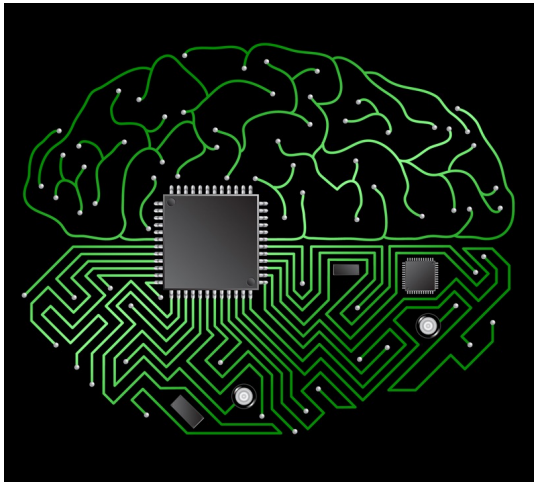


Inteligencia Artificial

Rafael Ramirez Melendez

Universitat Pompeu Fabra, Barcelona

rafael.ramirez@upf.edu

Que es la Inteligencia Artificial?

Cual es el estado de la IA en la actualidad?

Cual es el potencial y el peligro de la IA?

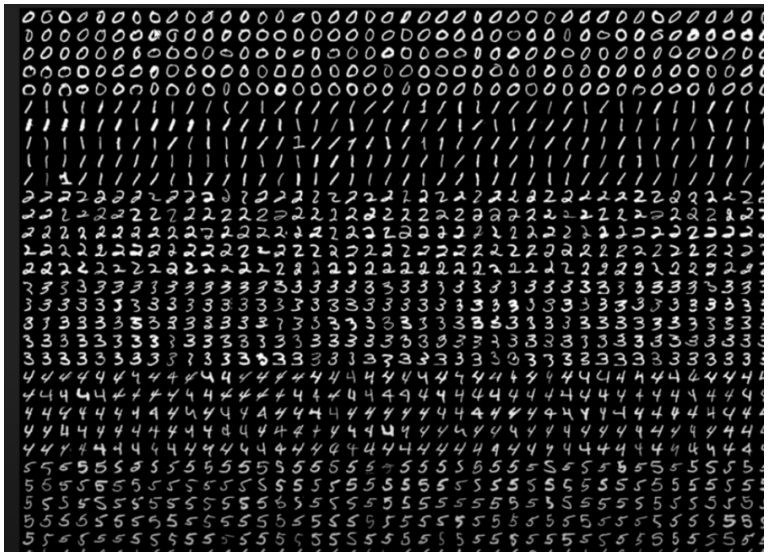
Que es la Inteligencia Artificial?

Why AI / Machine Learning?

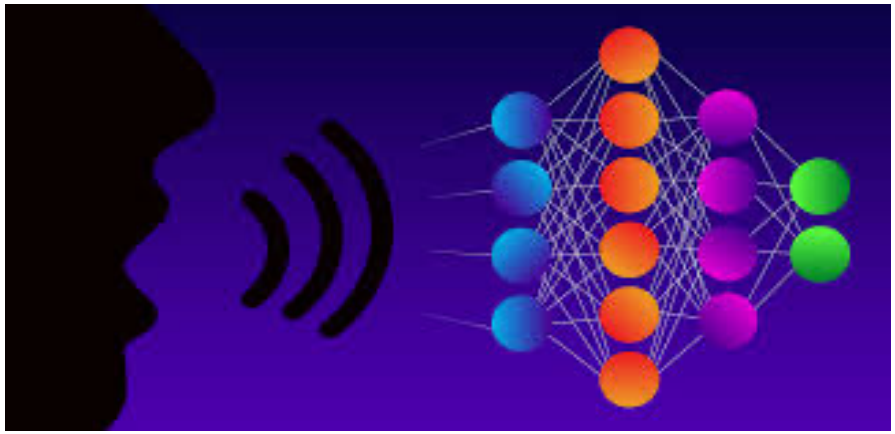
- Machine learning extremely useful in most if not all areas

Autonomous vehicles,
handwriting recognition,
Natural Language Processing,
Computer Vision
…

Self-customizing programs:
Amazon, Netflix
product recommendations, …



Handwriting recognition

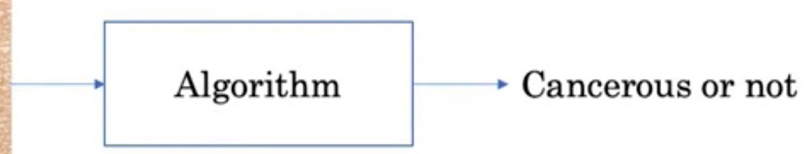


Natural Language Processing - text from/to speech, translation,…

Self-customizing programs Amazon, Netflix, product recommendations, …

ML in Biomedicine

Dermatology



ML in Biomedicine

Images

Labels

Cancerous

...

