness (perimeter^2 / area - 1.0)
g) concavity (severity of concave portions of the contour)
h) concave points (number of concave portions of the contour)
i) symmetry
j) fractal dimension ("coastline approximation" - 1)



```
842302,M,17.99,10.38,122.8,1001,0.1184,0.2776,0.3001,0.1471,0.2419,0.07871,1.095,0.9053,8.589,153.4,0.006399,0.04904,0.05373,0.01587,0.03003,0.006193,25.38,17.33,184.6,2019,0.1622,0.6656,0.7119,0.2654,0.4601,0.1189
842517,M,20.57,17.77,132.9,1326,0.08474,0.07864,0.0869,0.07017,0.1812,0.05667,0.5435,0.7339,3.398,74.08,0.005225,0.01308,0.0186,0.0134,0.01389,0.003532,24.99,23.41,158.8,1956,0.1238,0.1866,0.2416,0.186,0.275,0.08902
84300903,M,19.69,21.25,130,1203,0.1096,0.1599,0.1974,0.1279,0.2069,0.05999,0.7456,0.7869,4.585,94.03,0.00615,0.04006,0.03832,0.02058,0.0225,0.004571,23.57,25.53,152.5,1709,0.1444,0.4245,0.4504,0.243,0.3613,0.08758
84348301,M,11.42,20.38,77.58,386.1,0.1425,0.2839,0.2414,0.1052,0.2597,0.09744,0.4956,1.156,3.445,27.23,0.00911,0.07458,0.05661,0.01867,0.05963,0.009208,14.91,26.5,98.87,567.7,0.2098,0.8663,0.6869,0.2575,0.6638,0.173
84358402,M,20.29,14.34,135.1,1297,0.1003,0.1328,0.198,0.1043,0.1809,0.05883,0.7572,0.7813,5.438,94.44,0.01149,0.02461,0.05688,0.01885,0.01756,0.005115,22.54,16.67,152.2,1575,0.1374,0.205,0.4,0.1625,0.2364,0.07678
843786,M,12.45,15.7,82.57,477.1,0.1278,0.17,0.1578,0.08089,0.2087,0.07613,0.3345,0.8902,2.217,27.19,0.00751,0.03345,0.03672,0.01137,0.02165,0.005082,15.47,23.75,103.4,741.6,0.1791,0.5249,0.5355,0.1741,0.3985,0.1244
844359,M,18.25,19.98,119.6,1040,0.09463,0.109,0.1127,0.074,0.1794,0.05742,0.4467,0.7732,3.18,53.91,0.004314,0.01382,0.02254,0.01039,0.01369,0.002179,22.88,27.66,153.2,1606,0.1442,0.2576,0.3784,0.1932,0.3063,0.08368
84458202,M,13.71,20.83,90.2,577.9,0.1189,0.1645,0.09366,0.05985,0.2196,0.07451,0.5835,1.377,3.856,50.96,0.008805,0.03029,0.02488,0.01448,0.01486,0.005412,17.06,28.14,110.6,897,0.1654,0.3682,0.2678,0.1556,0.3196,0.1151
844981,M,13,21.82,87.5,519.8,0.1273,0.1932,0.1859,0.09353,0.235,0.07389,0.3063,1.002,2.406,24.32,0.005731,0.03502,0.03553,0.01226,0.02143,0.003749,15.49,30.73,106.2,739.3,0.1703,0.5401,0.539,0.206,0.4378,0.1072
84501001,M,12.46,24.04,83.97,475.9,0.1186,0.2396,0.2273,0.08543,0.203,0.08243,0.2976,1.599,2.039,23.94,0.007149,0.07217,0.07743,0.01432,0.01789,0.01008,15.09,40.68,97.65,711.4,0.1853,1.058,1.105,0.221,0.4366,0.2075
845636,M,16.02,23.24,102.7,797.8,0.08206,0.06669,0.03299,0.03323,0.1528,0.05697,0.3795,1.187,2.466,40.51,0.004029,0.009269,0.01101,0.007591,0.0146,0.003042,19.19,33.88,123.8,1150,0.1181,0.1551,0.1459,0.09975,0.2948,0.08452
