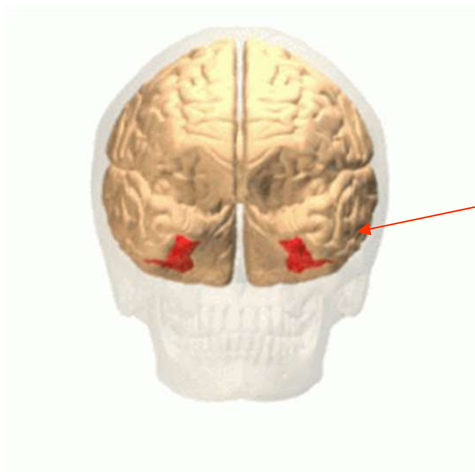
Machine learning

Chapter 1

Introduction

Searching for patterns in data is a fundamental problem humans have been facing throughout history. Millions of years ago, the ability to classify and identify poisonous plants or dangerous animals was a valuable skill for our ancestors, and it gave them a clear evolutionary advantage.

Brain support: as an example, prosopagnosia is a cognitive disorder of face perception in which the ability to recognize familiar faces, including one's own face, is impaired, while other aspects of visual processing (e.g., object discrimination) and intellectual functioning (e.g., decision-making) remain intact.



Nowadays, pattern recognition has evolved to automatic algorithms that can be applied in many scientific and technological fields, such as medical diagnosis, musicology, business, and financial sciences

Objective

To define learning techniques based on observed data with the objective of taking automatic smart decisions. The techniques are applicable to problems of very different nature.

Goals...

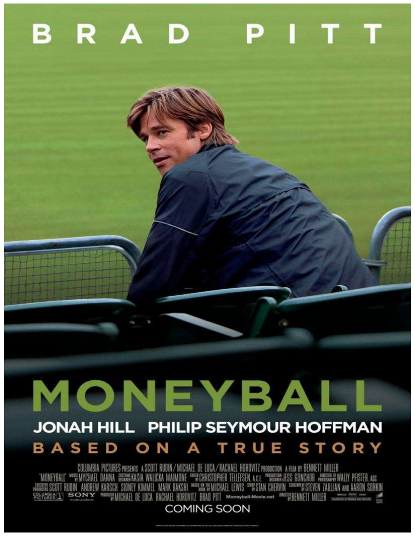
- Take better decisions than a human expert
- Reduction of costs

The art of winning an unfair game

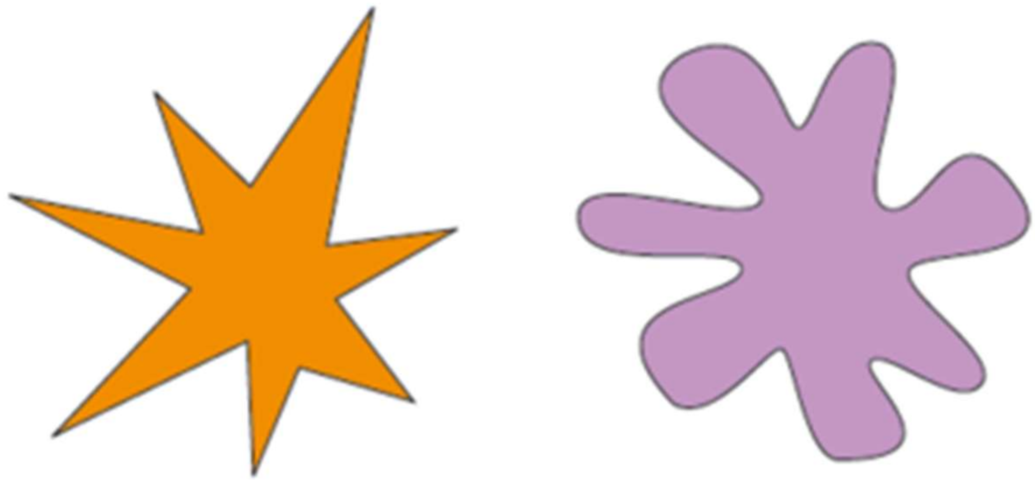
Billy Beane is the general manager of Oakland Athletics, which has lost another season in the baseball league. In 2002 he decides to relaunch the team and, with the help of the young Harvard graduated economist Peter Brand, will use statistical data to contract the most suitable players, instead of using the reports of experts. Oakland beats an old record of 20 consecutive victories in the Word Series.

Two years later, the Boston Red Sox used the same methods and wins the World Series after 86 years.

Why?



Cognitive biases



Which figure is called *bouba* and which one *kiki*?

50 congnitive biases

Memoria | Social | Aprendizaje | Creencias | Dinero | Política

Error de atribución fundamental
Juzgamos a los demás por su personalidad o carácter fundamental, pero nos juzgamos a nosotros mismos por la situación.

Sally llega tarde a clase, es vaga. Llegas tarde a clase; fue una mala mañana.

Sesgo de autoservicio
Nuestros fracasos son situacionales, pero nuestros éxitos son nuestra responsabilidad.

Ganó ese premio debido al trabajo duro en lugar de la ayuda o la suerte. Mientras tanto, falló una prueba porque no había dormido lo suficiente.

Favoritismo dentro del grupo
Favorecemos a las personas que están en nuestro grupo interno en lugar de un grupo externo

Francis está en tu iglesia, así que te gusta más Francis que Sally

Efecto de arrastre
Las ideas, modas y creencias crecen a medida que más personas las adoptan

Sally cree que los fidget spinners ayudan a sus hijos, Francis también lo hace

Pensamiento de grupo
Debido al deseo de conformidad y armonía en el grupo, tomamos decisiones irracionales, a menudo para minimizar el conflicto.

Sally quiere ir por un helado. Francis quiere comprar una camiseta. Sugieres conseguir una camiseta con imágenes de helado.

Efecto halo
Si ve a una persona como alguien que tiene un rasgo positivo, esa impresión positiva se extenderá a sus otros rasgos. (que también funciona para rasgos negativos)

"Taylor nunca podría ser mala; ella es tan linda"

Suerte moral
Una mejor posición moral ocurre debido a un resultado positivo; peor posición moral ocurre debido a un resultado negativo

"Una cultura X ganó una guerra X porque eran moralmente superiores a los perdedores"

Falsos concensos
Creemos que hay más personas de acuerdo con nosotros de lo que realmente es.

"Todos piensan que"

Maldición del conocimiento
Una vez que sabemos algo, asumimos que todos los demás también lo saben.

Alice es maestra y lucha por comprender la perspectiva de sus nuevos alumnos.

Efecto de foco
Sobrestimamos la cantidad de personas que prestan atención a nuestro comportamiento y apariencia

Sally está preocupada de que todos se den cuenta de lo patética que es su camiseta de helado.

Disponibilidad heurística
Nos basamos en ejemplos inmediatos que nos vienen a la mente al emitir juicios.

Al intentar decidir qué tienda visitar, elige la tienda para la que vio un anuncio más recientemente. es tan preciso"

Atribución defensiva
Como Testigo que en secreto teme ser vulnerable a un percance grave, culparemos a la víctima menos a un atacante más si nos relacionamos con la víctima

Sally esperó demasiado tiempo en un semáforo en verde porque estaba jugando con su teléfono. La chocaron por detrás. Greg, quien es conocido por enviar mensajes de texto y conducir, salió y le gritó a la persona que la golpeó.

Hipótesis de un mundo justo
Tendemos a creer que el mundo es justo; Por lo tanto, asumimos que los actos de injusticia son merecidos.

"El bolso de Sally fue robado porque ella era mala con Francis sobre su camiseta y tenía mal karma"

Realismo ingenuo
Creemos que observamos la realidad objetiva y que otras personas son irracionales, desinformadas o sesgadas.

"Veo el mundo como es la realidad - otras personas son tontas"

Cinismo ingenuo
Creemos que observamos la realidad objetiva y que otras personas tienen un sesgo egocéntrico más alto de lo que realmente tienen en sus acciones/ intenciones.

"La única razón por la que esta persona está haciendo algo bueno es para sacarme algo."

Efecto Forer (también conocido como efecto Barnum)
Atribuimos fácilmente nuestra personalidad a declaraciones vagas, incluso si pueden aplicarse a una amplia gama de personas

"Este horóscopo es tan preciso"

Efecto Dunning-Kruger
Cuanto menos sepa, más confianza tendrá. Cuanto más sepa, menos confianza tendrá

Francis asegura con confianza al grupo que no hay ayuda en el helado. No trabajan en la industria láctea.

Fondeo
Dependemos en gran medida de la primera información que se introduce al tomar decisiones.

"¿ese 50% de descuento? debe ser mucho"

Sesgo de automatización
Confiamos en sistemas automatizados, algo que confía demasiado en la corrección automatizada de decisiones realmente correctas

Su teléfono lo corrige automáticamente, por lo que asume que es correcto

Efecto Google (también conocido como efecto amnesia)
Tendemos a olvidar la información que se busca fácilmente en los motores de búsqueda.

"¿Cómo se llamaba ese actor en esa película lunática? Lo he buscado como ocho veces ..."

Resistencia reactiva
Hacemos lo contrario de lo que nos dicen, especialmente cuando percibimos amenazas a las libertades personales

Uno de los estudiantes de Alice se niega a hacer su tarea, a pesar de que tanto ella como sus padres se lo dicen

Sesgo de confirmación
Tendemos a encontrar y recordar infortaciones que confirman nuestra percepción.

Puede confirmar una teoría de la conspiración basada en escasa evidencia mientras ignora la evidencia contraria

Efecto contraproducente
Refutar la evidencia de algo tiene el efecto injustificado de confirmar nuestras creencias.

La evidencia que refuta su teoría de la conspiración, la guerra probablemente falseada por el gobierno.

Efecto tercera persona
Creemos que los demás se ven más afectados por el consumo de los medios de comunicación que nosotros mismos.

"Los medios de comunicación te han lavado el cerebro claramente"

Sesgo de creencias
Juzgamos la fuerza de un argumento no por la fuerza con que apoya la conclusión, sino por lo plausible que es la conclusión en nuestras propias mentes.

Sally menciona su teoría de apoyo sobre tu teoría de la conspiración, que adoptas de todo corazón a pesar de que tiene muy poca evidencia al respecto.

Cascada disponible
Vinculado a nuestra necesidad de aceptación social, las creencias colectivas se vuelven más plausibles a través de la repetición pública

Una historia sobre hojas de afeitar que aparecen en los dulces finalmente llevó a que muchas personas ya no ofrecieran golosinas caseras en Halloween en Estados Unidos.

Declinismo
Tendemos a romantizar el pasado y ver el futuro de manera negativa, creyendo que las sociedades / instituciones están en general en declive.

"En mi época, las niños tenían más respeto".

Sesgo Status Quo
Tendemos a preferir que las cosas sigan igual; los cambios desde la línea de base se consideran una pérdida

A pesar de que el término de servicio de una aplicación invade la privacidad de Sally, ella prefiere no cambiar a otra aplicación.

Falacia de costos hundidos (o escalada del compromiso)
Invertimos más en cosas que nos han costado algo en lugar de alterar nuestras inversiones, incluso si enfrentamos resultados negativos.

"en por un centavo, en por una libra"

Falacia del jugador
Creemos que las posibilidades futuras se ven afectadas por eventos pasados.

Alice ha perdido nueve lanzamientos de monedas seguidos, ¡así que seguro que ganará el próximo!

Sesgo cero riesgo
Preferimos reducir el riesgo pequeño a cero, incluso si podemos reducir más riesgo en general con otra opción

0%

"Probablemente deberías por la garantía"

Efecto enmarcado
A menudo sacamos conclusiones diferentes de la misma información dependiendo de cómo se presente.

Alice escucha que uno de los candidatos favoritos lo está matando con un índice de aprobación del 45%. Sally escucha que el candidato está decepcionando al país con una calificación del 45%. Tienen interpretaciones muy diferentes de la misma estadística.

Estereotipos
Adoptamos creencias generalizadas de que los miembros de un grupo tendrán ciertas características a pesar de no tener información sobre el individuo.

"Ese tipo del bigote elegante es un hipster. Probablemente tenga una colección de vinilos"

Sesgo de homogeneidad de grupos externos
Percibimos a los miembros de fuera del grupo como homogéneos y tu propio grupo como más diverso.

Alice no es una jugadora, pero cree que "todos los jugadores son iguales".

Sesgo de autoridad
Confiamos y estamos más influenciados por las opciones de las figuras de la autoridad.

"Mi profesora me dijo que estaba bien"

Efecto placebo
Si creemos que un tratamiento funcionará, a menudo tendrá un pequeño efecto fisiológico.