Non-cancerous

129,000 images

Convolutional Neural Network

ML in Biomedicine

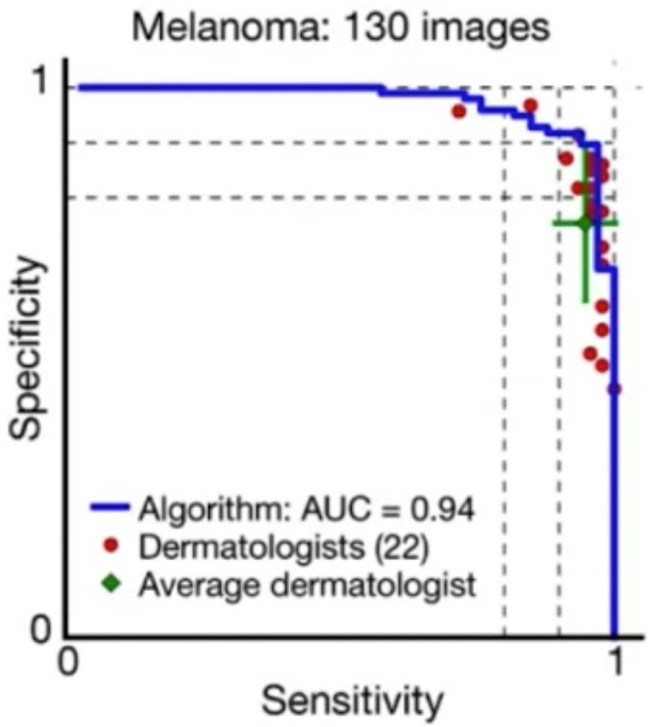
Cancerous or not

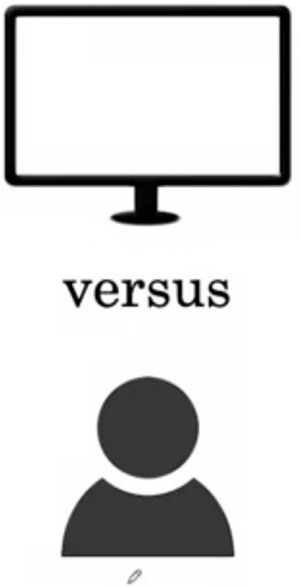
versus

AI vs. Human performance: Dermatología

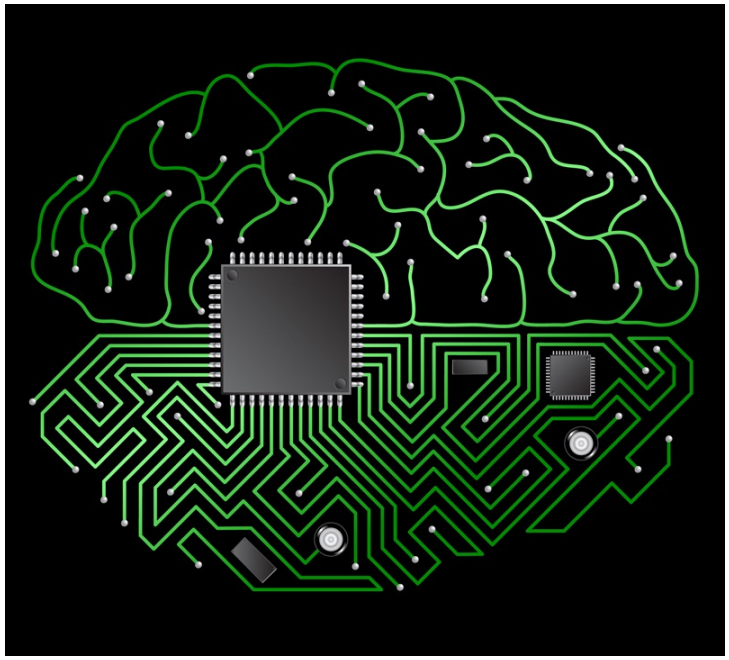
ML in Biomedicine

Sensitivity=TP/P
Specificity=TN/N





What is Artificial Intelligence?



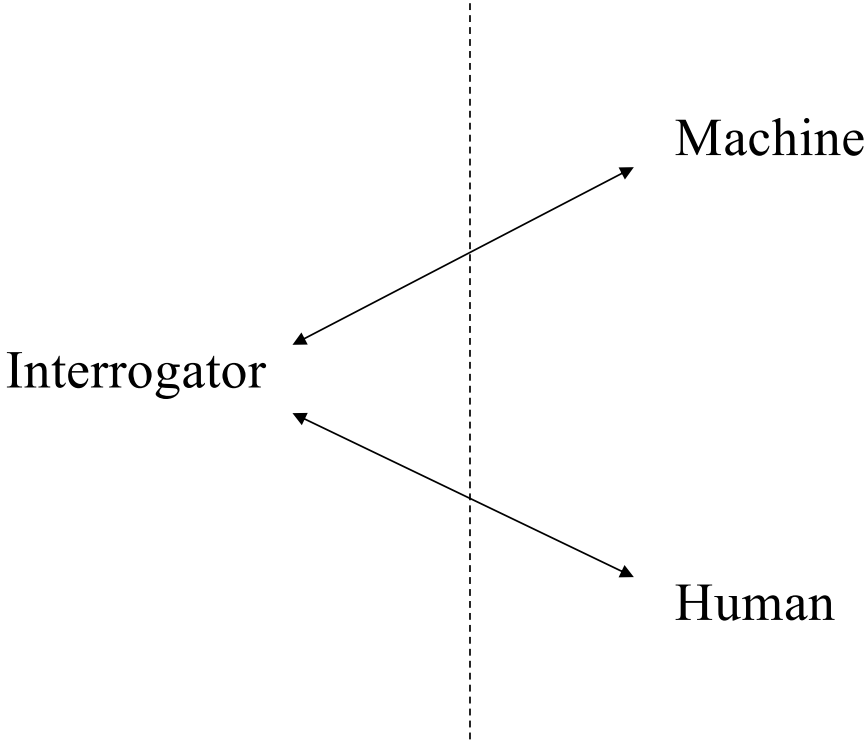
“The science and engineering of making intelligent machines”

(John McCarthy, 1955)

Turing's test

From Computer Desktop Encyclopedia
Reproduced with permission.
© 2001 The Computer Museum History Center

Alan Turing (1912-1954)



Deep Blue vs Kasparov

IBM 1996, 1997

Deep blue moment of glory

What about learning?

Que es el aprendizaje?

El aprendizaje es mas que la memorización de observaciones:

=> Noción central: GENERALISACIÓN

Asi que, que es el aprendizaje?

El aprendizaje se basa en experiencia:

- Un aprendiz necesita observaciones / ejemplos para aprender
- Un aprendiz necesita feedback para evaluar si lo hace bien

El aprendizaje requiere por lo menos memorizar:

- Un aprendiz necesita memoria y la habilidad de modificar la memoria

Que mas?

Machine Learning

Field of study that gives computers the ability to learn without being explicitly programmed.

