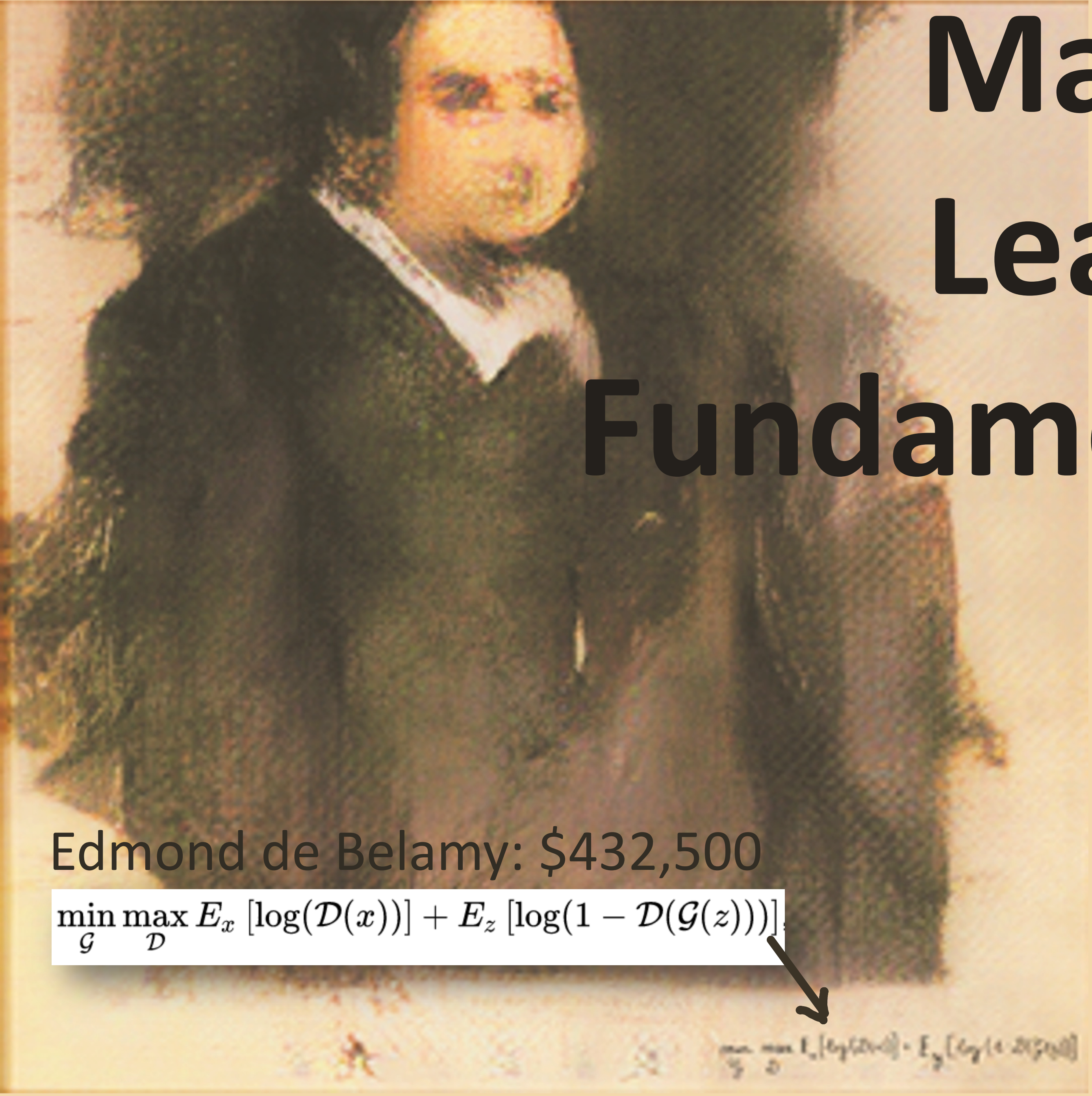
Machine Learning Fundamentals

Instituto Balseiro

Machine Learning Fundamentals



Edmond de Belamy: $432,500

$$\min_{\mathcal{G}} \max_{\mathcal{D}} E_x \left[\log(\mathcal{D}(x))\right] + E_z \left[\log(1 - \mathcal{D}(\mathcal{G}(z)))\right]$$

Instituto Balseiro

La cátedra

Karina Laneri
FiEstIn

Martín Onetto
Física Forense

Laila D. Kazimierski
DFM

Luis G. Moyano
FiEstIn
luis.moyano@ib.edu.ar

What's ML?

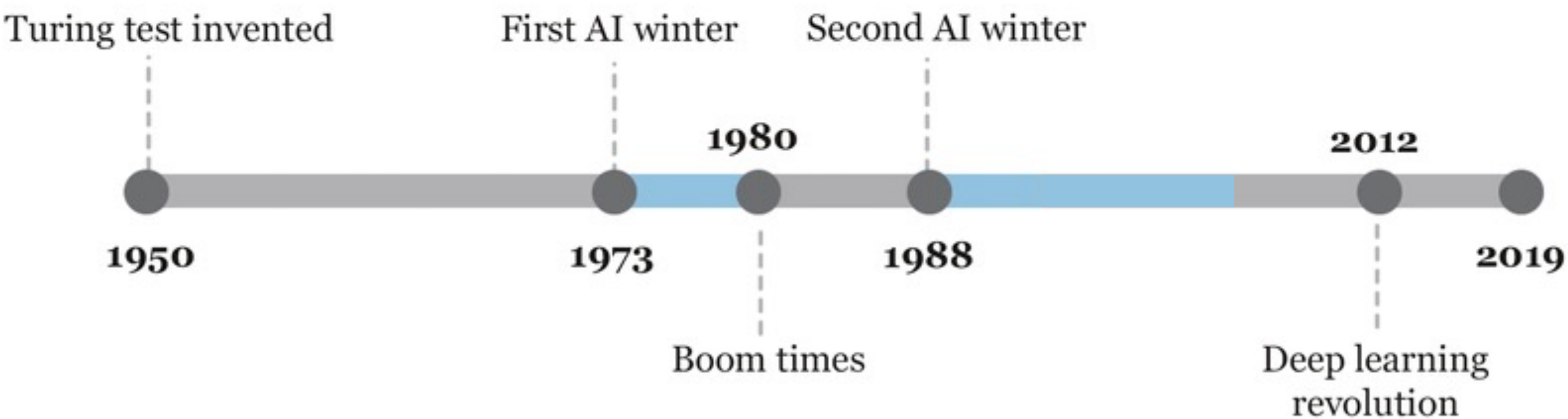
- **Machine learning (aprendizaje automático),** is a field of Computer Science that studies how to extract knowledge from data.

- ML lives at the intersection of statistics, artificial intelligence, and computer science.

- AKA, pattern recognition, statistical modeling, data mining, knowledge discovery, predictive analytics, data science, adaptive systems, self-organizing systems, etc.

> [Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.
>
> —Arthur Samuel, *1959*

Winter is coming...




002 5 1

Playing Atari with Deep Reinforcement Learning

Volodymyr Mnih **Koray Kavukcuoglu** **David Silver** **Alex Graves** **Ioannis Antonoglou**
Daan Wierstra **Martin Riedmiller**

DeepMind Technologies

{vlad,koray,david,alex.graves,ioannis,daan,martin.riedmiller} @ deepmind.com

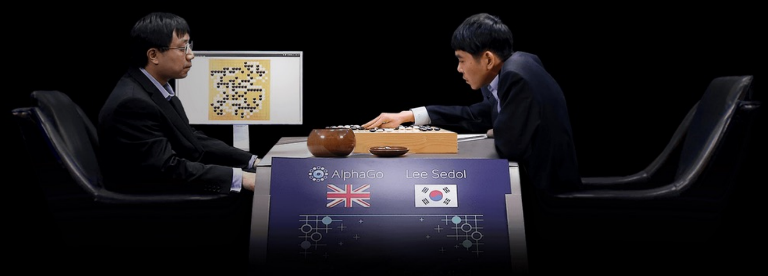
Abstract

We present the first deep learning model to successfully learn control policies directly from high-dimensional sensory input using reinforcement learning. The model is a convolutional neural network, trained with a variant of Q-learning, whose input is raw pixels and whose output is a value function estimating future rewards. We apply our method to seven Atari 2600 games from the Arcade Learning Environment, with no adjustment of the architecture or learning algorithm. We find that it outperforms all previous approaches on six of the games and surpasses a human expert on three of them.