A Alice le dieron un plaebo para su dolor, y su dolor disminuyó.

Sesgo de supervivencaia
Tendemos a enfocarnos en aquellas cosas que sobrevivieron a un proceso y pasamos por alto las que fallaron.

Greg le dice a Alice que el negocio de las carteras va a ser genial porque una empresa de moda exitosa tenía la misma estrategia. (Pero otras 10 empresas fallidas también tenían la misma estrategia)

Taquipsiquia
Nuestras percepciones del tiempo cambian según el trauna, el uso de drogas y el esfuerzo físico.

"cuando el auto casi me golpea, el tiempo se ralentizó"

Ley de la trivialidad (o desprendimiento de bicicletas)
Damos un peso desproporcionado a los problemas triviales, a menudo evitando problemas más complejos.

En lugar de averiguar cómo ayudar a las personas sin hogar, ya que el gobierno local de la ciudad pasa mucho tiempo discutiendo la instalación de bicisendas y cobertizos para bicicletas.

Efecto Zeigarnik
Recordamos las tareas incompletas más que las completadas

Greg se siente culpable por no haber hecho nada, hasta que ve todas las tareas que ha marcado en su lista de tareas.

Efecto IKEA
Le damos más valor a las cosas que creamos parcialmente nosotros mismos.

¿No te encanta esta taza en la que gasté $ 20? la pinto yo misma

Efecto Ben Franklin
Nos gusta hacer favores; es más probable que le hagamos otro favor a alguien si ya le hemos hecho un favor que si hubiéramos recibido un favor de esa persona.

Greg le prestó un bolígrafo a Francis. Cuando Francis pidió un préstamo de $ 5, Greg lo hizo fácilmente.

Efecto espectador
Cuantas más personas haya alrededor. es menos probable que ayudemos a una victima

En una multitud de estudiantes, nadie llamó al 911 cuando alguien resultó herido en una pelea

Sugestión
Nosotros, especialmente los niños, a veces confundimos ideas sugeridas por preguntas con recuerdos.

Entonces, ¿te caíste del sofá antes o después de que tu mamá te golpeara?

Memoria falsa
Confundimos la imaginación con recuerdos reales

Greg está seguro de que Sally dijo un chiste muy divertido sobre las piñas, cuando ese chiste en realidad salió de una televisión.

Criptomnesia
Confundimos los recuerdos reales con la imaginación

Greg cree que visitó un cementerio, pero está bastante seguro de que acaba de tener

Ilusión de agrupamiento
Encontramos patrones y grupos en datos aleatorios.

"Esa nube se parece a tu gata, Alice"

Sesgo de pesimismo
A veces sobrestimamos la probabilidad o los malos resultados

"Nada mejorará jamás"

Sesgo de optimismo
A veces somos demasiado optimistas acerca de los buenos resultados

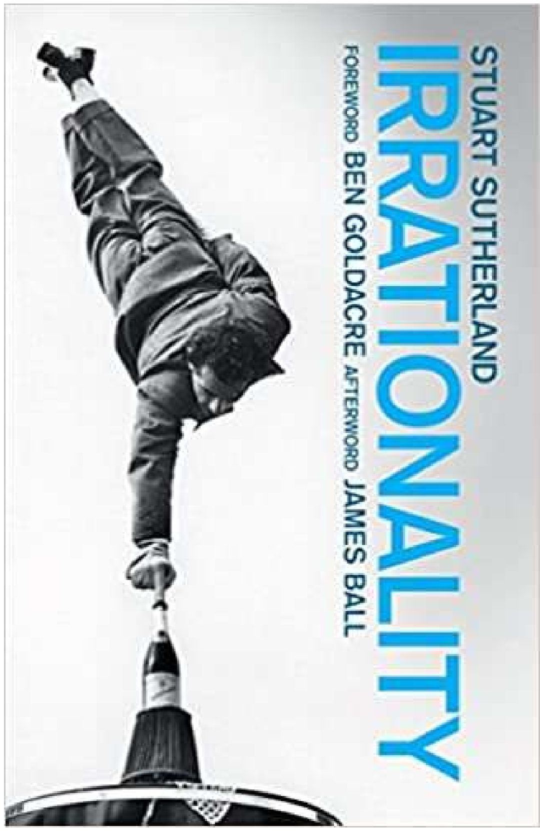
"Va a resultar genial"

Sesgo de punto ciego
No creemos que tengamos prejuicios, y lo vemos en los demás más que en nosotros mismos.

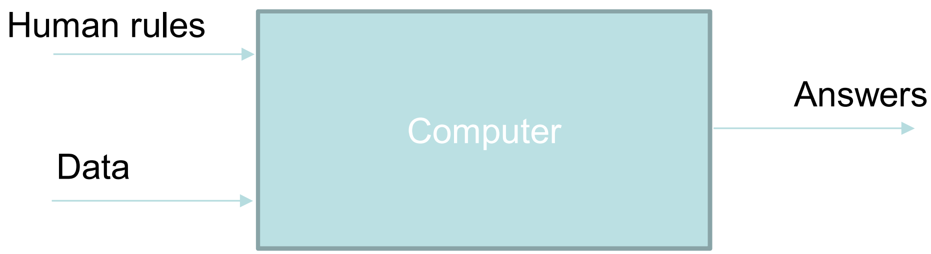
"no soy parcial"

Fundamental causes of human irrational behavior

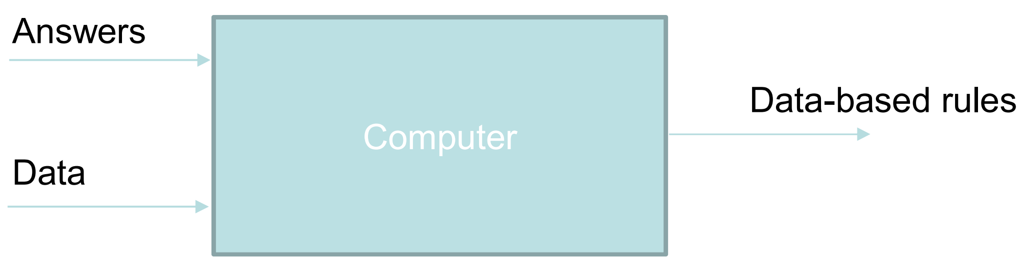
- Evolutionary causes: fear of loosing esteem, groupthink, conformism, shame,...
- Physiological causes: when we learn, some neurons connect but some other existing connections weaken
- Mental sloth
- Unability to manage many variables simultaneously
- Unability to handle basic concepts of probability calculus
- Self-interests



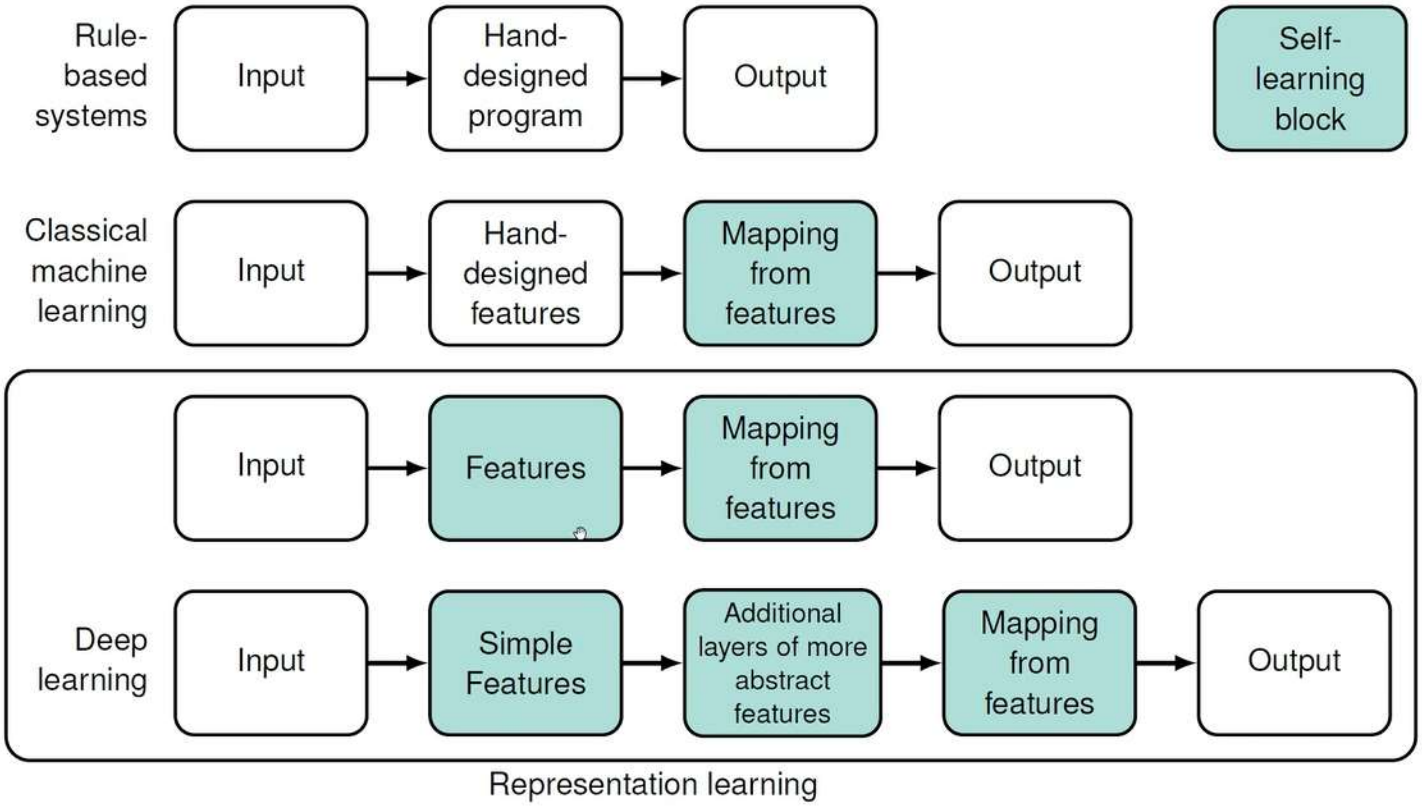
Traditional use of computers



Machine learning at training phase



Differences between AI/ML/DL



Some aplications benefiting from decision taking based on data analysis...

- Prediction

  Medical diagnose, product recommendation (Amazon, Netflix), future values of stock market or currency exchange rates,...

- Pattern classification

  Voice, faces, facial expressions, objets in real scenes, written characters, biometric identification, spam detection,...

- Data mining

  Clicks on webpages, searches on Google, ...

- Anomaly detection

  Abnormal credit card transactions, anomalous recordings of sensors in a nuclear plant,...

- Understanding of the brain behavior

Three types of learning

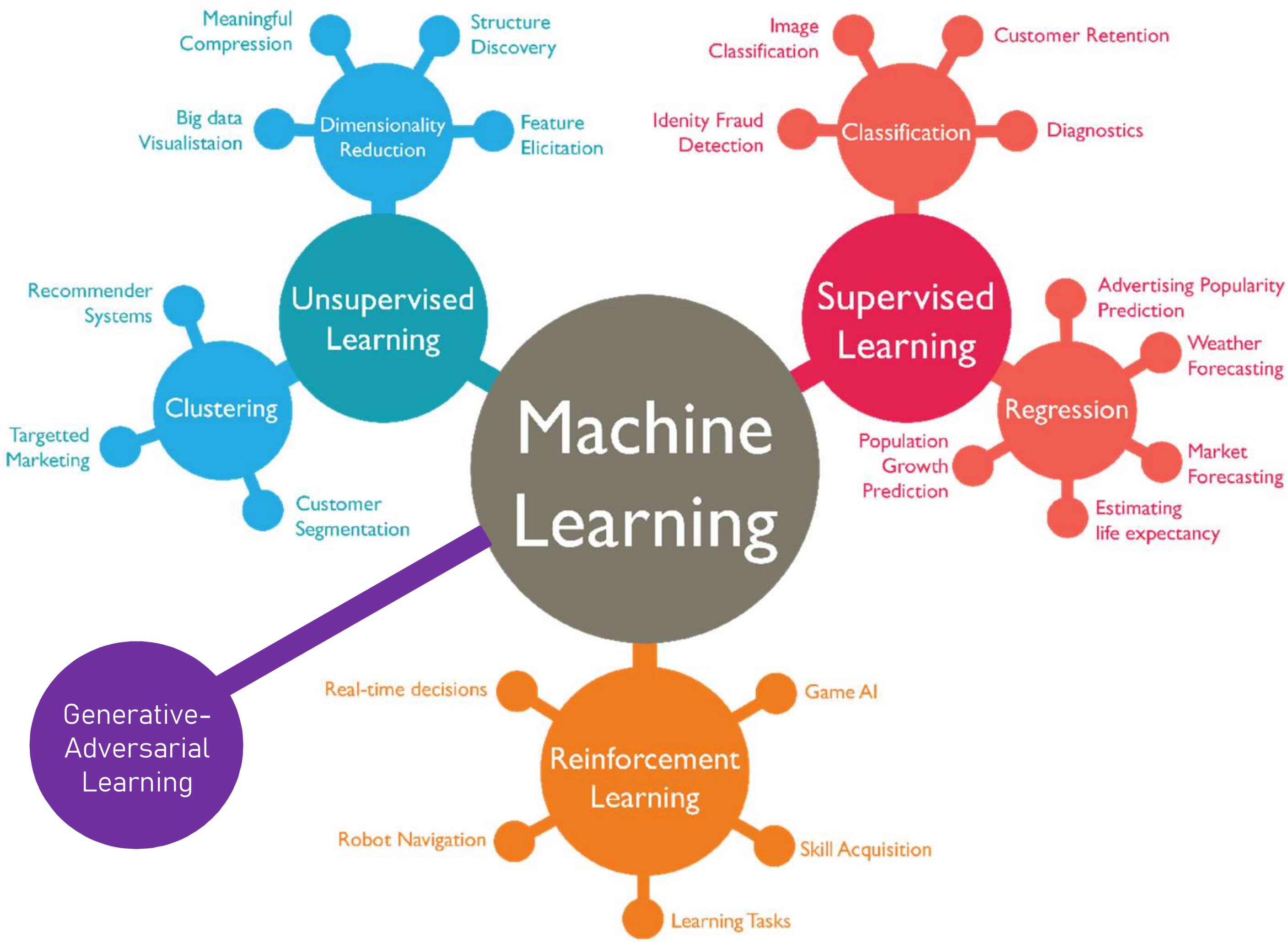
We want to build a machine that learns how to take decisions out of sensorial data (feature vectors) $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$,...