Arthur Samuel (1901-1990)

Machine Learning

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.

---Tom Mitchell

Data Mining

“The automatic analysis of large quantities of data to extract previously unknown interesting patterns”

Induction vs Deduction

Deduction (logically justified):
known: general rule ***p*** → ***q*** (“whenever ***p*** is true, ***q*** is also true”)
known/observed: ***p*** (“***p*** is true”)
deductive conclusion: ***q*** (“***q*** is true”)

Abduction
known: general rule ***p*** → ***q*** (“whenever ***p*** is true, ***q*** is also true”)
known/observed: ***q*** (“***q*** is true”)
deductive conclusion: ***p*** (“***p*** is true”)

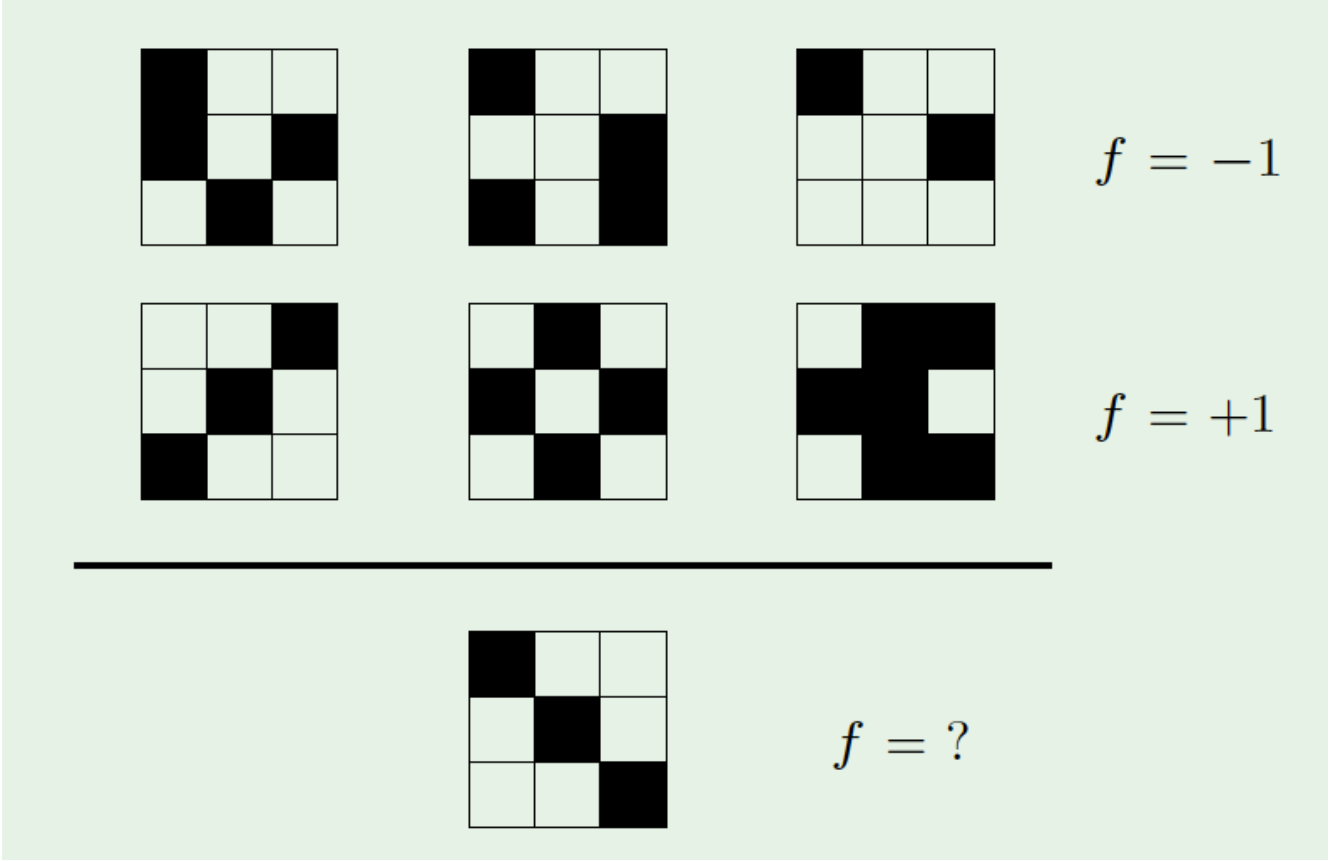
Induction vs Deduction

Induction (logically unjustified):