ALPHAGO

HOME FILMMAKERS GALLERY PRESS SEE THE FILM CONTACT TRAILER

On March 9, 2016, the worlds of Go and artificial intelligence collided in South Korea.

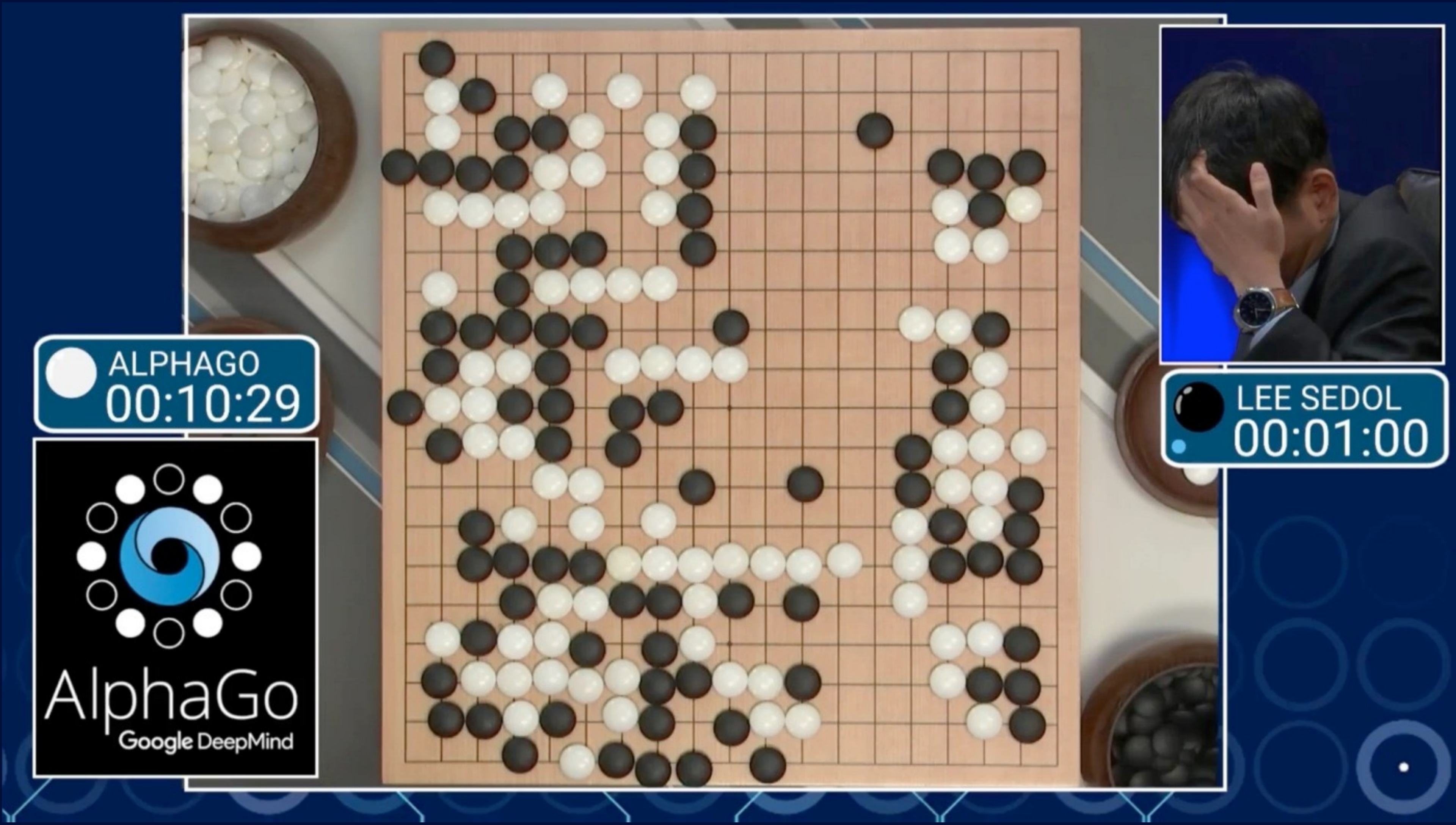


With more board configurations than there are atoms in the universe, the ancient Chinese game of Go has long been considered a grand challenge for artificial intelligence. On March 9, 2016, the worlds of Go and artificial intelligence collided in South Korea for

Directed by Greg Kohs with an original score by Academy Award nominee, Hauschka, *AlphaGo* chronicles a journey from the halls of Oxford, through the backstreets of Bordeaux, past the coding terminals of Google DeepMind in London, and ultimately,