There are three main approaches to machine learning:

1. **Supervised learning**: the learning rule is provided with the desired outputs y_1, y_2, y_3, y_4 ,... for each given input, and the objective is generating the right output when fed with a new input.

2. **Unsupervised learning**: a model is built for the values of $\mathbf{x}$ that can be used to explain some phenomenon, to understand the structure of data, to predict, to communicate, etc.

3. **Reinforcement learning**: the machine decides some actions a_1, a_2, a_3,... affecting the environment, and gets some reward (or punishment) r_1, r_2, r_3,... The objective is learning to act in a way that long term accumulated rewards are maximized.


Machine Learning
Unsupervised Learning
Dimensionality Reduction
Meaningful Compression
Structure Discovery
Big data Visualistaion
Feature Elicitation
Clustering
Recommender Systems
Targetted Marketing
Customer Segmentation
Supervised Learning
Classification
Image Classification
Customer Retention
Idenity Fraud Detection
Diagnostics
Regression
Advertising Popularity Prediction
Weather Forecasting
Population Growth Prediction
Market Forecasting
Estimating life expectancy
Reinforcement Learning
Real-time decisions
Game AI
Robot Navigation
Skill Acquisition
Learning Tasks
Generative-Adversarial Learning

Description of the supervised learning...

Data base

- Feature vectors (observations): $\mathbf{x}_i \quad i=1,\ldots, N$
- Each belonging to one of the c classes (nature states): y_i

Objective

Design a procedure that learns to associate $\mathbf{x}_i$ and y_i through the function $h_{\boldsymbol{\theta}}$

$$\hat{y}_j = h_\theta(\mathbf{x}_j)$$

that depends on some parameters $\boldsymbol{\theta}$, using the data base and optimising some loss function

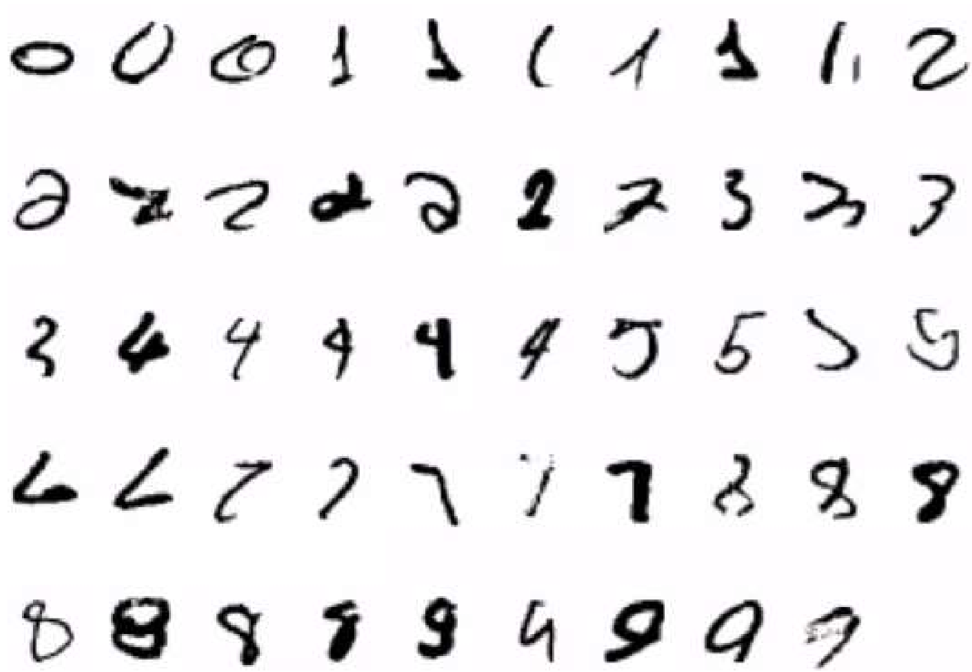
$$\theta = \arg\min_\theta L\left(y, h_\theta(\mathbf{x})\right)$$

Conditioning

Is the statistics of vectors known?

What are the criteria used to determine the procedure that learns $h_{\boldsymbol{\theta}}$?

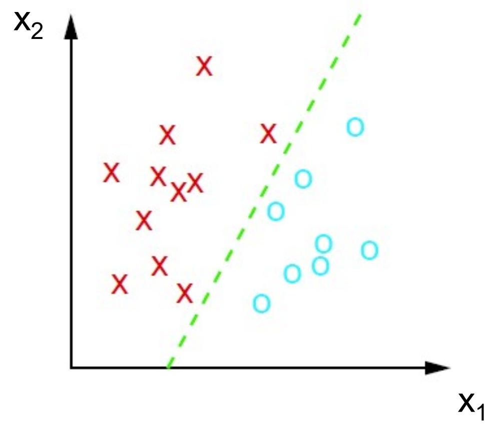
An example: define a machine that identifies handwritten digits between 0 and 9?



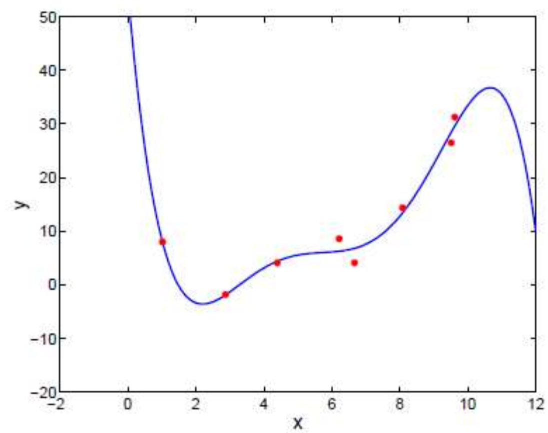
The decision algorithm will select one or another digit without explicit rules, but learning from examples. What are in this case $\mathbf{x}_i$ and y_i?

Two main types of supervised learning...

Classification
The desired outputs y_i are discrete class labels. The goal is to classify new inputs $\mathbf{x}$ correctly.



Regression
The desired outputs y_i are continous valued. The goal is to predict the output correctly for new inputs.



Practical examples

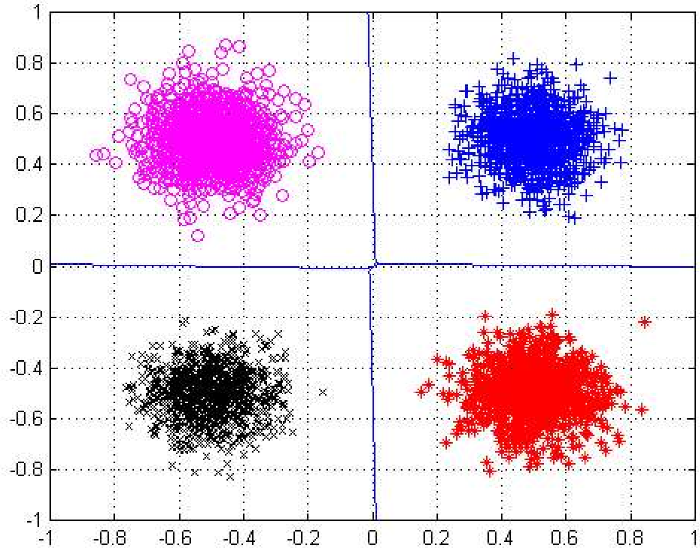
Example 1: Symbol detection in a Gaussian communication channel

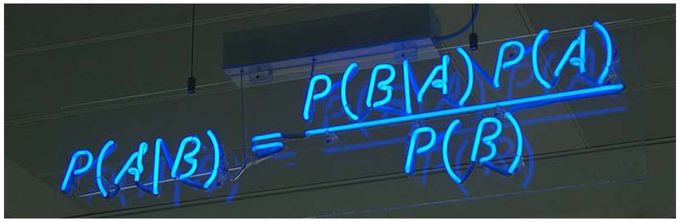
Processing:

Down-conversion and matched filters

Feature extraction: I/Q Sampling

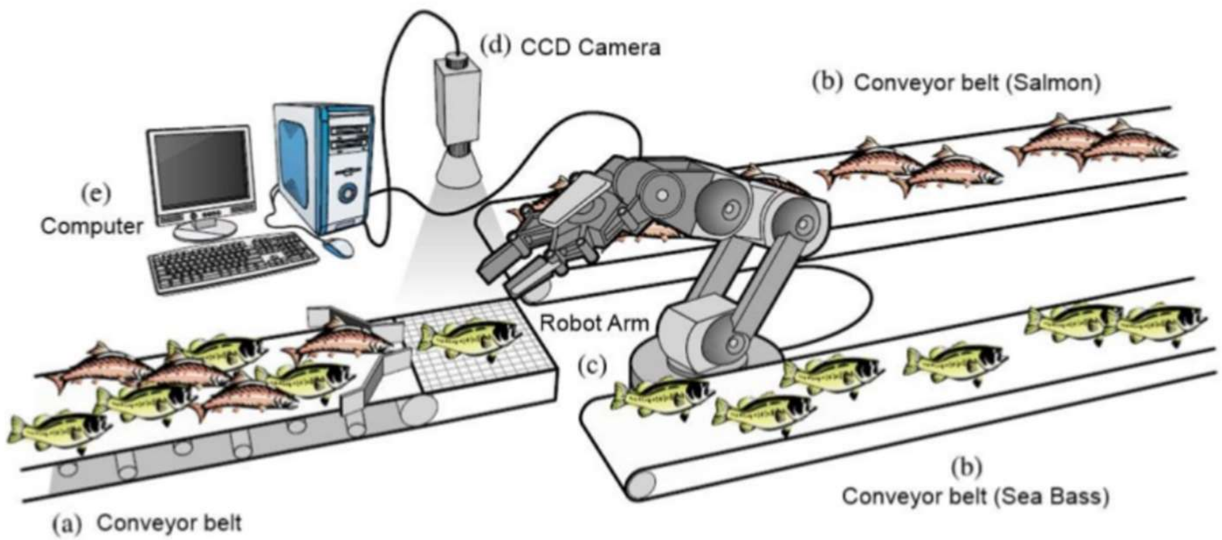
Decision: MAP detection





Example 2: Classification of salmons or sea bass

Features are extracted the CCD image (e.g. brightness, length, width) and a computer takes decisions that are executed by the robot.



Example 3: Diagnose of heart disease

Data base SHEART

Feature vectors (blood analitics, tobacco, family history, obesity, alcohol consumption, age, etc.) can be used to

- Predict the risk of a heart attack onset.
- Identify the most relevant features for prediction.