84610002,M,15.78,17.89,103.6,781,0.0971,0.1292,0.09954,0.06606,0.1842,0.06082,0.5058,0.9849,3.564,54.16,0.005771,0.04061,0.02791,0.01282,0.02008,0.004144,20.42,27.28,136.5,1299,0.1396,0.5609,0.3965,0.181,0.3792,0.1048
846226,M,19.17,24.8,132.4,1123,0.0974,0.2458,0.2065,0.1118,0.2397,0.078,0.9555,3.568,11.07,116.2,0.003139,0.08297,0.0889,0.0409,0.04484,0.01284,20.96,29.94,151.7,1332,0.1037,0.3903,0.3639,0.1767,0.3176,0.1023
846381,M,15.85,23.95,103.7,782.7,0.08401,0.1002,0.09938,0.05364,0.1847,0.05338,0.4033,1.078,2.903,36.58,0.009769,0.03126,0.05051,0.01992,0.02981,0.003002,16.84,27.66,112,876.5,0.1131,0.1924,0.2322,0.1119,0.2809,0.06287
84667401,M,13.73,22.61,93.6,578.3,0.1131,0.2293,0.2128,0.08025,0.2069,0.07682,0.2121,1.169,2.061,19.21,0.006429,0.05936,0.05501,0.01628,0.01961,0.008093,15.03,32.01,108.8,697.7,0.1651,0.7725,0.6943,0.2208,0.3596,0.1431
84799002,M,14.54,27.54,96.73,658.8,0.1139,0.1595,0.1639,0.07364,0.2303,0.07077,0.37,1.033,2.879,32.55,0.005607,0.0424,0.04741,0.0109,0.01857,0.005466,17.46,37.13,124.1,943.2,0.1678,0.6577,0.7026,0.1712,0.4218,0.1341
848406,M,14.68,20.13,94.74,684.5,0.09867,0.072,0.07395,0.05259,0.1586,0.05922,0.4727,1.24,3.195,45.4,0.005718,0.01162,0.01998,0.01109,0.0141,0.002085,19.07,30.88,123.4,1138,0.1464,0.1871,0.2914,0.1609,0.3029,0.08216
84862001,M,16.13,20.68,108.1,798.8,0.117,0.2022,0.1722,0.1028,0.2164,0.07356,0.5692,1.073,3.854,54.18,0.007026,0.02501,0.03188,0.01297,0.01689,0.004142,20.96,31.48,136.8,1315,0.1789,0.4233,0.4784,0.2073,0.3706,0.1142
849014,M,19.81,22.15,130,1260,0.09831,0.1027,0.1479,0.09498,0.1582,0.05395,0.7582,1.017,5.865,112.4,0.006494,0.01893,0.03391,0.01521,0.01356,0.001997,27.32,30.88,186.8,2398,0.1512,0.315,0.5372,0.2388,0.2768,0.07615
8510426,B,13.54,14.36,87.46,566.3,0.09779,0.08129,0.06664,0.04781,0.1885,0.05766,0.2699,0.7886,2.058,23.56,0.008462,0.0146,0.02387,0.01315,0.0198,0.0023,15.11,19.26,99.7,711.2,0.144,0.1773,0.239,0.1288,0.2977,0.07259
8510653,B,13.08,15.71,85.63,520,0.1075,0.127,0.04568,0.0311,0.1967,0.06811,0.1852,0.7477,1.383,14.67,0.004097,0.01898,0.01698,0.00649,0.01678,0.002425,14.5,20.49,96.09,630.5,0.1312,0.2776,0.189,0.07283,0.3184,0.08183
8510824,B,9.504,12.44,60.34,273.9,0.1024,0.06492,0.02956,0.02076,0.1815,0.06905,0.2773,0.9768,1.909,15.7,0.009606,0.01432,0.01985,0.01421,0.02027,0.002968,10.23,15.66,65.13,314.9,0.1324,0.1148,0.08867,0.06227,0.245,0.07773
8511133,M,15.34,14.26,102.5,704.4,0.1073,0.2135,0.2077,0.09756,0.2521,0.07032,0.4388,0.7096,3.384,44.91,0.006789,0.05328,0.06446,0.02252,0.03672,0.004394,18.07,19.08,125.1,980.9,0.139,0.5954,0.6305,0.2393,0.4667,0.09946
```