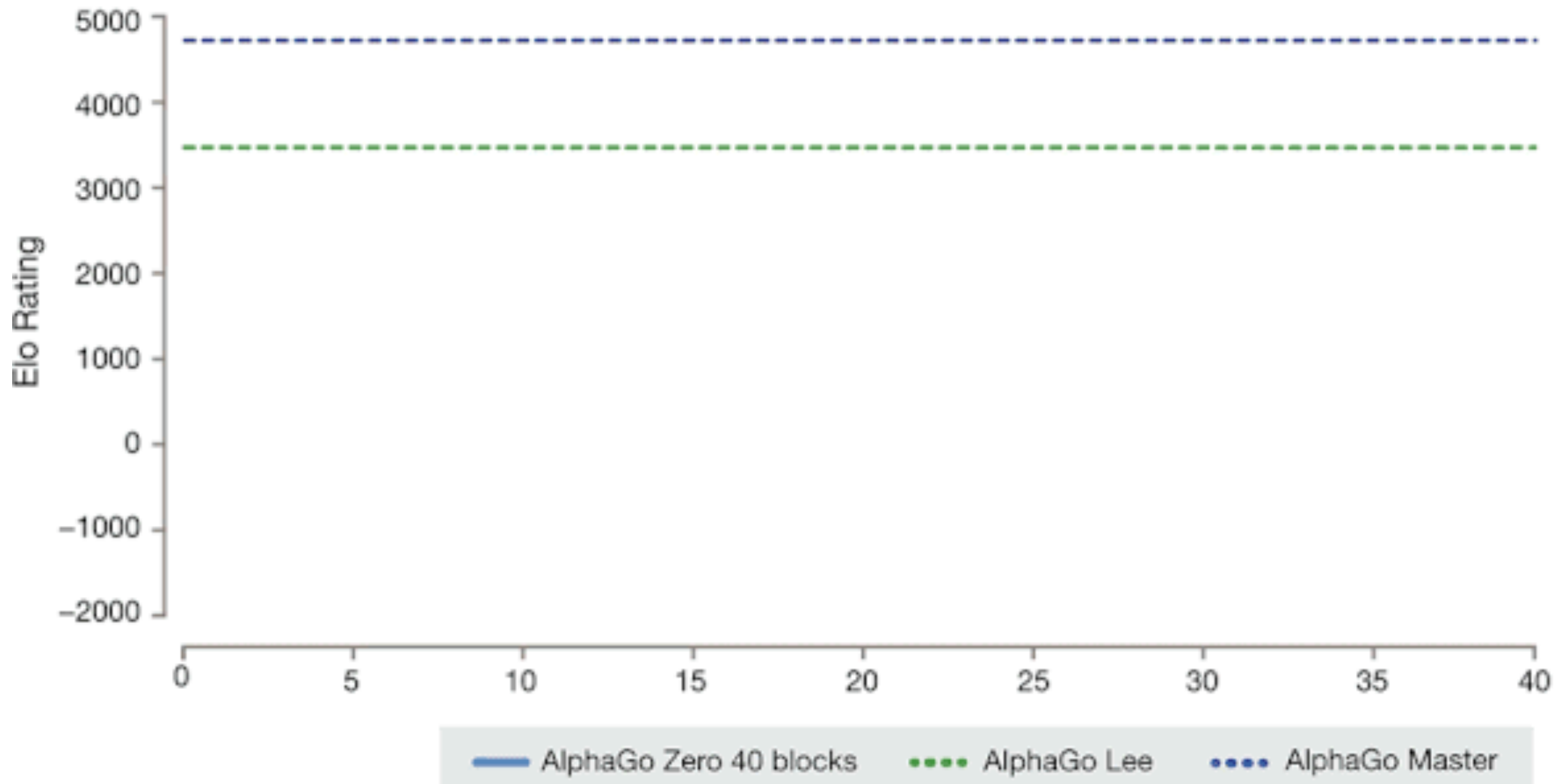
NETFLIX

Computers are useless, they can only give you answers." - Picasso

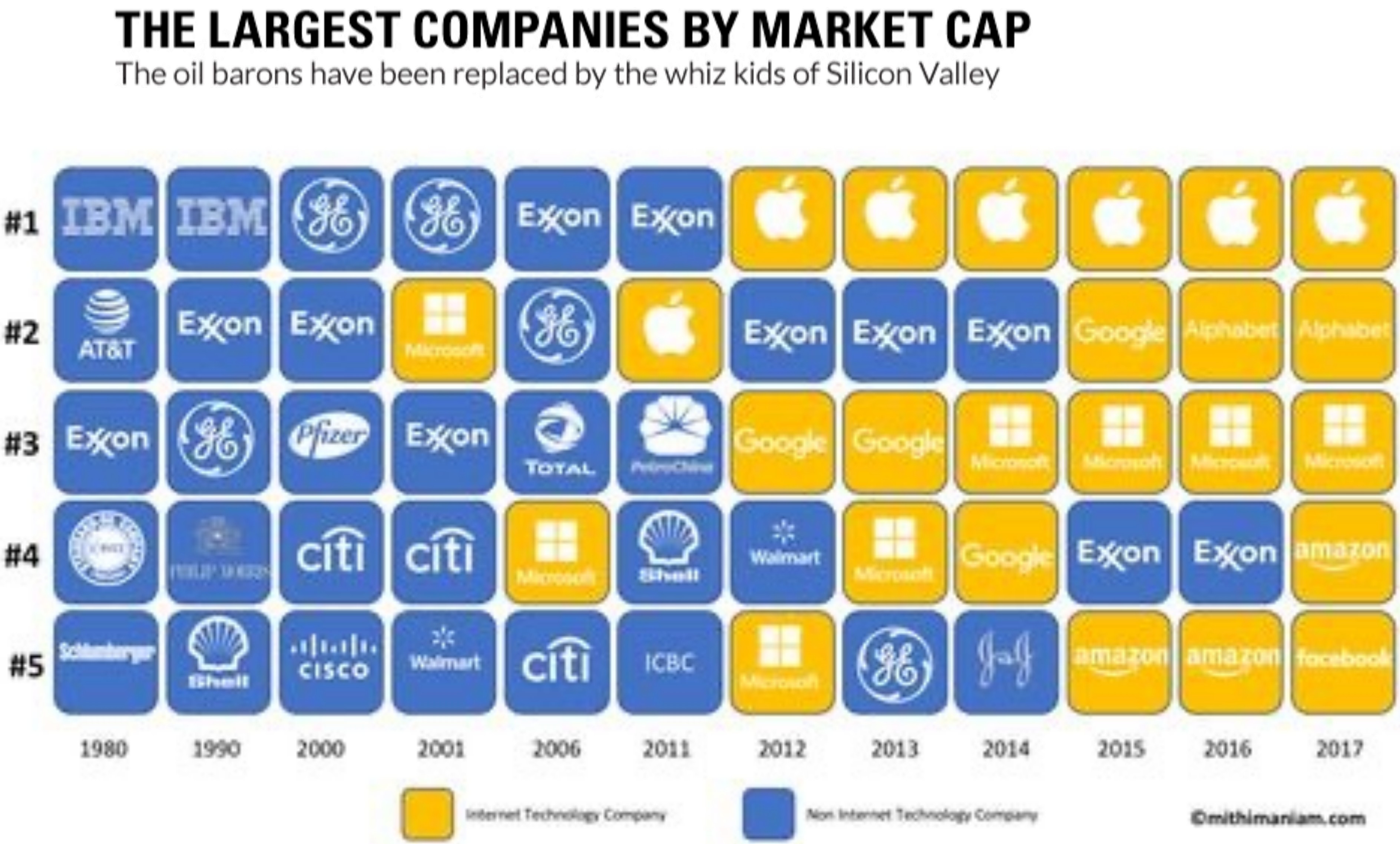


How to become the best
(in less than 6 weeks)