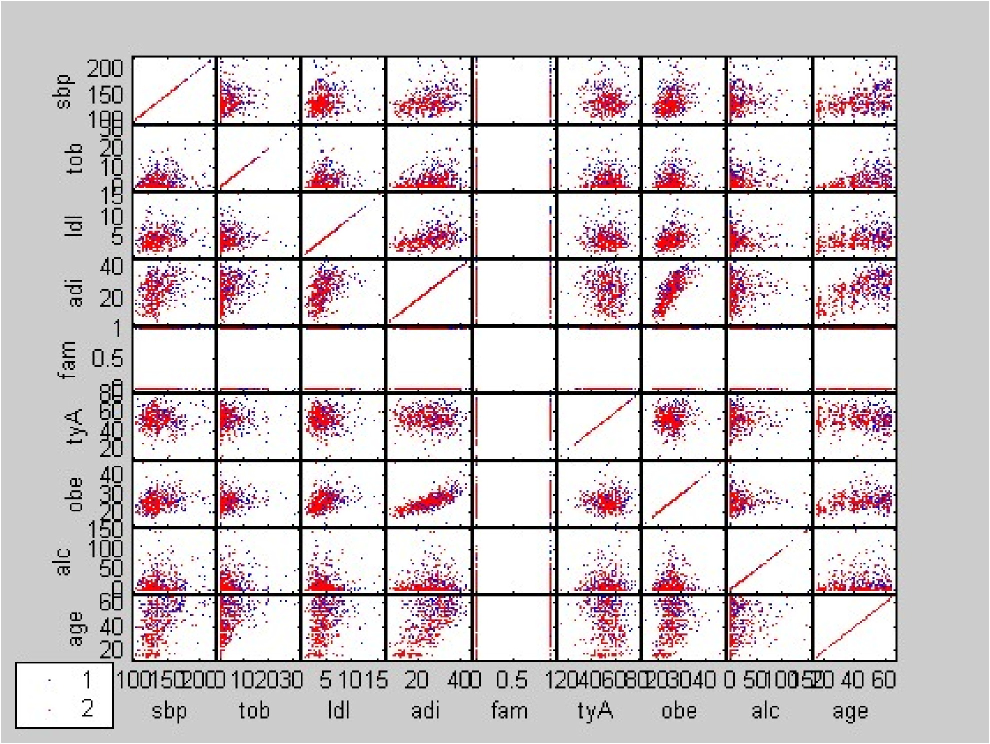
SHEART. Sample of males in a heart-disease high-risk region of the Western Cape, South Africa.

• sbp	systolic blood pressure
• tobacco	cumulative tobacco (kg)
• ldl	low density lipoprotein cholesterol
• adiposity	
• famhist	family history of heart disease (Present, Absent)
• typea	type-A behavior
• obesity	
• alcohol	current alcohol consumption
• age	age at onset
• chd	response, coronary heart disease

sbp	tob	ldl	Adip.	fa	A	Ob.	alc	age
160	12	5.73	23.11	1	49	25.3	97.2	52

Scatter plot of SHEART

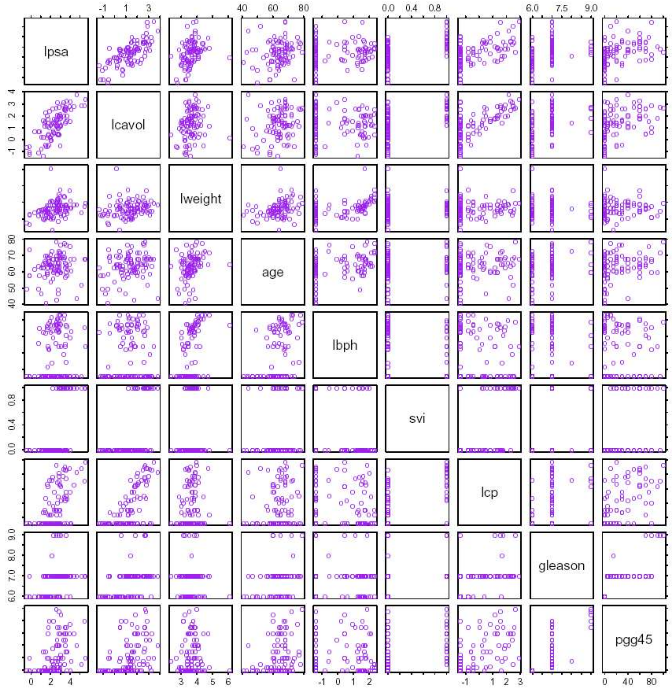
None of the features is Gaussian!



Example 3: Diagnose of prostate cancer

Search for correlation between the prostate specific antigen (PSA) and variables that may help in predicting the cancer condition:

- log cancer volume (lcavol)
- log prostate weight (lweight)
- age
- log benign prostatic hyperplasia (lbph)
- seminal vesicle invasion (svi)
- Gleason score (gleason)
- percent of Gleason scores 4 or 5 (pgg45)



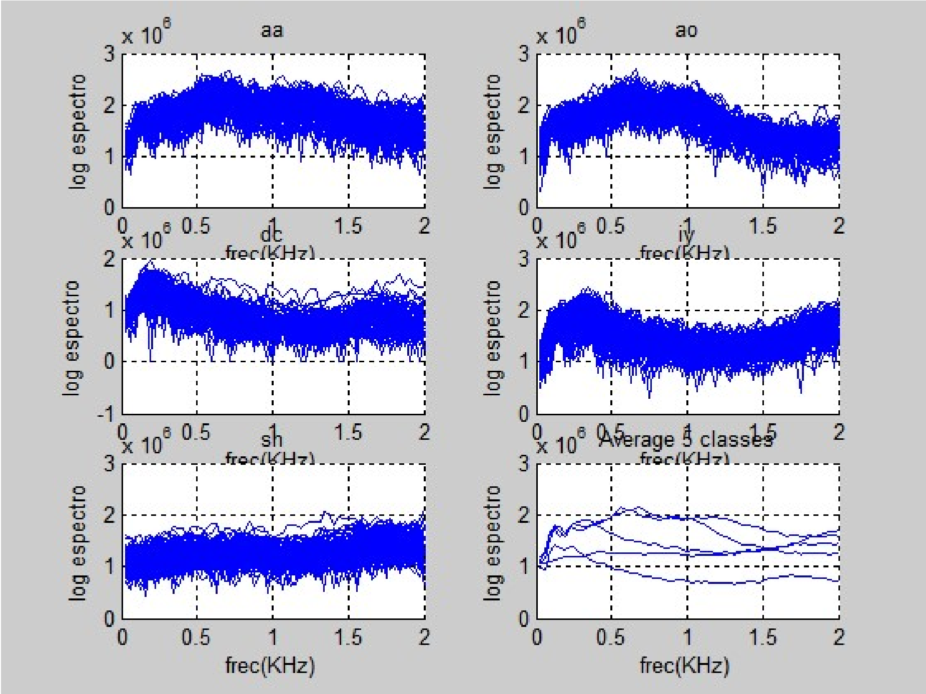
Example 4: PHONEME data base

Sampling freq. = 8KHz

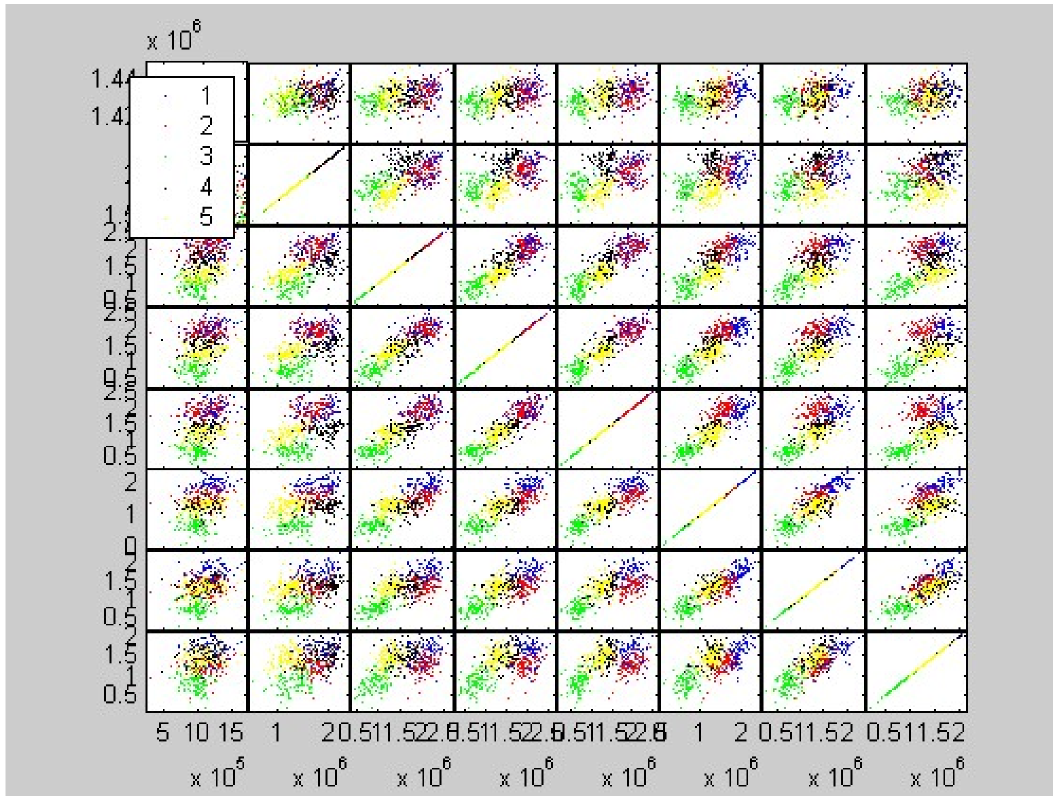
Log Spectrum

aa	ao	dcl	iy	sh
695	1022	757	1163	872

Feature vector size = 64



Scatter plot



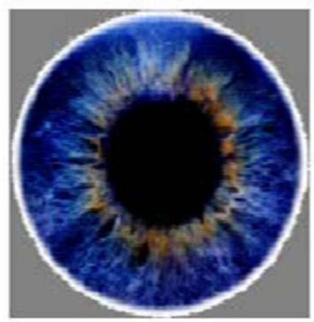
Example 5: Biometrics

(a) Fingerprint

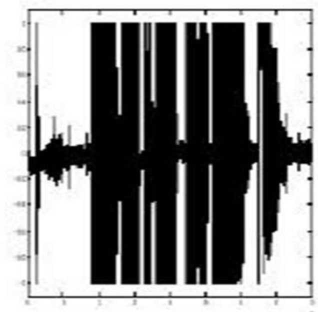
(b) Face

(c) Hand Geometry

(d) Signature



(e) Iris



(f) Voice

Comparison of biometrical identification methods

	Ojo - Iris	Ojo - Retina	Huellas dactilares	Geometria de la mano	Escritura - Firma	Voz
Fiabilidad	Muy alta	Muy alta	Alta	Alta	Alta	Alta
Facilidad de uso	Media	Baja	Alta	Alta	Alta	Alta
Prevención de ataques	Muy Alta	Muy alta	Alta	Alta	Media	Media
Aceptación	Media	Media	Media	Alta	Muy alta	Alta
Estabilidad	Alta	Alta	Alta	Media	Media	Media
Identificación y autenticación	Ambas	Ambas	Ambas	Autenticación	Ambas	Autenticación
Estándars	-	-	ANSI/NIST, FBI	-	-	SVAPI
Interferencias	Gafas	Irritaciones	Suciedad, heridas, asperezas ...	Artritis, reumatismo ...	Firmas fáciles o cambiantes	Ruido, resfriados ...
Utilización	Instalaciones nucleares, servicios médicos, centros penitenciarios	Instalaciones nucleares, servicios médicos, centros penitenciarios	Policía, industrial	General	Industrial	Accesos remotos en bancos o bases de datos