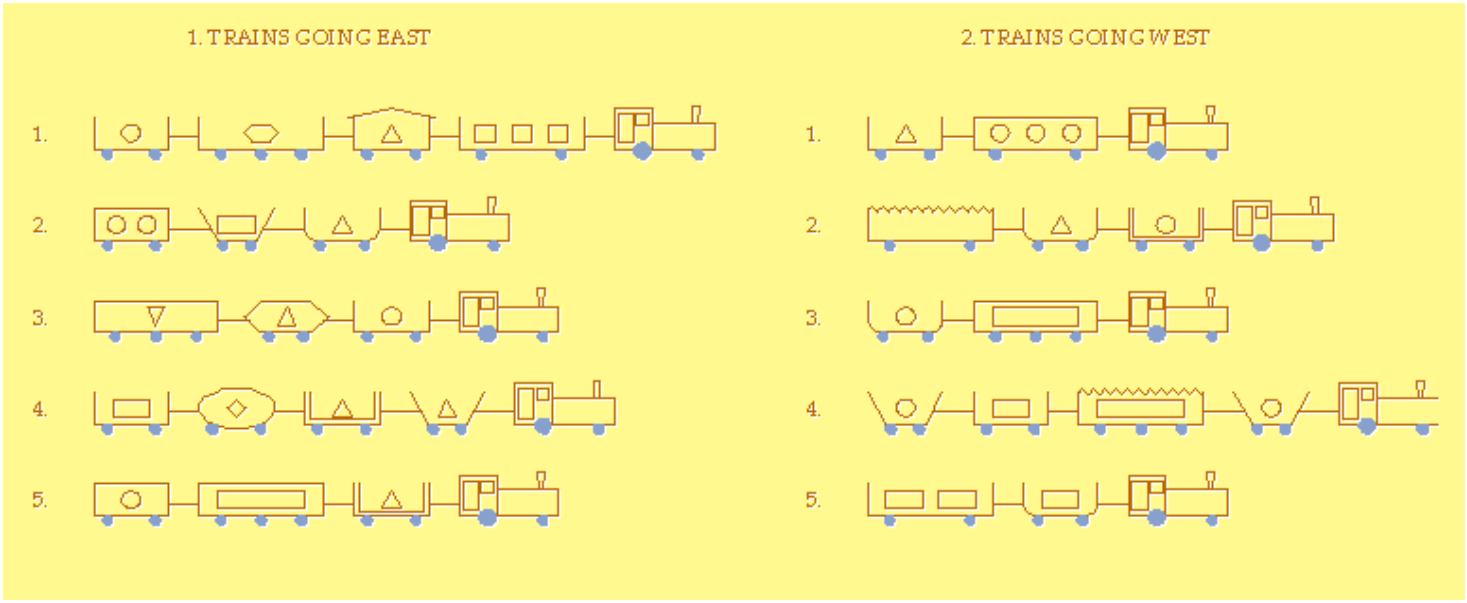
observed sets of facts:
p, q, r (***p***, ***q***, and ***r*** occurred together)
p, q, s, t, … (***p, q, s, t, …*** occurred together)
p, q, s, u, v, …

inductive conclusion: ***p*** → ***q*** (“whenever ***p*** is true, ***q*** is also true”)
or ***q*** → ***p*** (“whenever ***q*** is true, ***p*** is also true”)

Induction



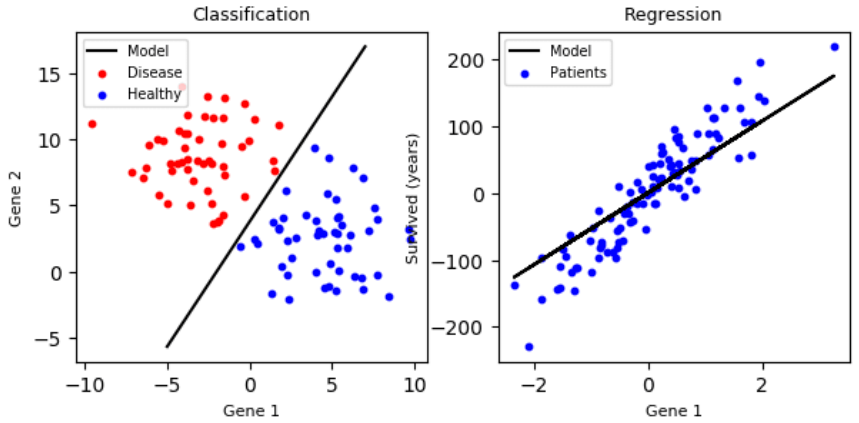
Induction

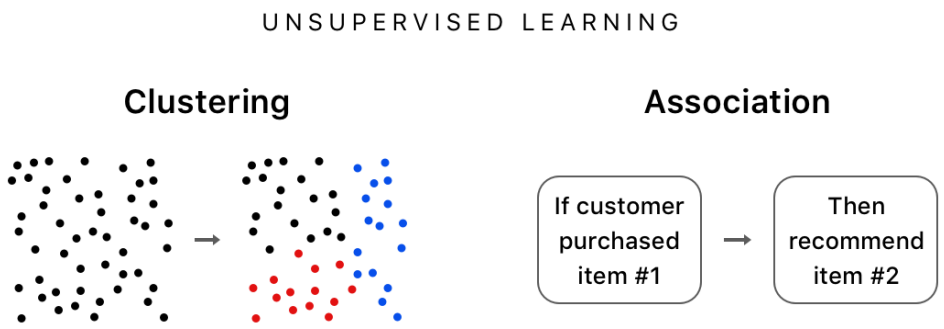


What makes a train go eastward ?

Induction tasks

- Classification
- Regression
- Clustering





Induction

Supervised learning

vs.

Non supervised learning

Induction

Supervised Learning

- Regression
- Classification

Regression Example: Treatment-Side effects

Training set of treatments

Treatment (hrs)	Side-effects
3	3.5
4.2	4.1
1.5	2.75
7.3	5.65
…	…

What is the side effect of a treatment of 9hrs?