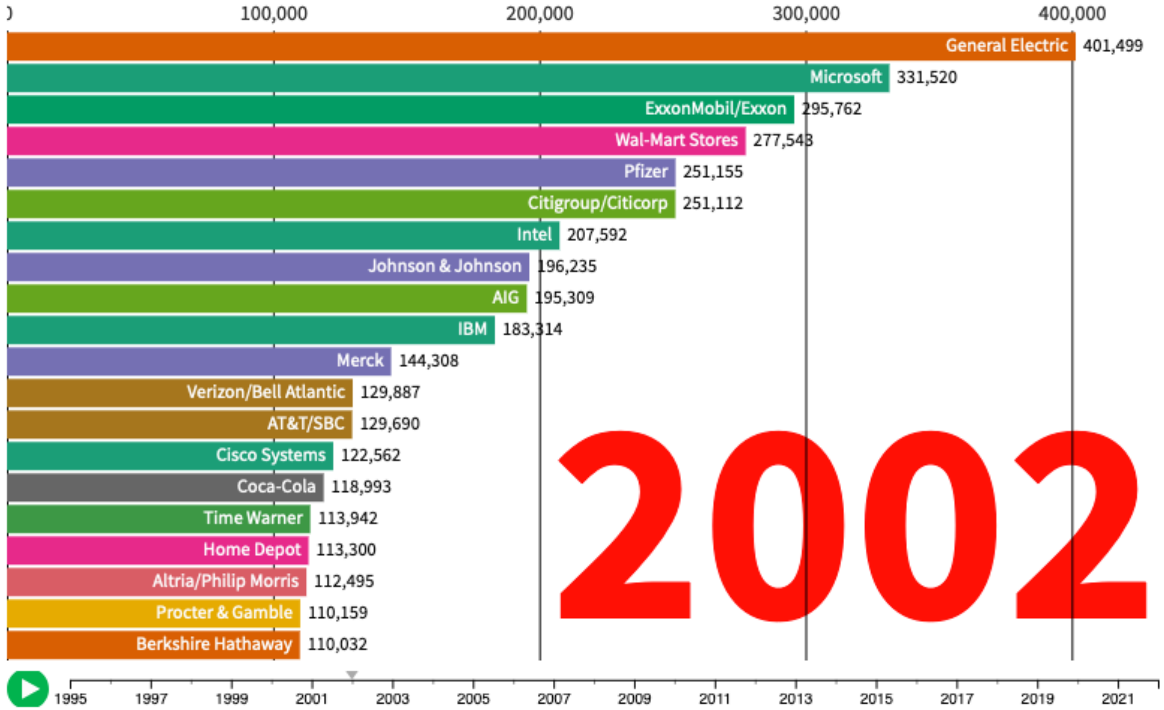
2018

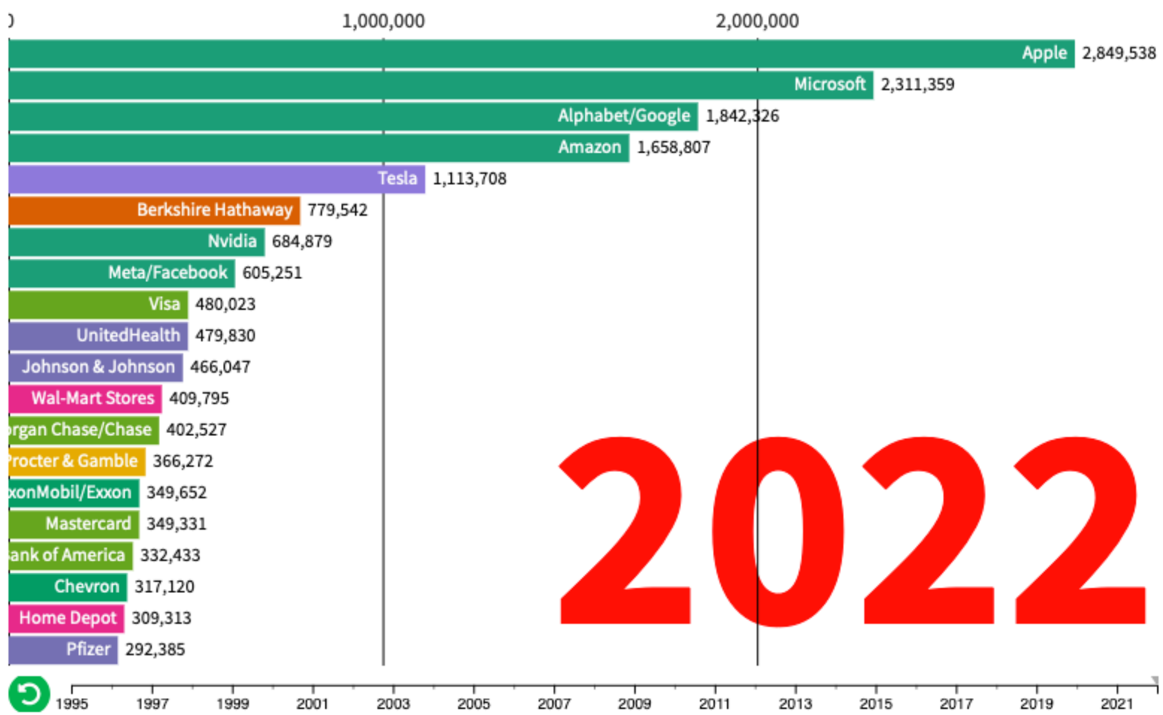


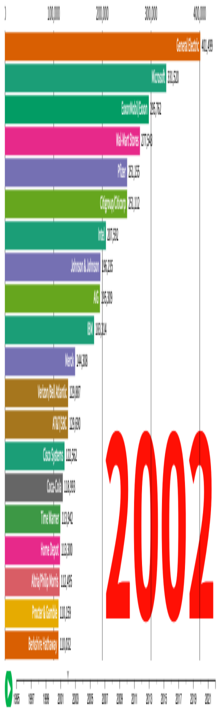
Crédito: https://deepmind.com/blog/alphago-zero-learning-scratch/

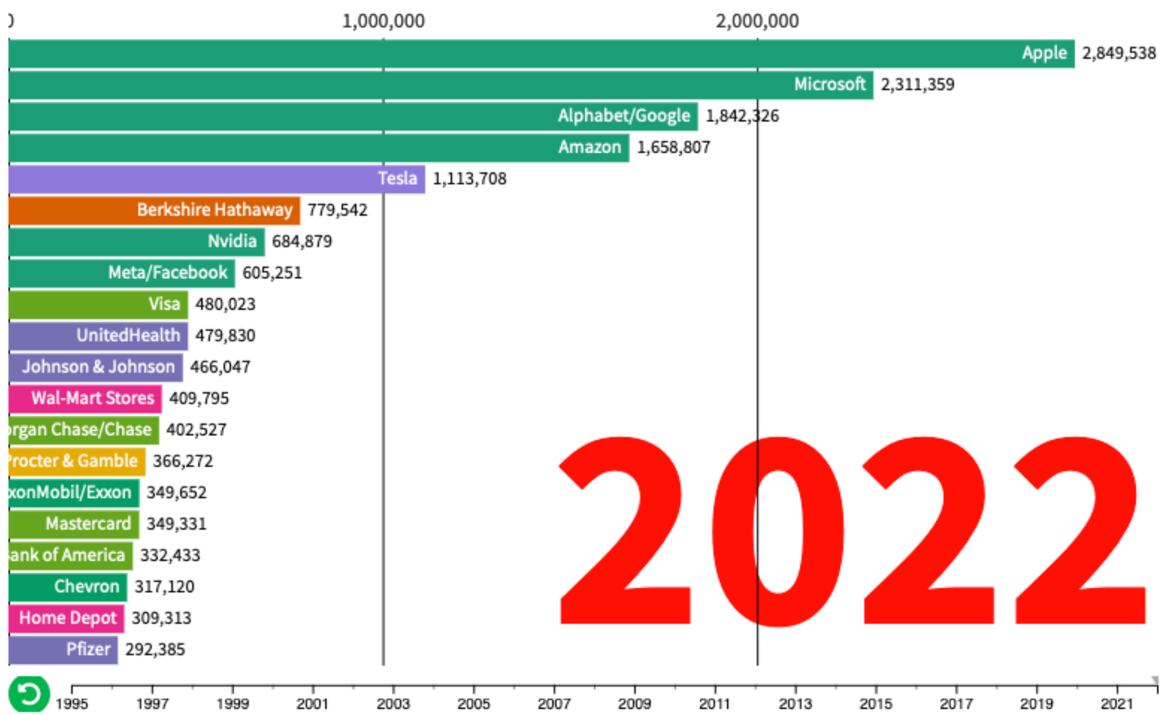


Crédito:http://www.visualcapitalist.com/chart-largest-companies-market-cap-15-years/


100,000
200,000
300,000
400,000
General Electric 401,499
Microsoft 331,520
ExxonMobil/Exxon 295,762
Wal-Mart Stores 277,548
Pfizer 251,155
Citigroup/Citicorp 251,112
Intel 207,592
Johnson & Johnson 196,235
AIG 195,309
IBM 183,314
Merck 144,308
Verizon/Bell Atlantic 129,887
AT&T/SBC 129,690
Cisco Systems 122,562
Coca-Cola 118,993
Time Warner 113,942
Home Depot 113,300
Altria/Philip Morris 112,495
Procter & Gamble 110,159
Berkshire Hathaway 110,032
2002
1995 1997 1999 2001 2003 2005 2007 2009 2011 2013 2015 2017 2019 2021


1,000,000
2,000,000
Apple 2,849,538
Microsoft 2,311,359
Alphabet/Google 1,842,326
Amazon 1,658,807
Tesla 1,113,708
Berkshire Hathaway 779,542
Nvidia 684,879
Meta/Facebook 605,251
Visa 480,023
UnitedHealth 479,830
Johnson & Johnson 466,047
Wal-Mart Stores 409,795
rgan Chase/Chase 402,527
rocter & Gamble 366,272
xonMobil/Exxon 349,652
Mastercard 349,331
ank of America 332,433
Chevron 317,120
Home Depot 309,313
Pfizer 292,385
2022
1995 1997 1999 2001 2003 2005 2007 2009 2011 2013 2015 2017 2019 2021


2002


1,000,000
2,000,000
Apple 2,849,538
Microsoft 2,311,359
Alphabet/Google 1,842,326
Amazon 1,658,807
Tesla 1,113,708
Berkshire Hathaway 779,542
Nvidia 684,879
Meta/Facebook 605,251
Visa 480,023
UnitedHealth 479,830
Johnson & Johnson 466,047
Wal-Mart Stores 409,795
402,527
366,272
349,652
Mastercard 349,331
332,433
Chevron 317,120
Home Depot 309,313
Pfizer 292,385
2022
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
2021

IEEE Open Journal of
Engineering in Medicine and Biology

Technology

COVID-19 Artificial Intelligence Diagnosis using only Cough Recordings

Jordi Laguarta[1], Ferran Hueto[1,2] and Brian Subirana[1,2,*]

COVID-19 Cough Test

Artificial intelligence model detects asymptomatic Covid-19 infections through cellphone-recorded coughs

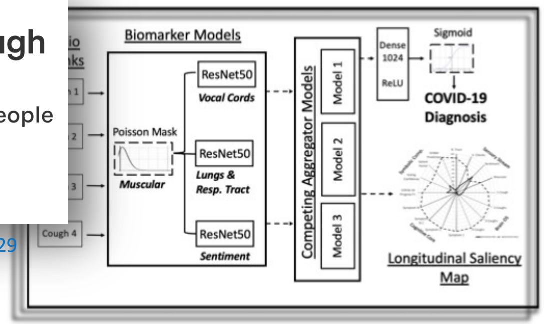
Results might provide a convenient screening tool for people who may not suspect they are infected.

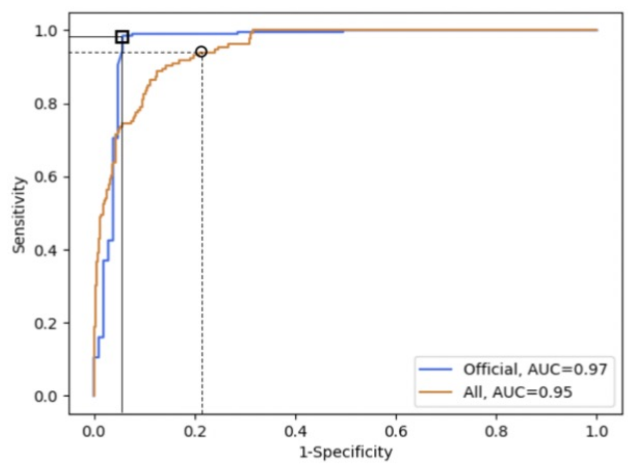
Jennifer Chu | MIT News Office
October 29, 2020

https://news.mit.edu/2020/covid-19-cough-cellphone-detection-1029

- Non-invasive
- Essentially free
- Unlimited throughput
- Real-time results
- Longitudinally monitor

AI Discrimination Model





Performance

- 98.5% sensitivity - 94.2% specificity on PCR/serology confirmed subjects
- 100% Asymptomatic detection rate

Use-Cases

- Daily Country-Wide Screening
- Outbreak Monitoring
- Test Pooling Candidate Selection

REVIEWS OF MODERN PHYSICS, VOLUME 91, OCTOBER–DECEMBER 2019

Machine learning and the physical sciences*

Giuseppe Carleo†

Center for Computational Quantum Physics, Flatiron Institute,
...ew York 10010, USA

Annual Review of Condensed Matter Phys...