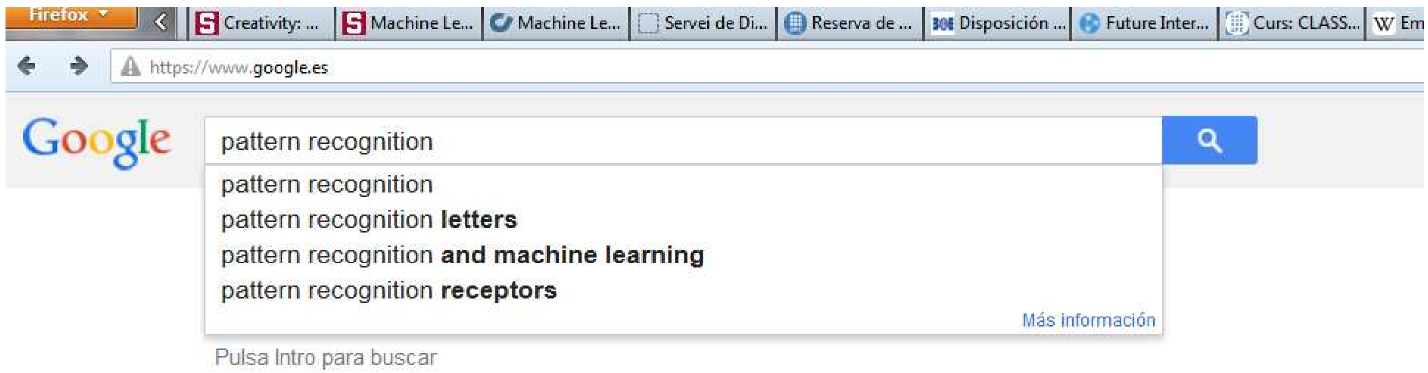
Example 6: Classification of spam email

Text file → Statistics of words/characters → Clasification

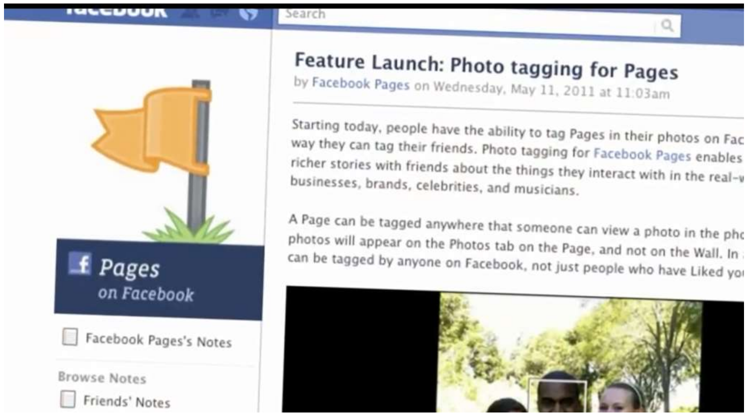
```
word_freq_make:         continuous.
word_freq_address:      continuous.
word_freq_all:          continuous.
word_freq_3d:           continuous.
word_freq_our:          continuous.
word_freq_over:         continuous.
word_freq_remove:       continuous.
word_freq_internet:     continuous.
word_freq_order:        continuous.
word_freq_mail:         continuous.
word_freq_receive:      continuous.
word_freq_will:         continuous.
word_freq_people:       continuous.
word_freq_report:       continuous.
word_freq_addresses:    continuous.
word_freq_free:         continuous.
```

```
word_freq_business:     continuous.
word_freq_email:        continuous.
word_freq_you:          continuous.
word_freq_credit:       continuous.
word_freq_your:         continuous.
word_freq_font:         continuous.
word_freq_000:          continuous.
word_freq_money:        continuous.
word_freq_hp:           continuous.
word_freq_hpl:          continuous.
word_freq_george:       continuous.
word_freq_650:          continuous.
word_freq_lab:          continuous.
word_freq_labs:         continuous.
word_freq_telnet:       continuous.
word_freq_857:          continuous.
word_freq_data:         continuous.
word_freq_415:          continuous.
word_freq_85:           continuous.
word_freq_technology:   continuous.
word_freq_1999:         continuous.
word_freq_parts:        continuous.
word_freq_pm:           continuous.
word_freq_direct:       continuous.
word_freq_cs:           continuous.
word_freq_meeting:      continuous.
word_freq_original:     continuous.
word_freq_project:      continuous.
word_freq_re:           continuous.
word_freq_edu:          continuous.
word_freq_table:        continuous.
word_freq_conference:   continuous.
char_freq_;:            continuous.
char_freq_(:            continuous.
char_freq_[:            continuous.
char_freq_!:            continuous.
char_freq_$:            continuous.
char_freq_#:            continuous.
capital_run_length_average: continuous.
capital_run_length_longest: continuous.
capital_run_length_total:   continuous.
1, 0.    | spam, non-spam classes
```

Example 7: Browsers



Example 8: Face recognition(Facebook), smile recognition(photo cameras)

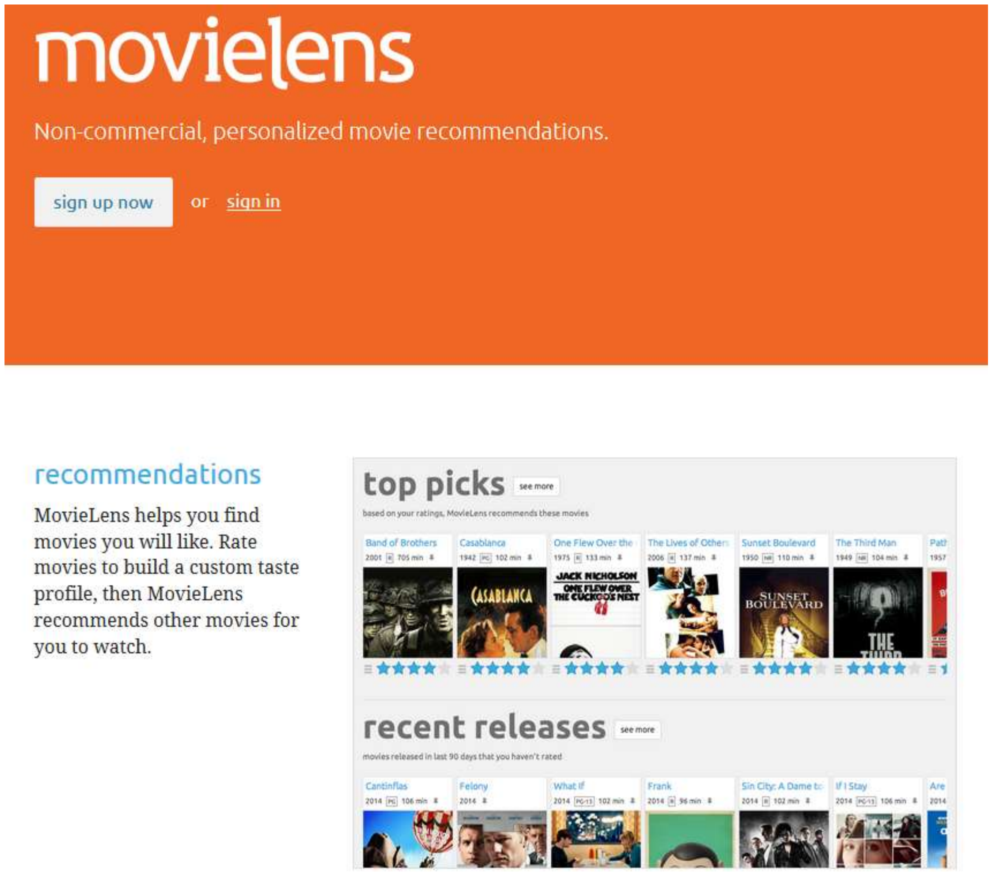


Audio clips (Shazam)

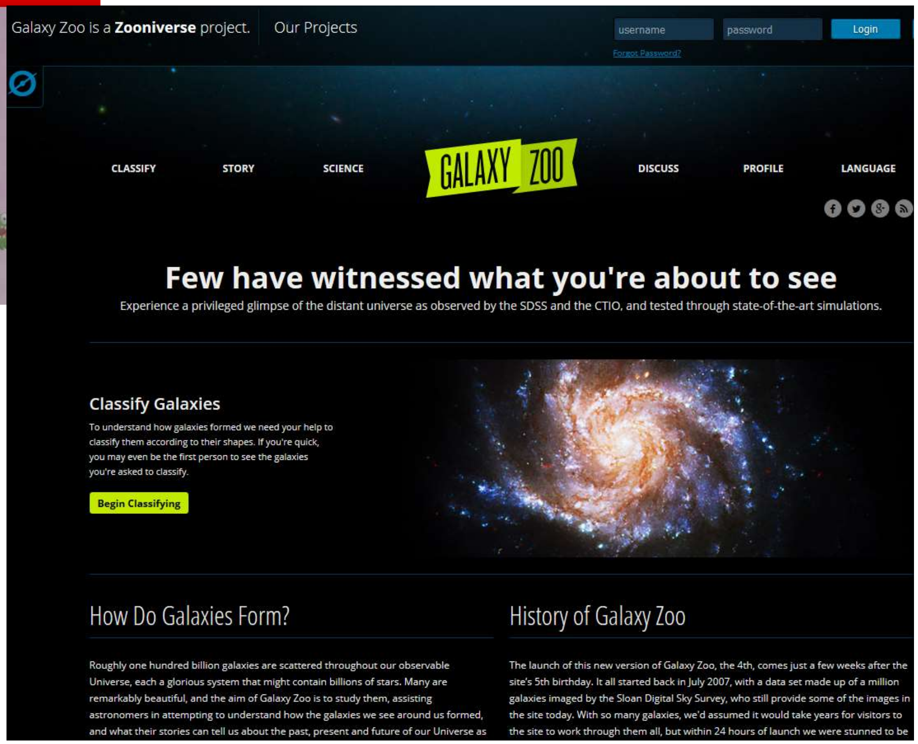
Example 9: Robotics, recognition of movements

Example 10: Reinforcement learning

Example 11: Recommendation systems (Amazon, Netflix)



Example 12: Crowdsourcing



Example 12: Natural language applications

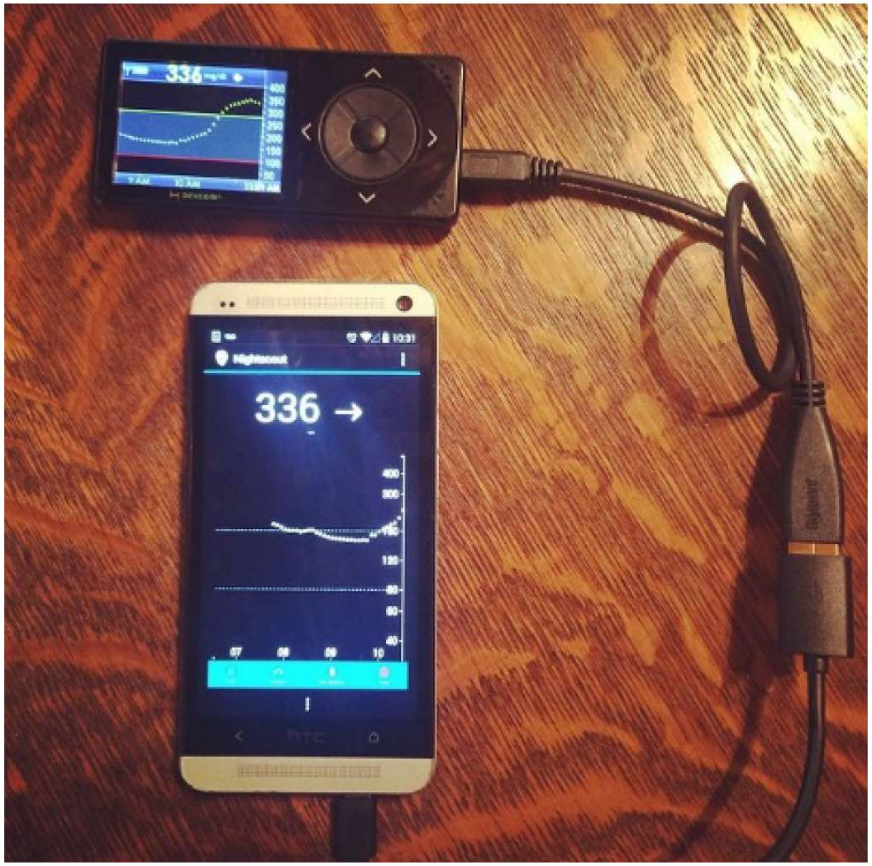
Microsoft
Translator

Google's Neural Machine Translation

...reduced translation errors by an average of 60% when compared to the prior Google Translate technology