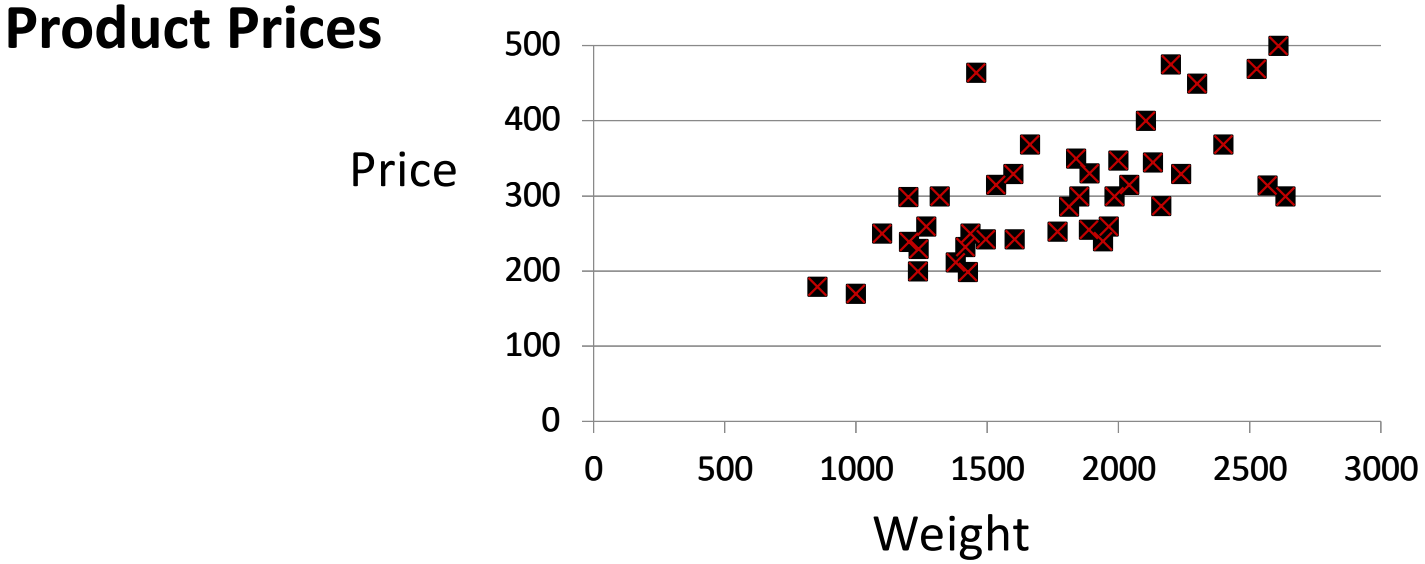
Regression Example: Rent Prices

Training set of rent prices

# Rooms	Price (€)
1	300
3	1100
4	1800
5	2700
…	…

What is the price of a house with 2 rooms?

Realistic Regression Example



Model: y = ax + b

Model: price= a*weight + b

Classification Example: Customers

Training set of customers' decision

Age	Income (€)	Buy?
27	4300	yes
25	3200	no
30	3800	no
18	3000	yes
43	5000	no
39	4400	no
32	4300	yes
20	3100	yes
…	…	…

Would a 22 year old earning 3000 buy your product?

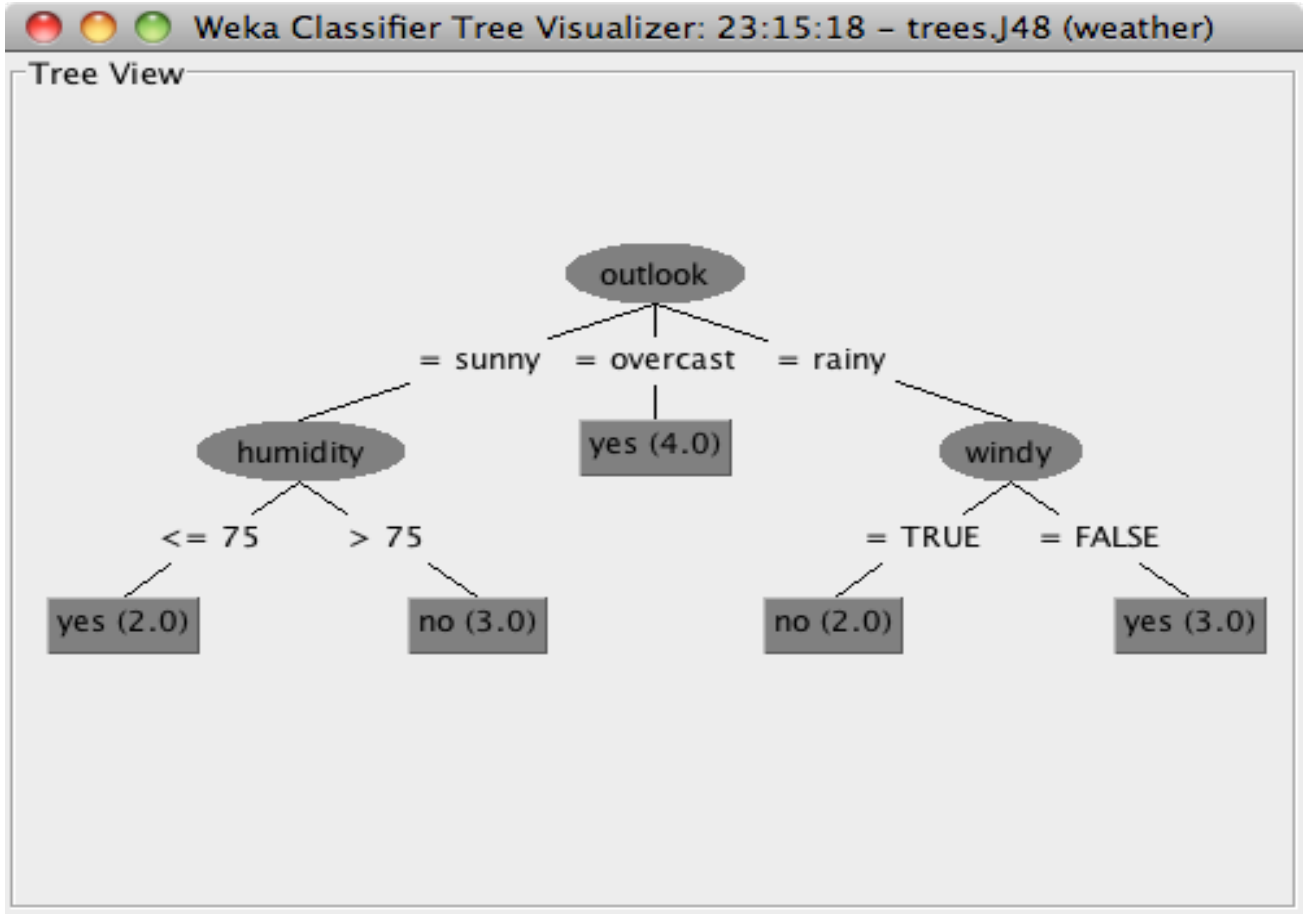
Classification Example: Tennis Player

No.	Outlook (nominal)	Temperature (numeric)	Humidity (numeric)	Windy (nominal)	Play? (nominal)
1	sunny	85.0	85.0	FALSE	no
2	sunny	80.0	90.0	TRUE	no
3	overcast	83.0	86.0	FALSE	yes
4	rainy	70.0	96.0	FALSE	yes
5	rainy	68.0	80.0	FALSE	yes
6	rainy	65.0	70.0	TRUE	no
7	overcast	64.0	65.0	TRUE	yes
8	sunny	72.0	95.0	FALSE	no
9	sunny	69.0	70.0	FALSE	yes
10	rainy	75.0	80.0	FALSE	yes
11	sunny	75.0	70.0	TRUE	yes
12	overcast	72.0	90.0	TRUE	yes
13	overcast	81.0	75.0	FALSE	yes
14	rainy	71.0	91.0	TRUE	no