Statistical Mechanics of Deep Learning

Yasaman Bahri,[1] Jonathan Kadmon,[2] Jeffrey Pennington,[1] Sam Schoenholz,[1] Jascha Sohl-Dickstein,[1] and Surya Ganguli[1,2]

[1]Google Brain, Google Inc., Mountain View, California 94043, USA
[2]Department of Applied Physics, Stanford University, Stanford, California 94035, USA; email: sganguli@stanford.edu

nature COMMUNICATIONS

ARTICLE

Received 19 Feb 2014 | Accepted 4 Jun 2014 | Published 2 Jul 2014 DOI: 10.1038/ncomms5308

Searching for exotic par... physics with deep learn...

P. Baldi[1], P. Sadowski[1] & D. Whiteson[2]

Article | Published: 21 June 2021

Automatic heterogeneous quantization of deep neural networks for low-latency inference on the edge for particle detectors

Claudionor N. Coelho Jr, Aki Kuusela, Shan Li, Hao Zhuang, Jennifer Ngadiuba, Thea Klaeboe Aarrestad, Vladimir Loncar, Maurizio Pierini, Adrian Alan Pol & Sioni Summers

Nature Machine Intelligence **3**, 675–686 (2021) | Cite this article

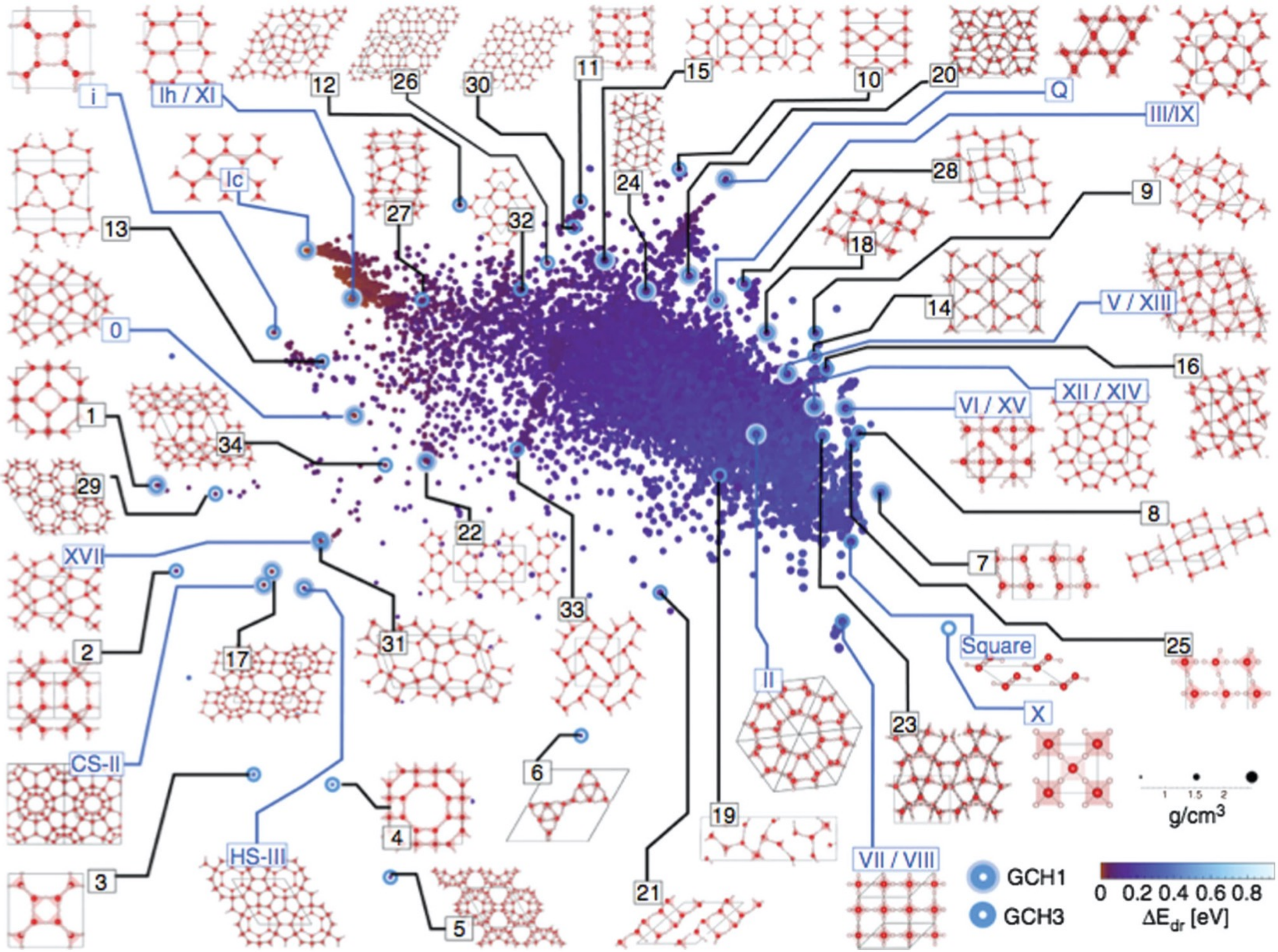
https://www.nature.com/articles/s42256-021-00356-5

Understanding deep learning is also a job for physicists

NATURE PHYSICS | www.nature.com/naturephysics

Automated learning from data by means of deep neural networks is finding use in an ever-increasing number of applications, yet key theoretical questions about how it works remain unanswered. A physics-based approach may help to bridge this gap.

Lenka Zdeborová



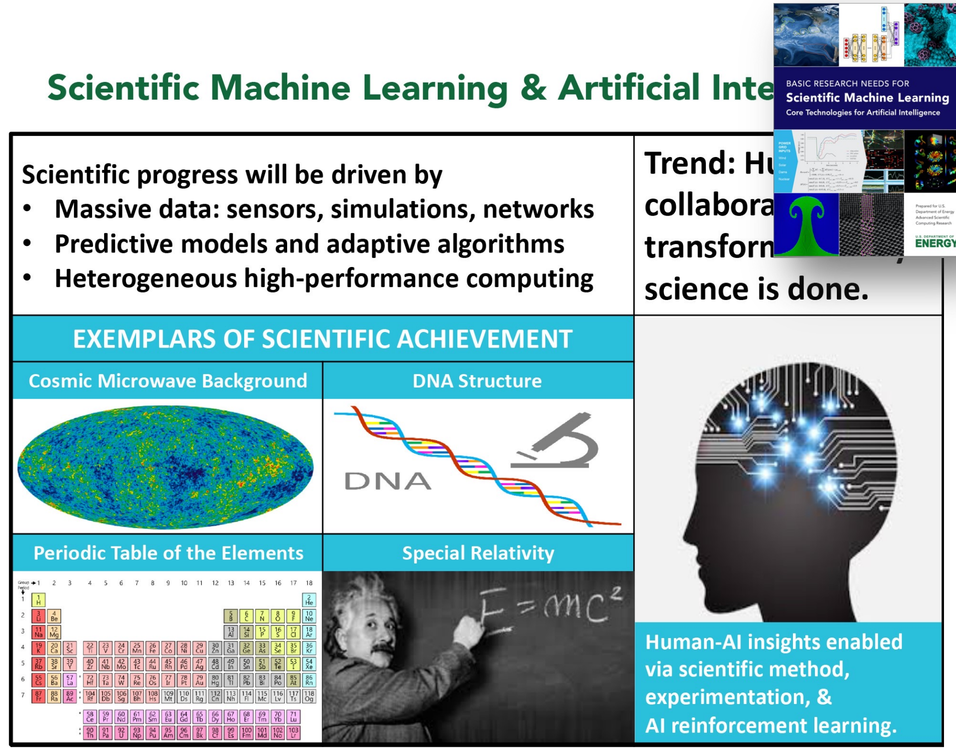
kernel PCA in Engel, E. A., A. Anelli, M. Ceriotti, C. J. Pickard, and R. J. Needs, 2018, Nat. Commun. 9, 2173.