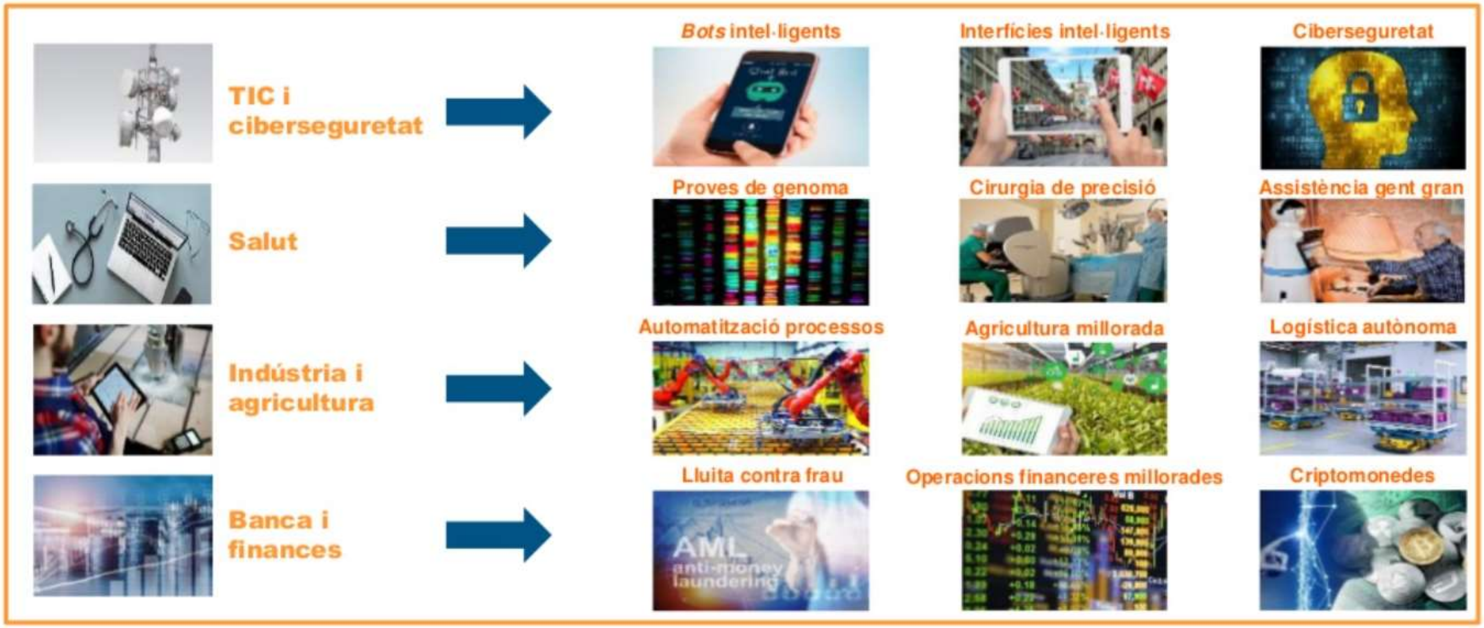
Amazon's Alexa

Example 13: Forecast of hipoglucemia on mobile devices

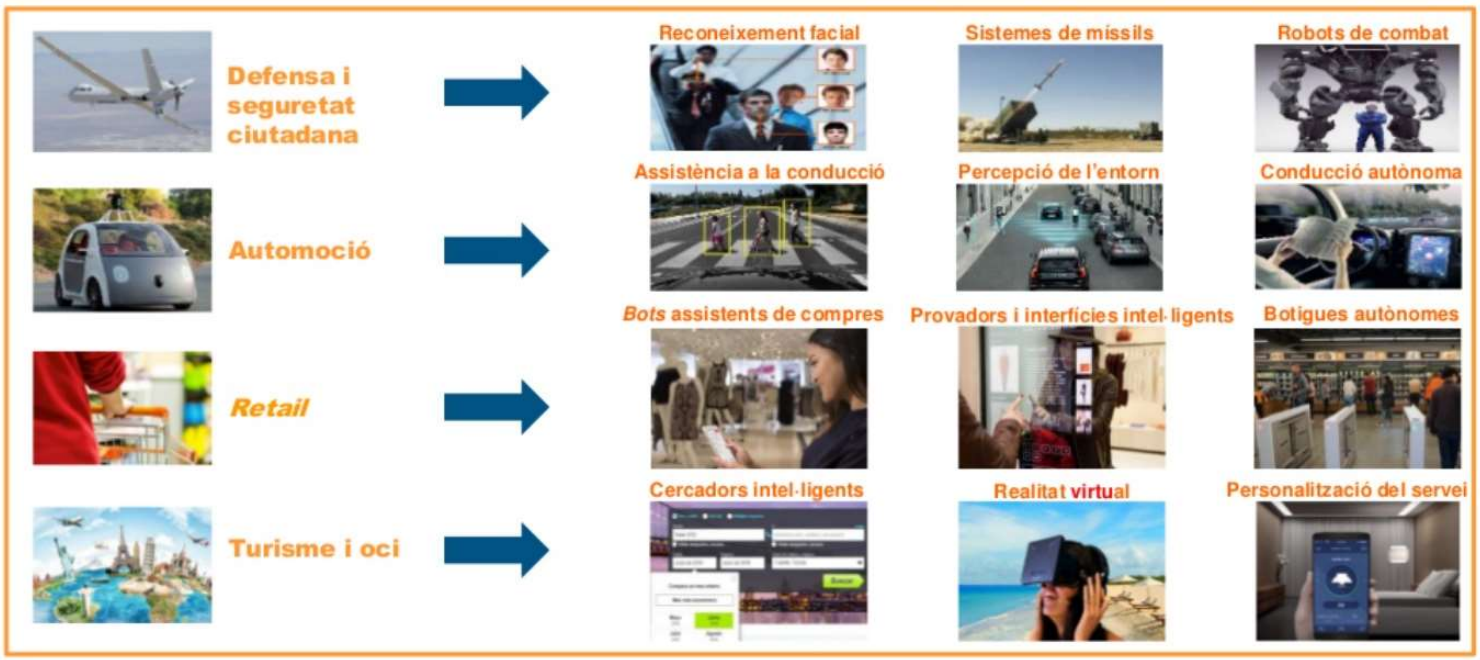


Example 14: Generative intelligence

Some applications per demand sector



Some applications per demand sector

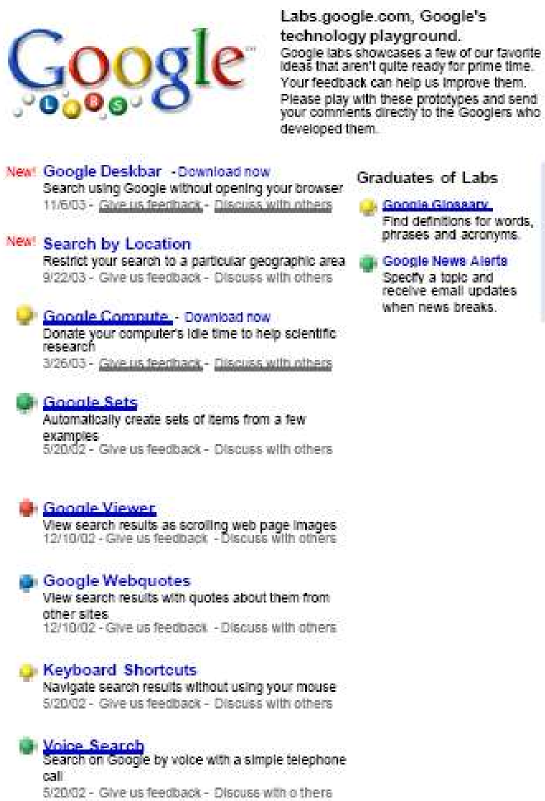


Main AI companies...

Almost any Internet-based company use machine learning: the business is no longer in owning the physical asset, but on how to manage it smartly.

Main world-wide investors

Jobs offers



Passionate about these topics? You should work at Google.

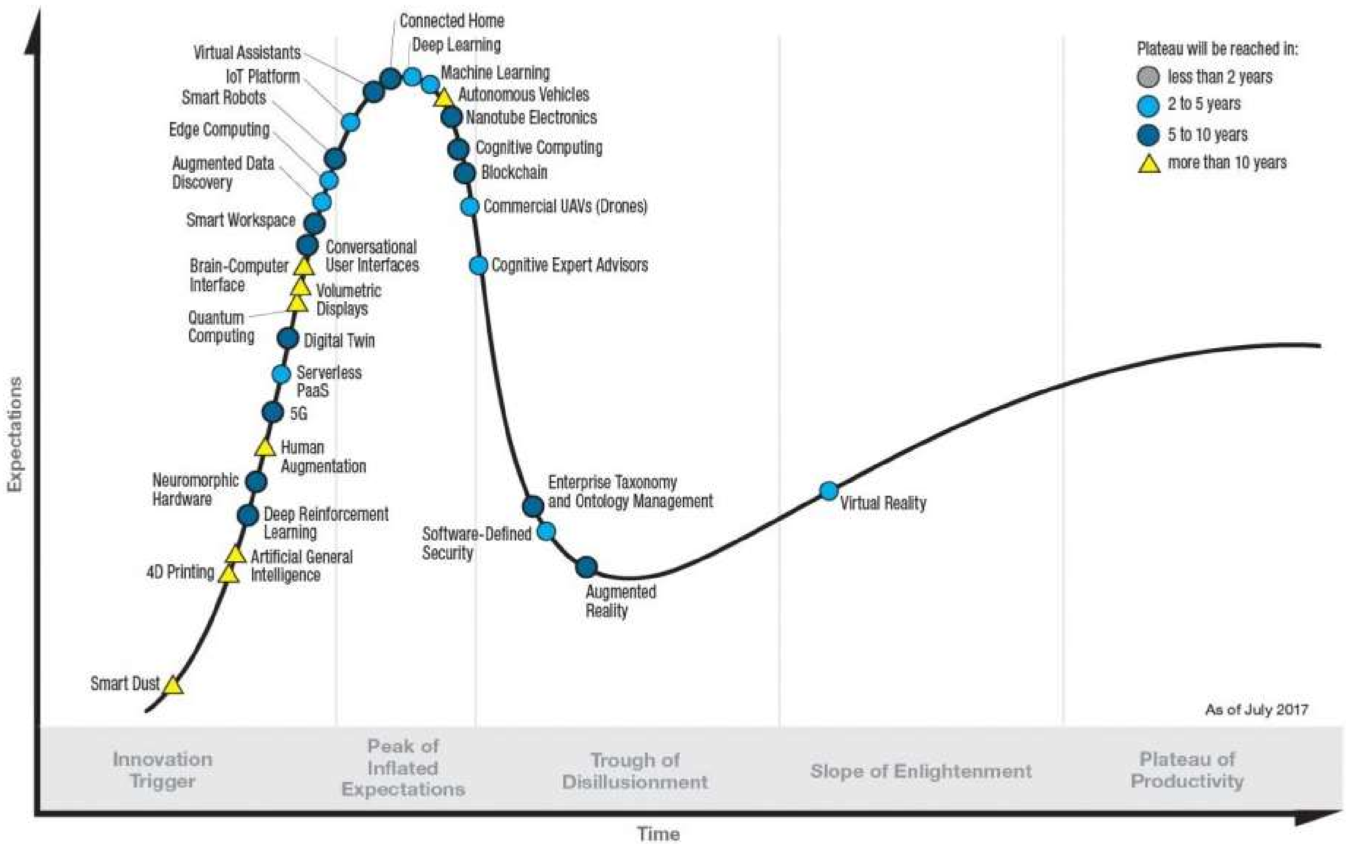
- algorithms
- artificial intelligence
- compiler optimization
- computer architecture
- computer graphics
- data compression
- data mining
- file system design
- genetic algorithms
- information retrieval
- machine learning
- natural language processing
- operating systems
- profiling
- robotics
- text processing
- user interface design
- web information retrieval
- and more!