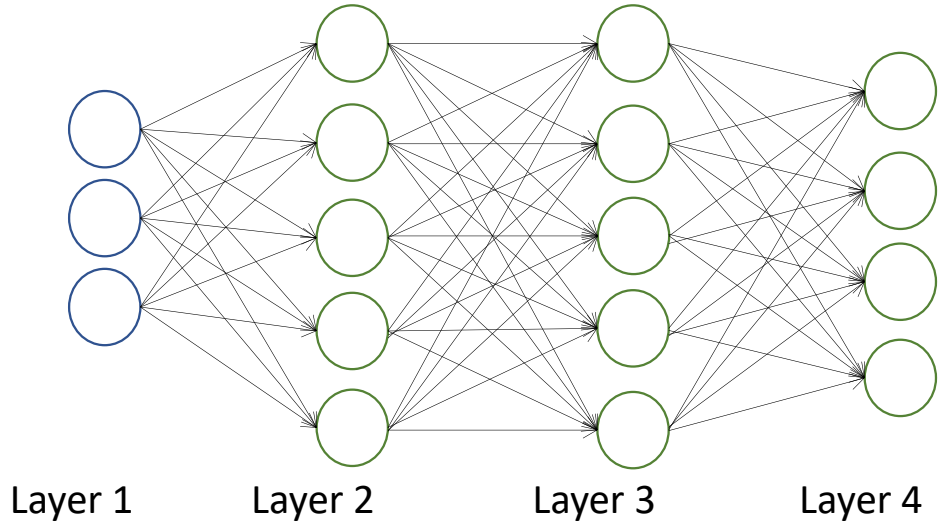
Linear Model

0.8436 * outlook=sunny
+ -0.9535 * outlook=overcast
+ 0.1099 * outlook=rainy
+ 0.5276 * temperature
+ 0.771 * humidity
+ -0.8901 * windy
- 0.8683