Scientific ML

Scientific machine learning benchmarks

Jeyan Thiyagalingam, Mallikarjun Shankar, Geoffrey Fox & Tony Hey

Nature Reviews Physics 4, 413–420 (2022) | Cite this article



U.S. DEPARTMENT OF ENERGY | Office of Science

DOE Applied Mathematics Research Program
Scientific Machine Learning Workshop (January 2018)

scML lines of research:

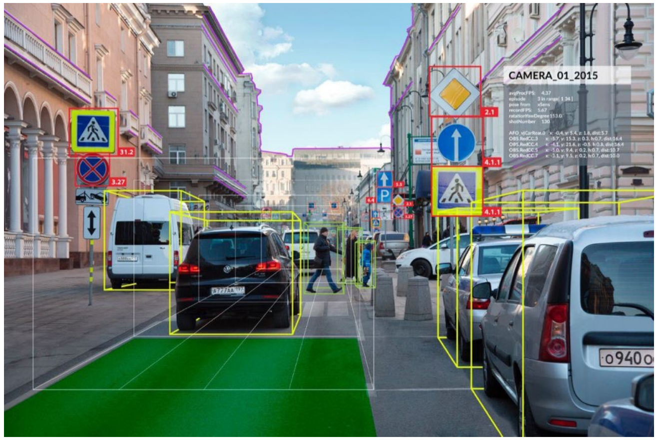
- Limited or low-quality labels or massive unlabeled (often class imbalanced) data sets
- Ground truth is unknown and benchmark data sets are unavailable.
- Scientific data are often high-dimensional, noisy, heterogeneous, low-signal-to-noise, and multiscale.
- Models should respect or incorporate physical laws, constraints, and other scientific domain knowledge.
- Robust methods and an ability to quantify uncertainty are required for scientific rigor.
- Extracting new scientific insights from data requires human-interpretable models or outputs.

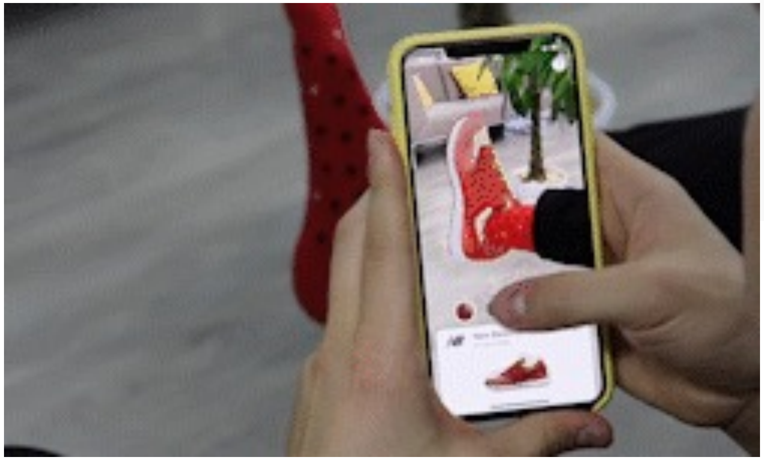
ML applications

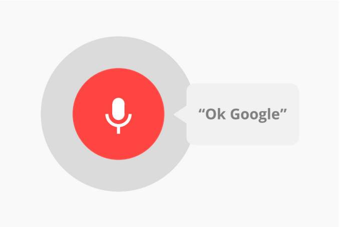
Applications [edit]

There are many applications for machine learning, including:

- Agriculture
- Anatomy
- Adaptive websites
- Affective computing
- Banking
- Bioinformatics
- Brain–machine interfaces
- Cheminformatics
- Citizen science
- Computer networks
- Computer vision
- Credit-card fraud detection
- Data quality
- DNA sequence classification
- Economics
- Financial market analysis [64]
- General game playing
- Handwriting recognition
- Information retrieval
- Insurance
- Internet fraud detection
- Linguistics
- Machine learning control
- Machine perception
- Machine translation
- Marketing
- Medical diagnosis
- Natural language processing
- Natural language understanding
- Online advertising
- Optimization
- Recommender systems
- Robot locomotion
- Search engines
- Sentiment analysis
- Sequence mining
- Software engineering
- Speech recognition
- Structural health monitoring
- Syntactic pattern recognition
- Telecommunication
- Theorem proving
- Time series forecasting
- User behavior analytics







We Asked GPT-3 to Write an Academic Paper about Itself—Then We Tried to Get It Published

An artificially intelligent first author presents many ethical questions—and could upend the publishing process

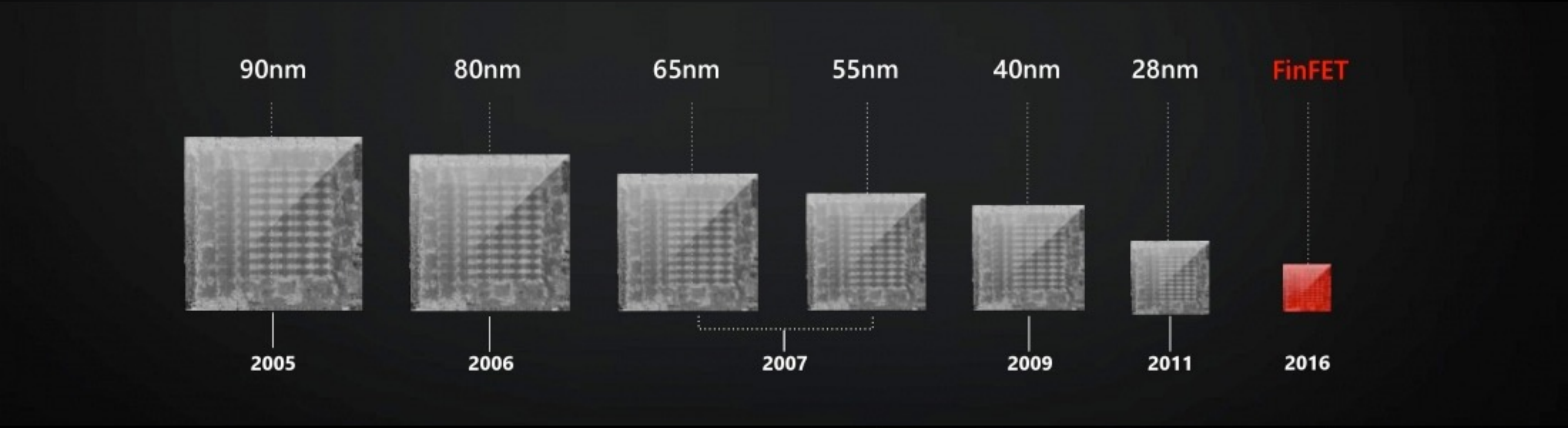
By Almira Osmanovic Thunström

https://www.scientificamerican.com/article/we-asked-gpt-3-to-write-an-academic-paper-about-itself-mdash-then-we-tried-to-get-it-published/

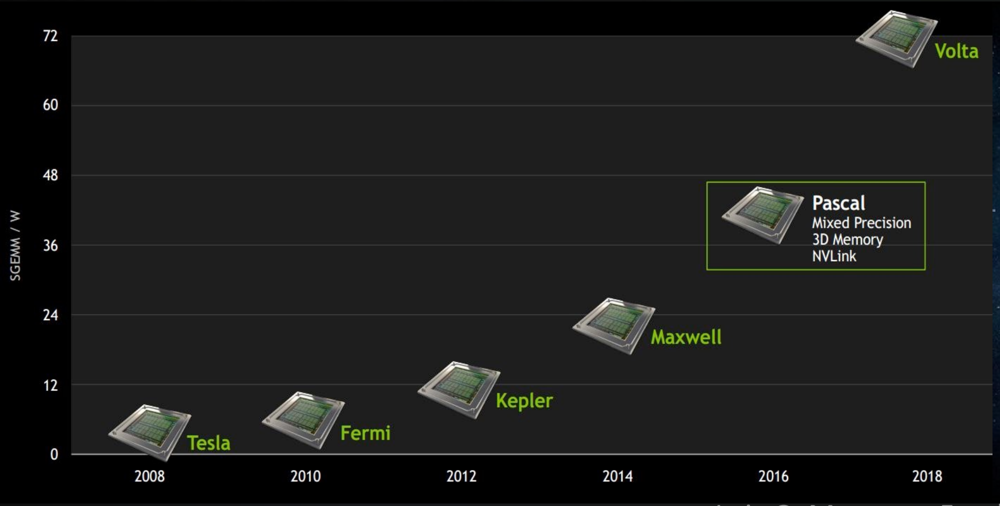
OpenAI

Once we designed this proof-of-principle test, the fun really began. In response to my prompts, GPT-3 produced a paper in just two hours. "Overall, we believe that the benefits of letting GPT-3 write about itself outweigh the risks," GPT-3 wrote in conclusion. "However, we recommend that any such writing be closely monitored by researchers in order to mitigate any potential negative consequences."