Generative - Discriminative

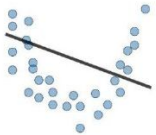
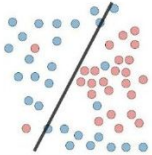
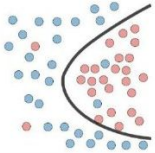
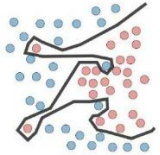


- Τα δεδομένα είναι διανύσματα σε ένα χώρο μεγάλων διαστάσεων

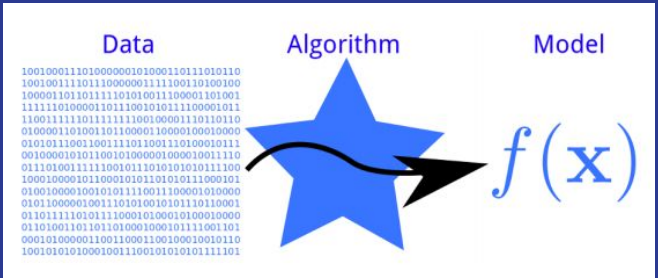
- Παραμετρικοί αλγόριθμοι (πχ GMM)
- Μη-Παραμετρικοί Αλγόριθμοι (πχ SVM)

Bias-Variance Tradeoff

(Ανταλλαγή Προκατάληψης-Διακύμανσης)

	Underfitting	Just right	Overfitting
Symptoms	- High training error - Training error close to test error - High bias	- Training error slightly lower than test error	- Low training error - Training error much lower than test error - High variance
Regression			
Classification			

Γενίκευση (Generalization)



Έχοντας δει κάποια παραδείγματα, μπορεί το σύστημα να αντιμετωπίσει “σωστά” νέα παραδείγματα που δεν έχει “δει” ποτέ; -> **Γενίκευση, Μάθηση**

Πώς μαθαίνουμε;

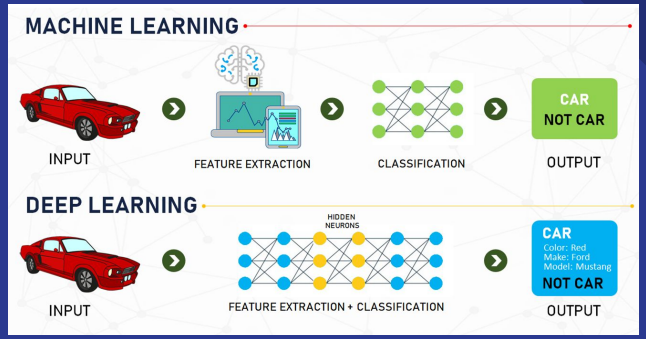
- *Υποσυμβολικό* επίπεδο αντίληψης (δεδομένα)
- *Συμβολικό* επίπεδο αντίληψης (?)
- *Γενίκευση*

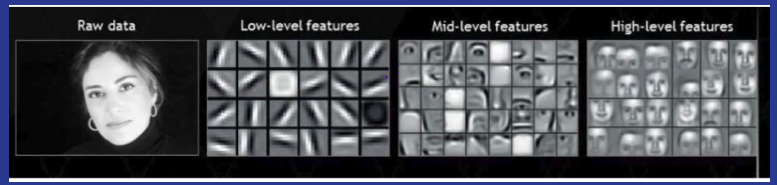
Καλή γενίκευση: “αντίστοιχο” της συμβολικής αναπαράστασης του δέντρου, δεν γνωρίζουμε τους δικούς μας μηχανισμούς μάθησης
Υπερεκπαίδευση: αναγνωρίζουμε μόνο ότι έχει τη μορφή κονοφόρου
Υποεκπαίδευση: ό,τι είναι πράσινο είναι δέντρο