Send your resume and a brief cover letter to great-engineers@google.com

Data Science Jobs Board

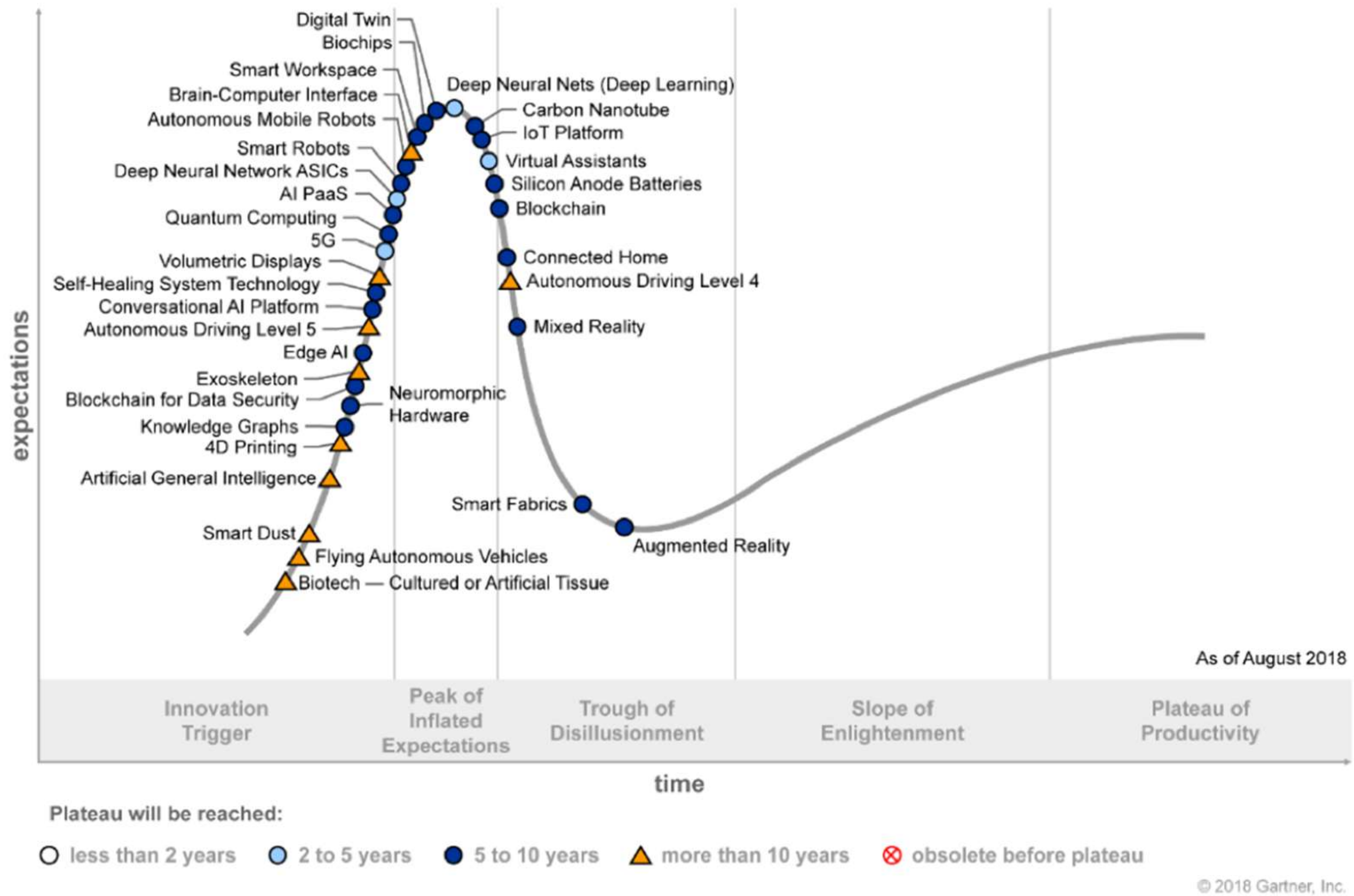
https://www.kaggle.com/jobs

Cycles for new technologies (2017)

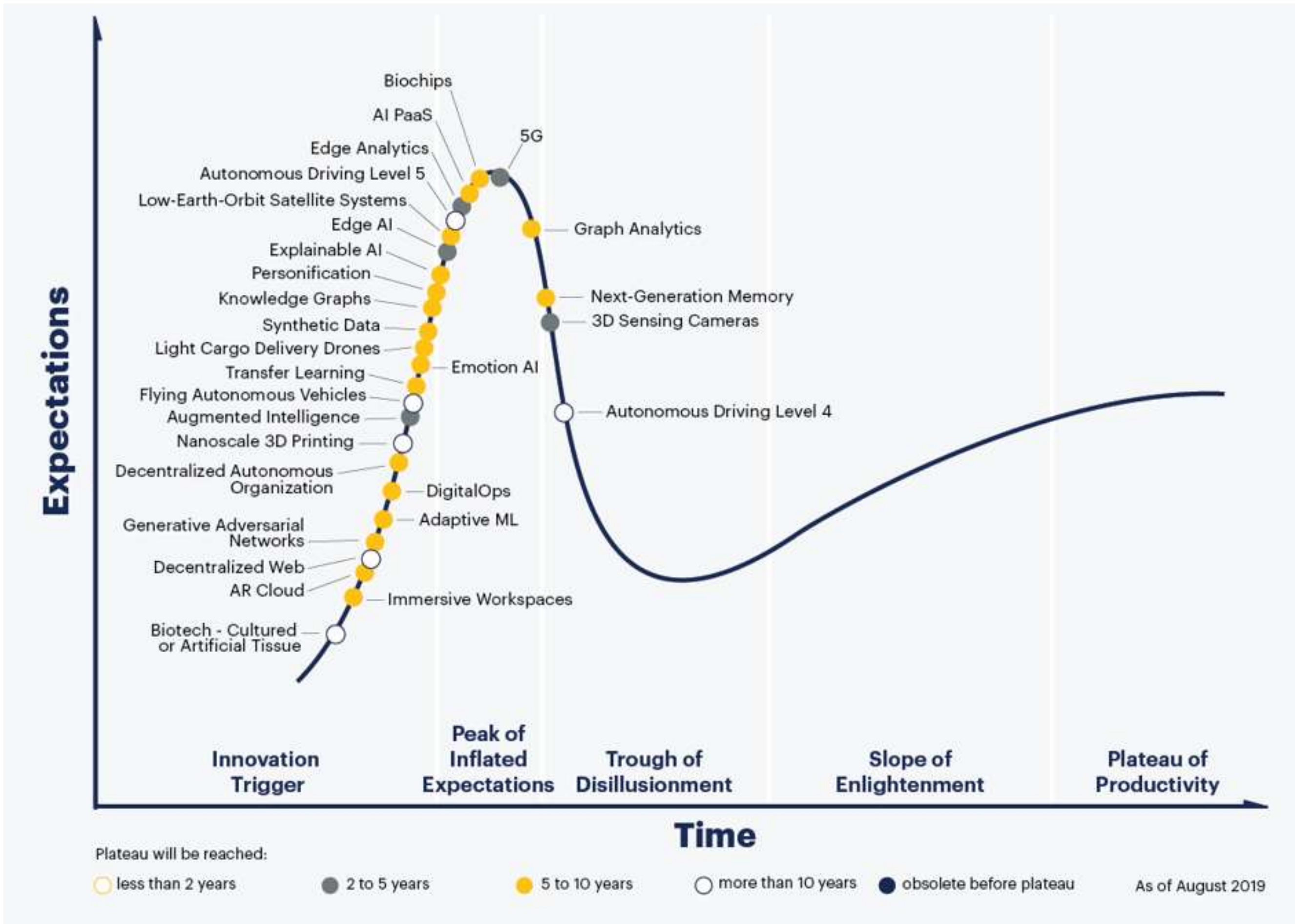


Source: https://www.gartner.com/smarterwithgartner/top-trends-in-the-gartner-hype-cycle-for-emerging-technologies-2017/

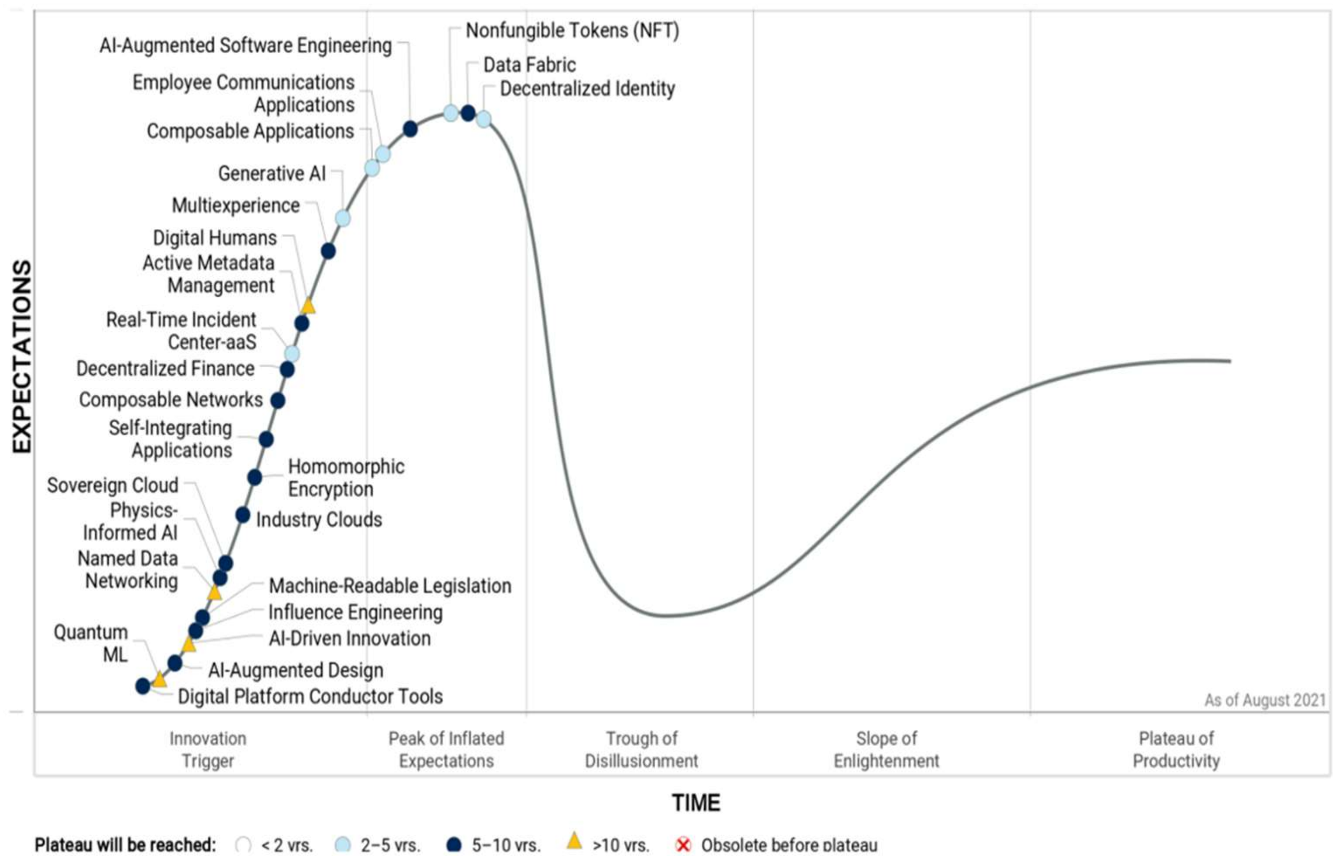
Cycles for new technologies (2018)



Cycles for new technologies (2019)

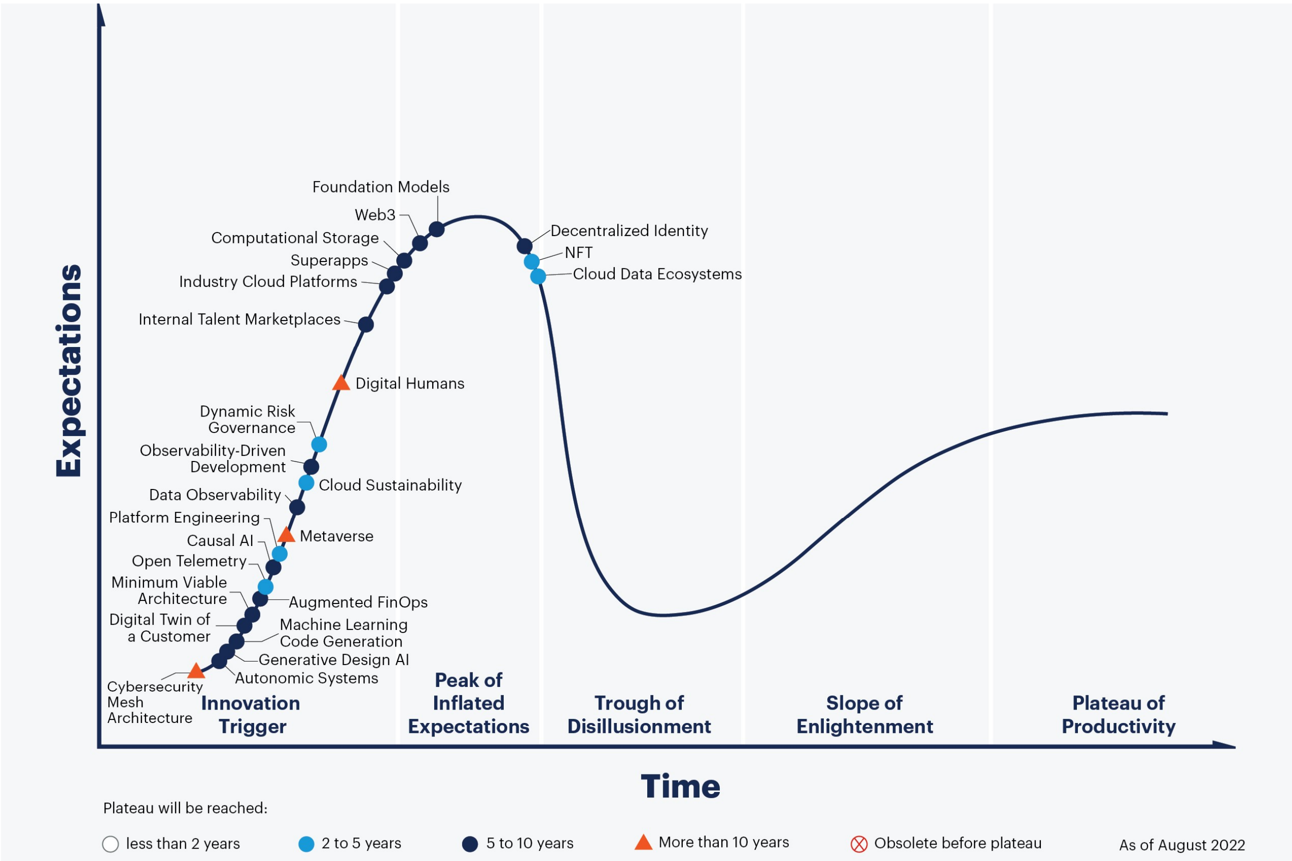


Cycles for new technologies (2021)