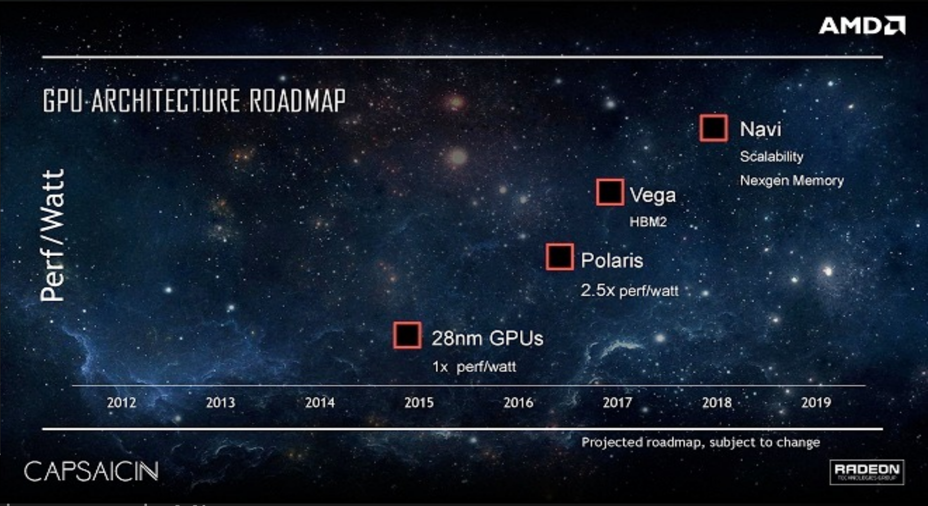
Hardware



90nm
80nm
65nm
55nm
40nm
28nm
FinFET
2005
2006
2007
2009
2011
2016



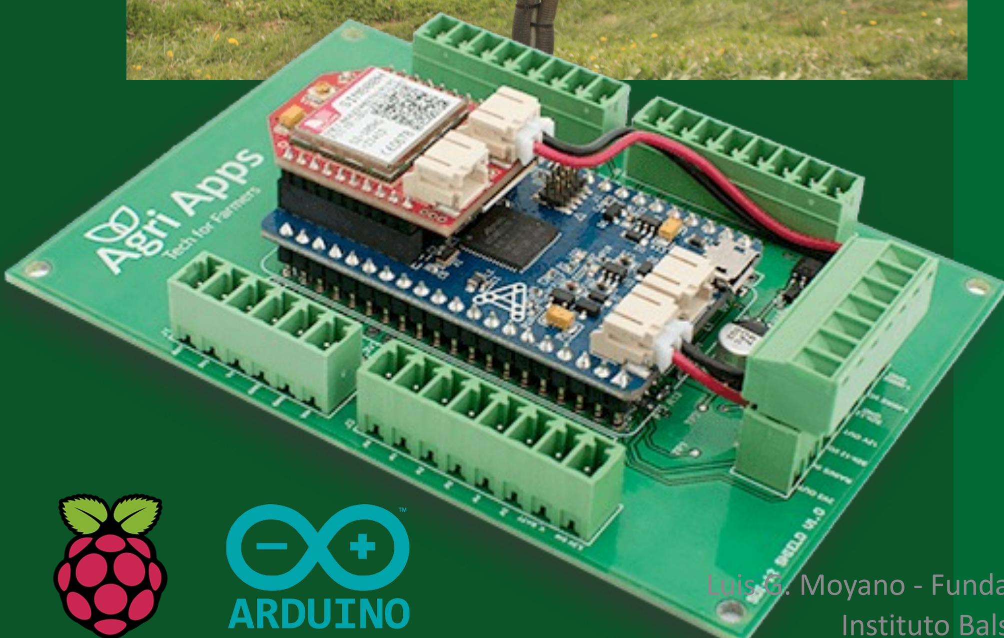
SGEMM / W
72
60
48
36
24
12
0
2008
2010
2012
2014
2016
2018
Tesla
Fermi
Kepler
Maxwell
Pascal
Mixed Precision
3D Memory
NVLink
Volta

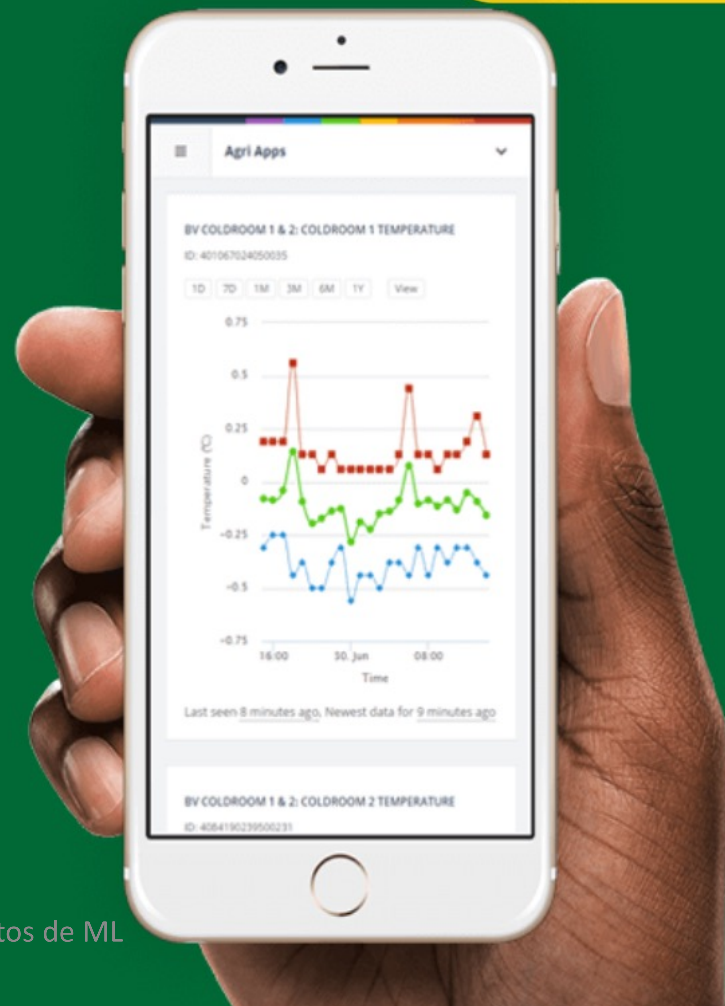


AMD
GPU ARCHITECTURE ROADMAP
Perf / Watt
Navi
Scalability
Nexgen Memory
Vega
HBM2
Polaris
2.5x perf/watt
28nm GPUs
1x perf/watt
2012
2013
2014
2015
2016
2017
2018
2019
Projected roadmap, subject to change
CAPSAICIN
RADEON

Crédito: https://www.agriapps.com
https://elpais.com/economia/2017/03/23/actualidad/1490286632_450842.html