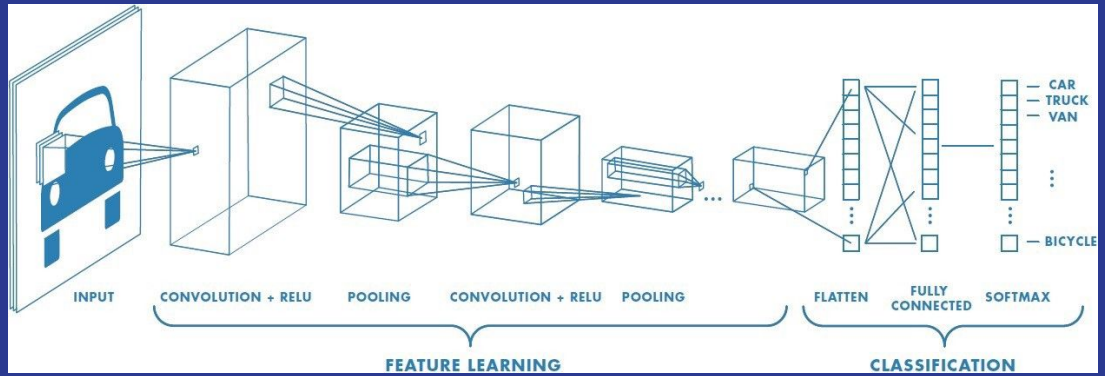
-> Η Γενίκευση σημαίνει αναπόφευκτα (αποδεκτό) **σφάλμα**
-> Η Γενίκευση σημαίνει όμως και **ευρωστία, προσαρμοστικότητα, ανακάλυψη**