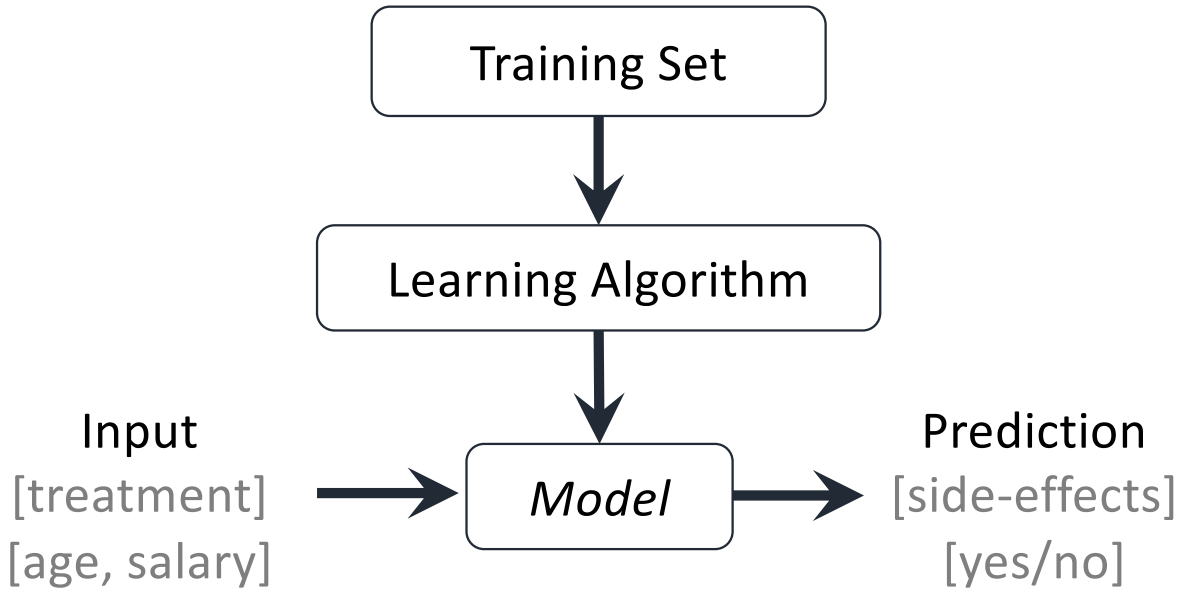
Tree-based Model



ANN Model



General Method



Real-life Example: Direct Marketing

Data mining task

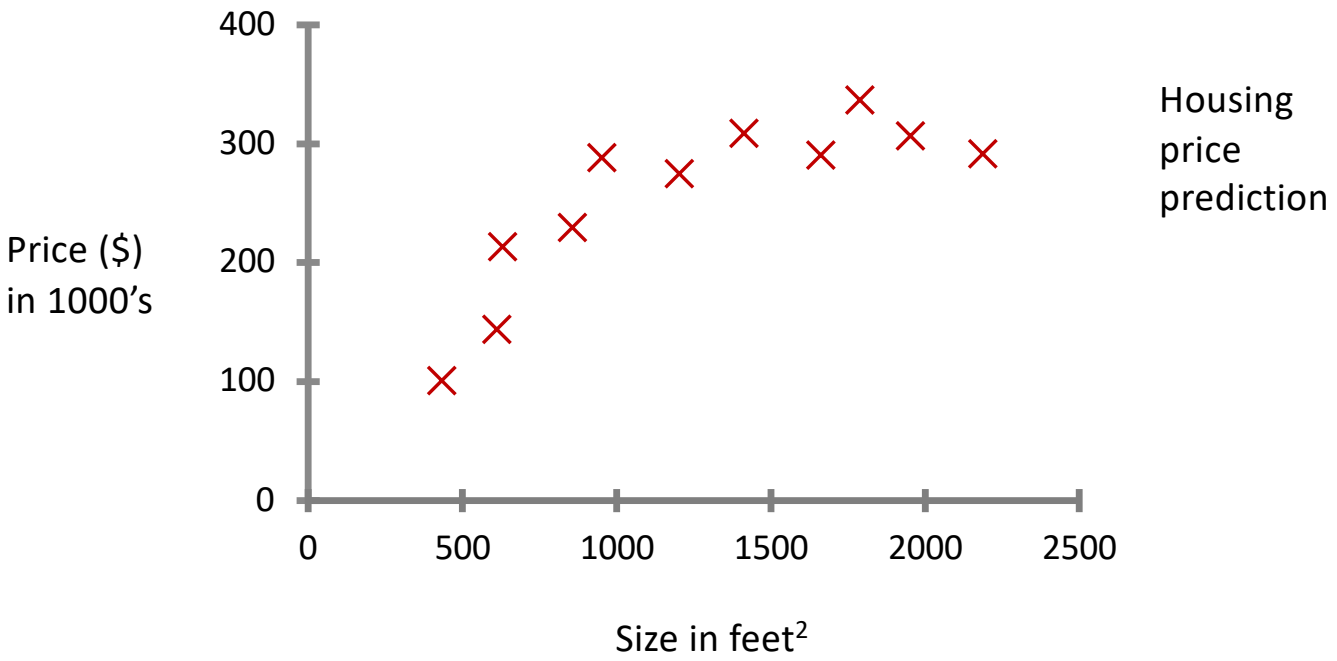
- Bank wants to conduct a *directed* marketing campaign for attracting new deposits
- Data provided of past marketing campaigns:
  - 17 campaigns between 2008-2010
  - 45,211 contacts
- For each contact: *Age, Marital status, Job, Education, Annual balance, Credit, Personal loan, Housing loan, Contact communication type, Day, Month, Duration, etc.*

- Advice about strategies to conduct the directed marketing campaign

Real-life Example: Direct Marketing

Name	Description and Values
Personal Client Information	
Age	Age at the contact date (Numeric ≥18)
Marital status	Married, single, divorced, widowed, separated (Nominal)
Job	Services, management, blue-collar, admin, technician, entrepreneur, student, housemaid, retired, unemployed, self-employed
Education	unknown, secondary, primary, tertiary (Nominal)
Bank Client Information	
Annual balance	Average yearly balance in *euros* (Numeric)
Credit	Yes or No (Nominal)
Personal loan	Yes or No (Nominal)
Housing loan	Yes or No (Nominal)
Last Contact Information	
Contact communication type	Unknown, telephone, cellular (Nominal)
Day	Last contact day of the month (Numeric)
Month	Last contact month of year (Nominal)
Duration	Of the contact in seconds (Numeric)
Other attributes	
Campaign	Number of contacts performed during this campaign and for this client (Numeric)
Prev days	Number of days that passed by after the client was last contacted from a previous campaign (Numeric, -1 means client was not previously contacted)
Previous	Number of contacts performed before this campaign and for this client (Numeric)
Poutcome	Outcome of the previous marketing campaign (Nominal: unknown, other, failure, success)

Supervised Learning - Regression



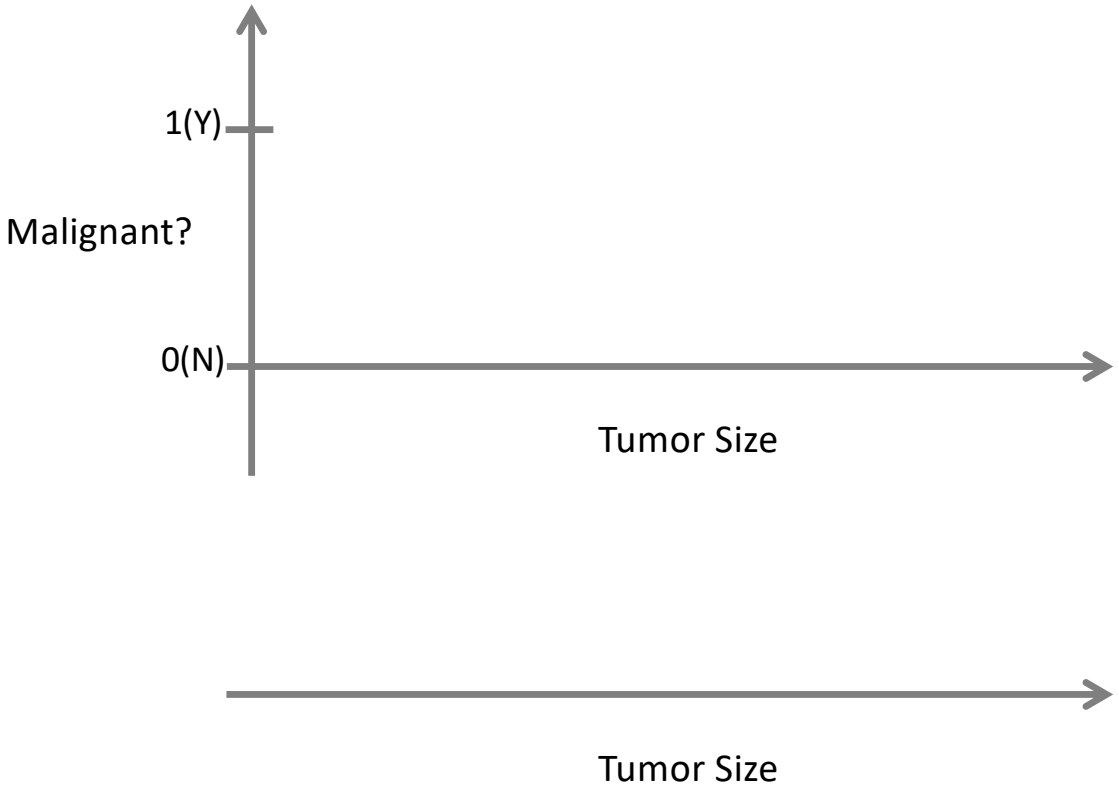
Supervised Learning

"right answers" given

Regression: Predict continuous valued output (price)