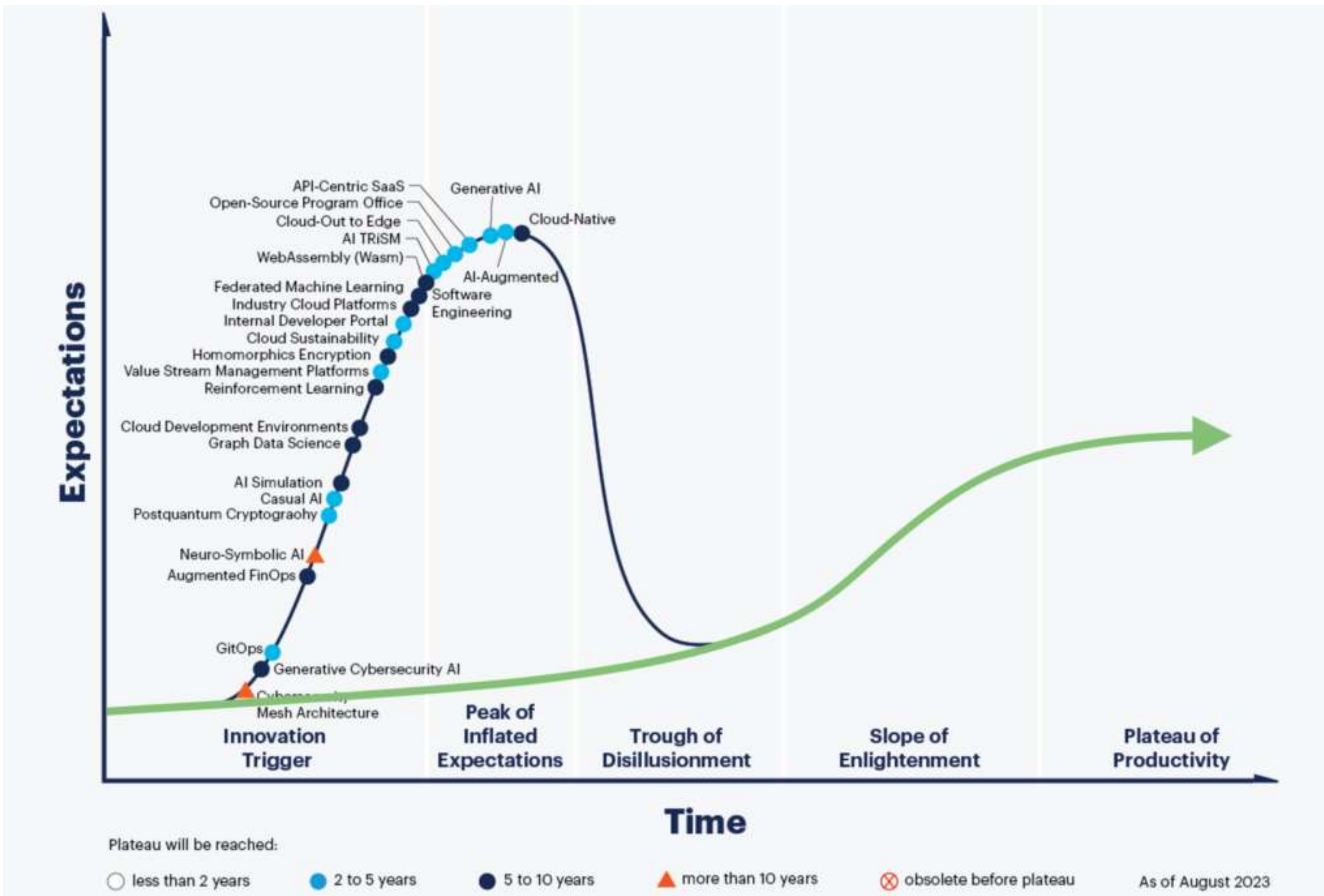


Source: Gartner (August 2021)

Cycles for new technologies (2022)

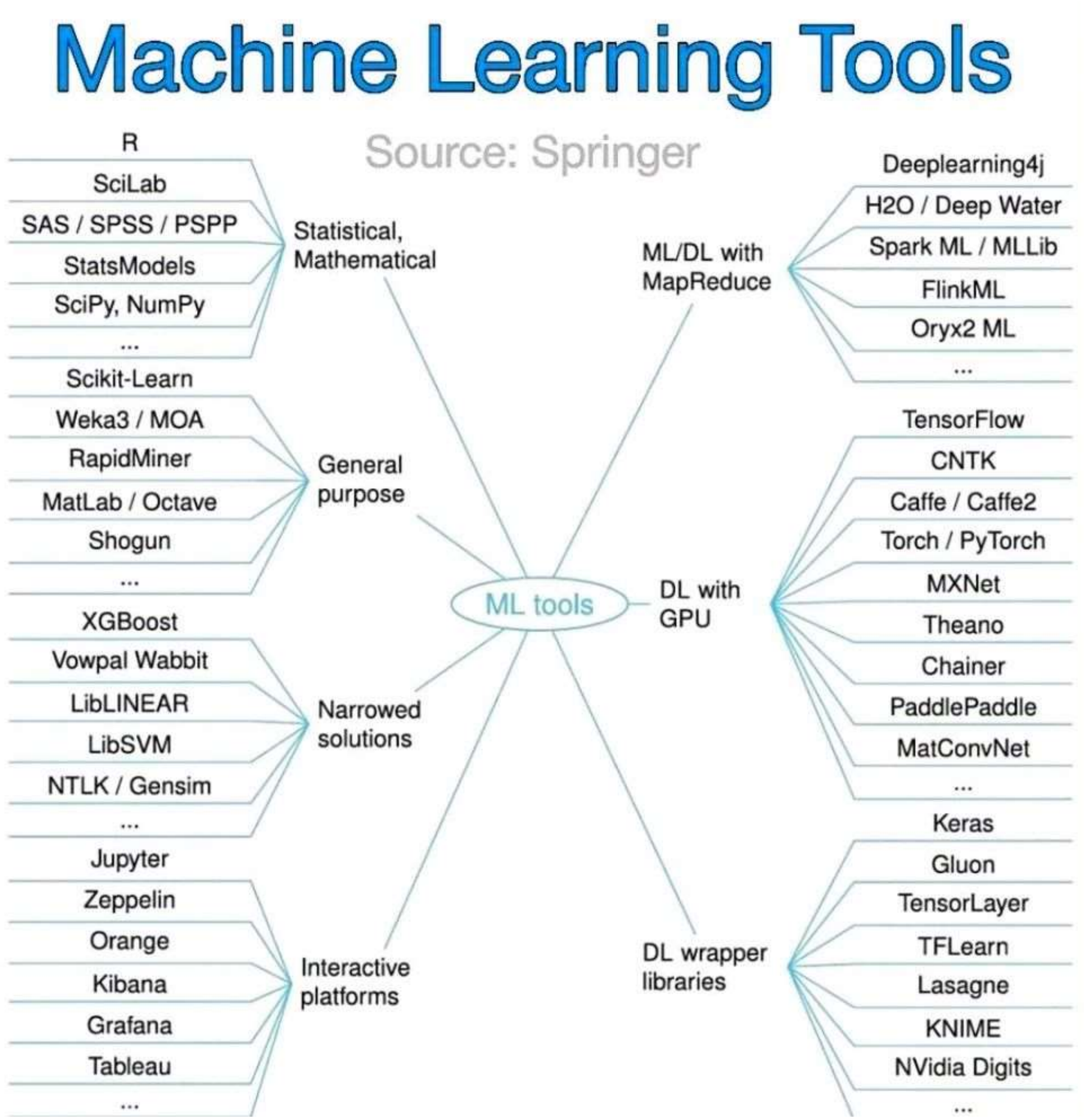


Cycles for new technologies (2023)



Other valuable skills in data science engineering (that will not be covered in this course)

- R/Python programming languages
- Deep learning tools (TensorFlow, Theano, Caffe)
- SQL and MySQL
- Data visualization tools (Tableau, Spotfire, Carto...)
- Data warehousing (Sparkle, Hadoop, Pig, Hive, MapReduce, Flume)


Machine Learning Tools
Source: Springer
ML tools
Statistical, Mathematical
R
SciLab
SAS / SPSS / PSPP
StatsModels
SciPy, NumPy
...
General purpose
Scikit-Learn
Weka3 / MOA
RapidMiner
MatLab / Octave
Shogun
...
Narrowed solutions
XGBoost
Vowpal Wabbit
LibLINEAR
LibSVM
NTLK / Gensim
...
Interactive platforms
Jupyter
Zeppelin
Orange
Kibana
Grafana
Tableau
...
ML/DL with MapReduce
Deeplearning4j
H2O / Deep Water
Spark ML / MLLib
FlinkML
Oryx2 ML
...
DL with GPU
TensorFlow
CNTK
Caffe / Caffe2
Torch / PyTorch
MXNet
Theano
Chainer
PaddlePaddle
MatConvNet
...
DL wrapper libraries
Keras
Gluon
TensorLayer
TFLearn
Lasagne
KNIME
NVidia Digits
...

A posible software architecture for data analysis

Course overview

Pre-processing

Feature selection

Classification

Techniques
PCA (lecture 3) MDA (lecture 3)

	Density function for classes	Classifier
Supervised	Parametrized Gaussian	Linear discriminant Quadratic discriminant[1] (lecture 2)
	Non-parametric	K-nearest neighbours[2] Fuzzy logic (lecture 4)
	Unknown	SVM[3] Neural networks[3] Decision trees[4] (lecture 4)
Unsupervised	Parametrized Gaussian	K-means (lecture 6)
	Unknown	Clustering (lecture 6)

1. Useful only if covariance matrices are not rank deficient.
2. Useful with the number of features is very large, even larger that the number of training vectors.
3. Imposes a structure to the classifier irrespective of the training data base.
4. Useful when non-numeric features are present.

Resources

Atenea: Handouts, Exercises, Practice in lab

Labs: Software in Python

Data bases

Awesome Public Datasets: https://github.com/awesomedata/awesome-public-datasets
Dept. of Statistics at Stanford Univ.: https://web.stanford.edu/~hastie/ElemStatLearn/
http://archive.ics.uci.edu/ml/
https://web.stanford.edu/~hastie/ElemStatLearn/data.html
http://www.cmaterju.org/cmaterdb
https://www.kaggle.com/datasets
https://physionet.org/physiobank/database/

Open competitions

http://www.kdd.org/kdd-cup
http://www.kaggle.com/competitions
https://datahack.analyticsvidhya.com/contest/all/
https://signalprocessingsociety.org/get-involved/signal-processing-cup

MOOC

https://www.coursera.org/learn/machine-learning

References

[**Duda, 2001**] R. O. Duda, P. E. Hart, D. G. Stork. "Pattern Classification", Ed. Wiley Interscience, 2002.

[**Bishop, 2006**] C. M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006.

[**Goodfellow, 2016**] I. Goodfellow et al, "Deep Learning", The MIT Press, Cambridge, MA, 2016

[**Hastie, 2001**] T. Hastie et al, "The Elements of Statistical Learning Springer", Verlag, 2001

[**Theodoridis, 2015**] S. Theodoridis, "Machine Learning", Academic Press, 2015

[**Bishop, 2024**] C. M. Bishop, "Deep Learning: Foundations and Concepts", Springer, 2024.

[**Murphy, 2012**] K. P. Murphy, "Machine Learning. A Probabilistic Perspective", The MIT Press, Cambridge, MA, 2012

[**Kuncheva, 2004**] L. I. Kuncheva, "Combining Pattern Classifiers: Methods and Algorithms", John Wiley, 2004