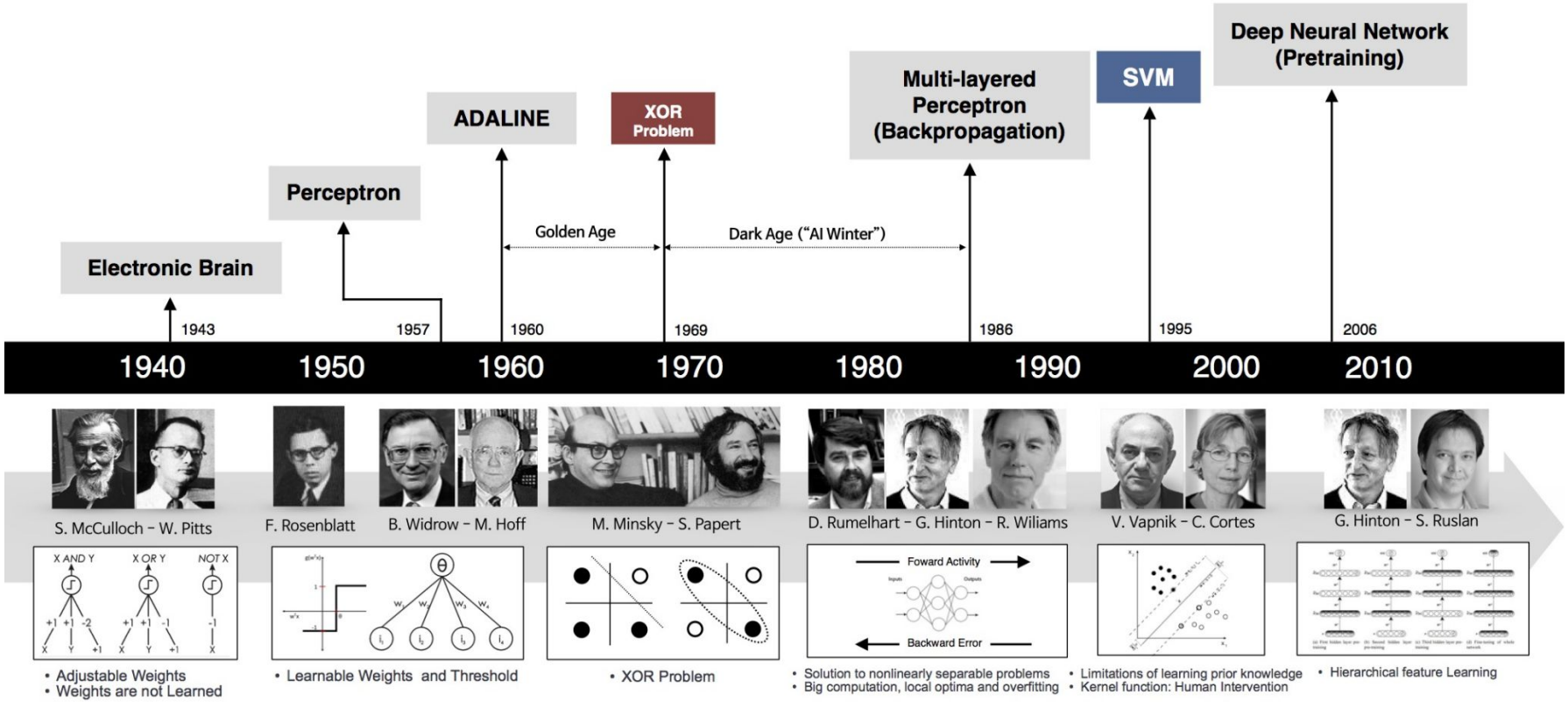
Βαθιά Μάθηση

- συνελικτικά δίκτυα
- αυτόματη εξαγωγή χαρακτηριστικών
- μεταφορά μάθησης





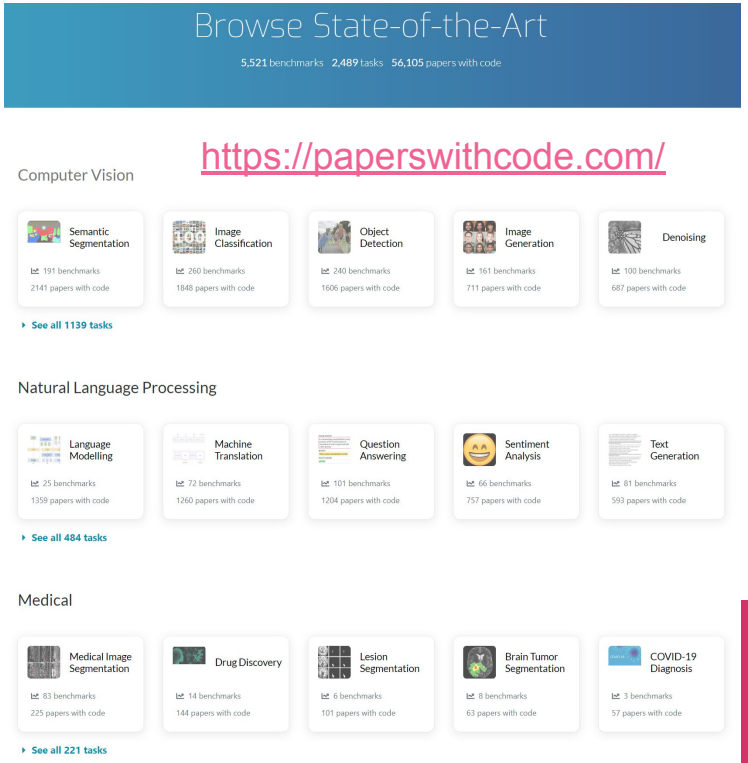



Deep Neural Network (Pretraining)
SVM
Multi-layered Perceptron (Backpropagation)
ADALINE
XOR Problem
Perceptron
Electronic Brain
Golden Age
Dark Age ("AI Winter")
1943
1957
1960
1969
1986
1995
2006
1940
1950
1960
1970
1980
1990
2000
2010
S. McCulloch - W. Pitts
F. Rosenblatt
B. Widrow - M. Hoff
M. Minsky - S. Papert
D. Rumelhart - G. Hinton - R. Wiliams
V. Vapnik - C. Cortes
G. Hinton - S. Ruslan
X AND Y
X OR Y
NOT X
Foward Activity
Backward Error
• Adjustable Weights
• Weights are not Learned
• Learnable Weights and Threshold
• XOR Problem
• Solution to nonlinearly separable problems
• Big computation, local optima and overfitting
• Limitations of learning prior knowledge
• Kernel function: Human Intervention
• Hierarchical feature Learning

Η εποχή της Μηχανικής Μάθησης

- Περισσότερα δεδομένα - Internet
- Αλγοριθμικές εξελίξεις
- Αλματώδης βελτίωση hardware
- Καλές βιβλιοθήκες (λογισμικό)
- **Strong AI vs Weak AI**

Computer Vision και NLP είναι τα πεδία με τα περισσότερα tasks



Η εποχή της Μηχανικής Μάθησης

Το Turing Award του 2019 απονεμήθηκε σε τρεις “γίγαντές” της ΜΜ (NY times article)

…αλλά πάντα υπάρχει (γόνιμος) αντίλογος -Schmidhuber-

Αναμφίβολα όμως, το βραβείο αυτό είναι η παγκόσμια επιστημονική (και όχι μόνο) αναγνώριση ότι η Μηχανική Μάθηση, παρότι τελείως αντισυμβατική (ένας υπολογιστής που... κάνει λάθη) είναι πλέον αποδεκτή ως ένα απολύτως έγκυρο και καθοριστικό για το μέλλον υπολογιστικό παράδειγμα.