Supervised Learning - Classification

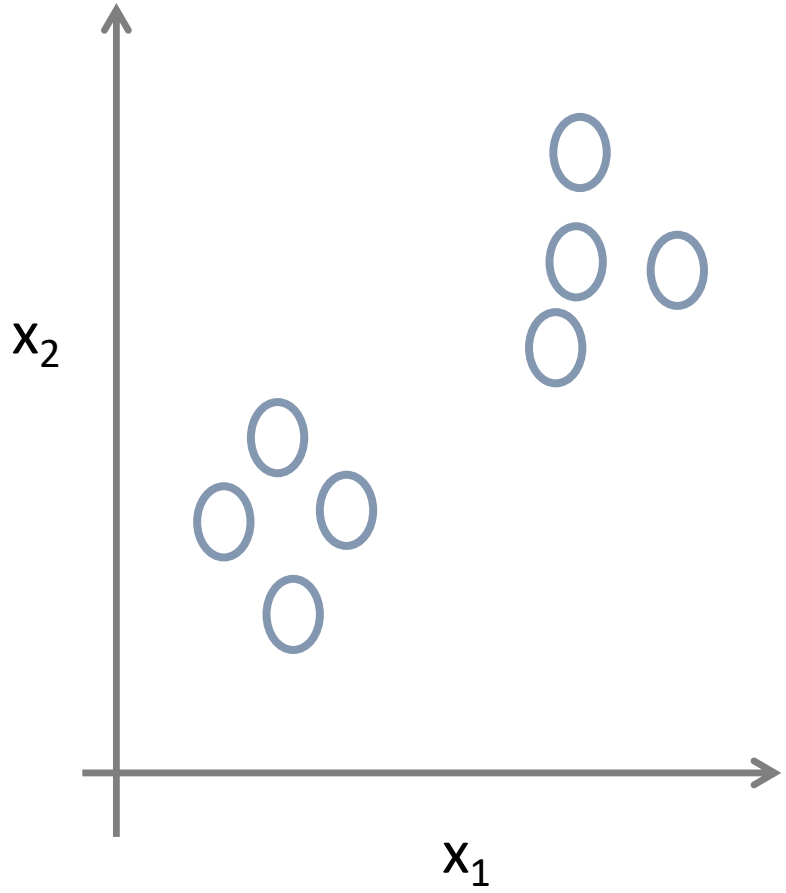
Breast cancer (malignant, benign)

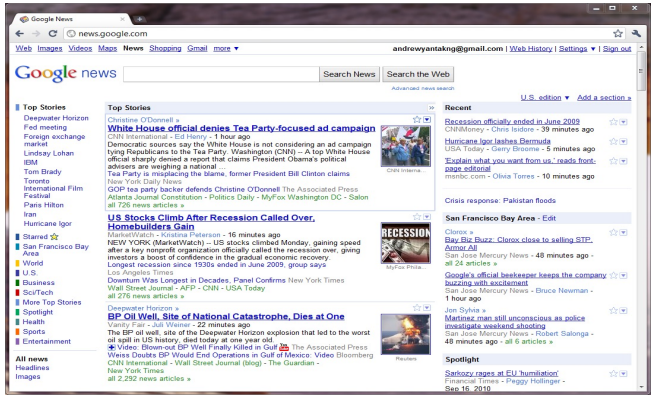


Classification

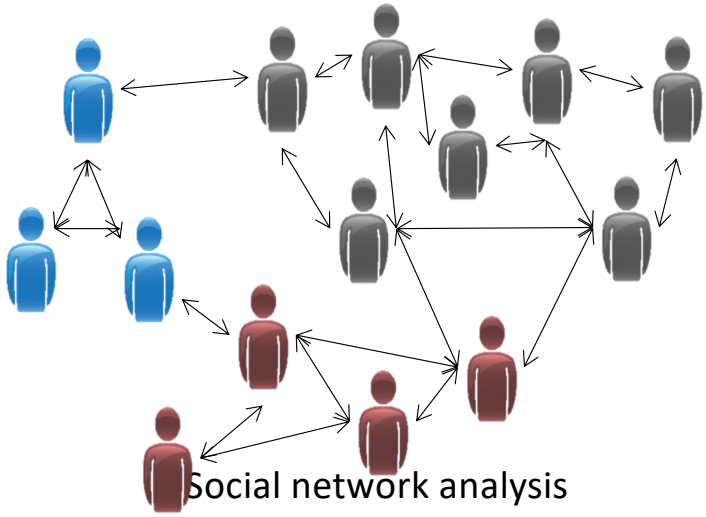
Discrete valued output (0 or 1)

Unsupervised Learning

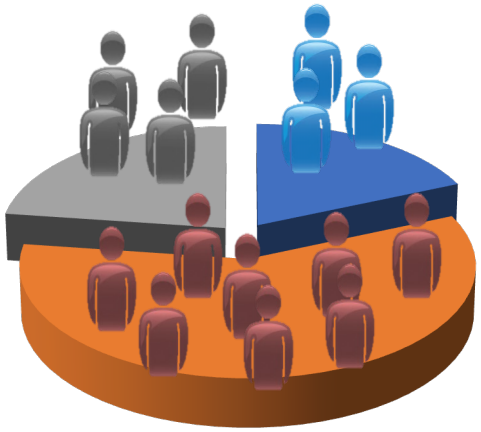




News clustering



Social network analysis



Market segmentation

Astronomical data analysis