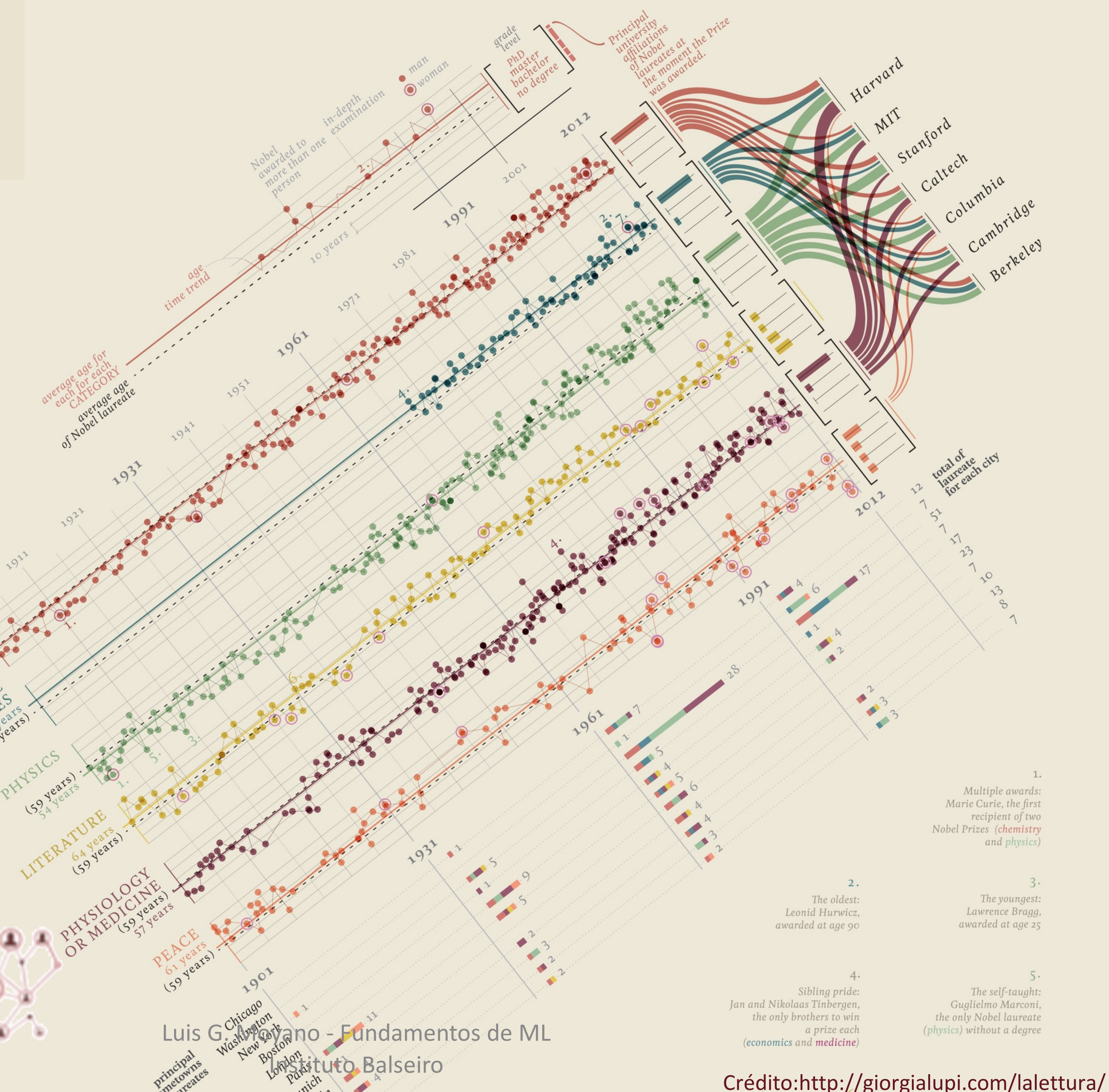
Hardware





ARDUINO

Data



IMAGENET

WIKIPEDIA
The Free Encyclopedia

Crédito:http://giorgialupi.com/lalettura/

Algos & SW

1952 Stochastic Gradient Descent

1958 Perceptron
- Learnable Weights

⋮

1986 Backpropagation
- Multi-Layer Perceptron

1995 Deep Convolutional NN
- Digit Recognition

⋮

2006 Deep belief nets

A fast learning algorithm for deep belief nets *

Geoffrey E. Hinton and **Simon Osindero**
Department of Computer Science University of Toronto
10 Kings College Road
Toronto, Canada M5S 3G4
{hinton, osindero}@cs.toronto.edu

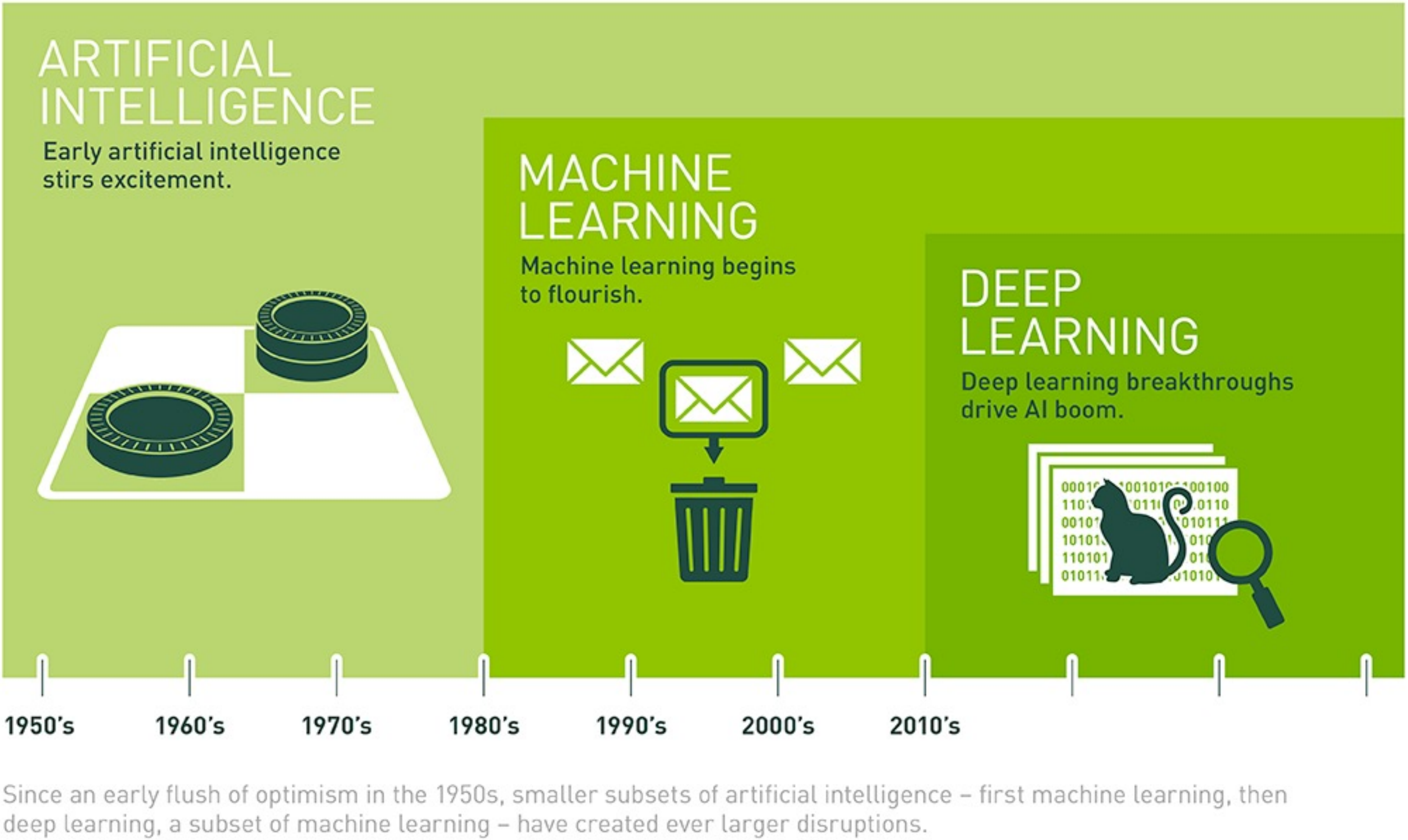
Yee-Whye Teh
Department of Computer Science
National University of Singapore
3 Science Drive 3, Singapore, 117543
tehyw@comp.nus.edu.sg

Abstract

We show how to use "complementary priors" to eliminate the explaining away effects that make inference difficult in densely-connected belief nets that have many hidden layers. Using complementary priors, we derive a fast, greedy algorithm that can learn deep, directed belief networks one layer at a time, provided the top two layers form an undirected associative memory. The fast, greedy algorithm is used to initialize a slower learning procedure that fine-tunes the weights using a contrastive version of the wake-sleep algorithm. After fine-tuning, a network with three hidden layers forms a very good generative model of the joint distribution of handwritten digit images and their labels. This generative model gives better digit classification than the best discriminative learning algorithms. The low-dimensional manifolds on which the digits lie are modelled by

remaining hidden layers form a directed acyclic graph that converts the representations in the associative memory into observable variables such as the pixels of an image. This hybrid model has some attractive features:

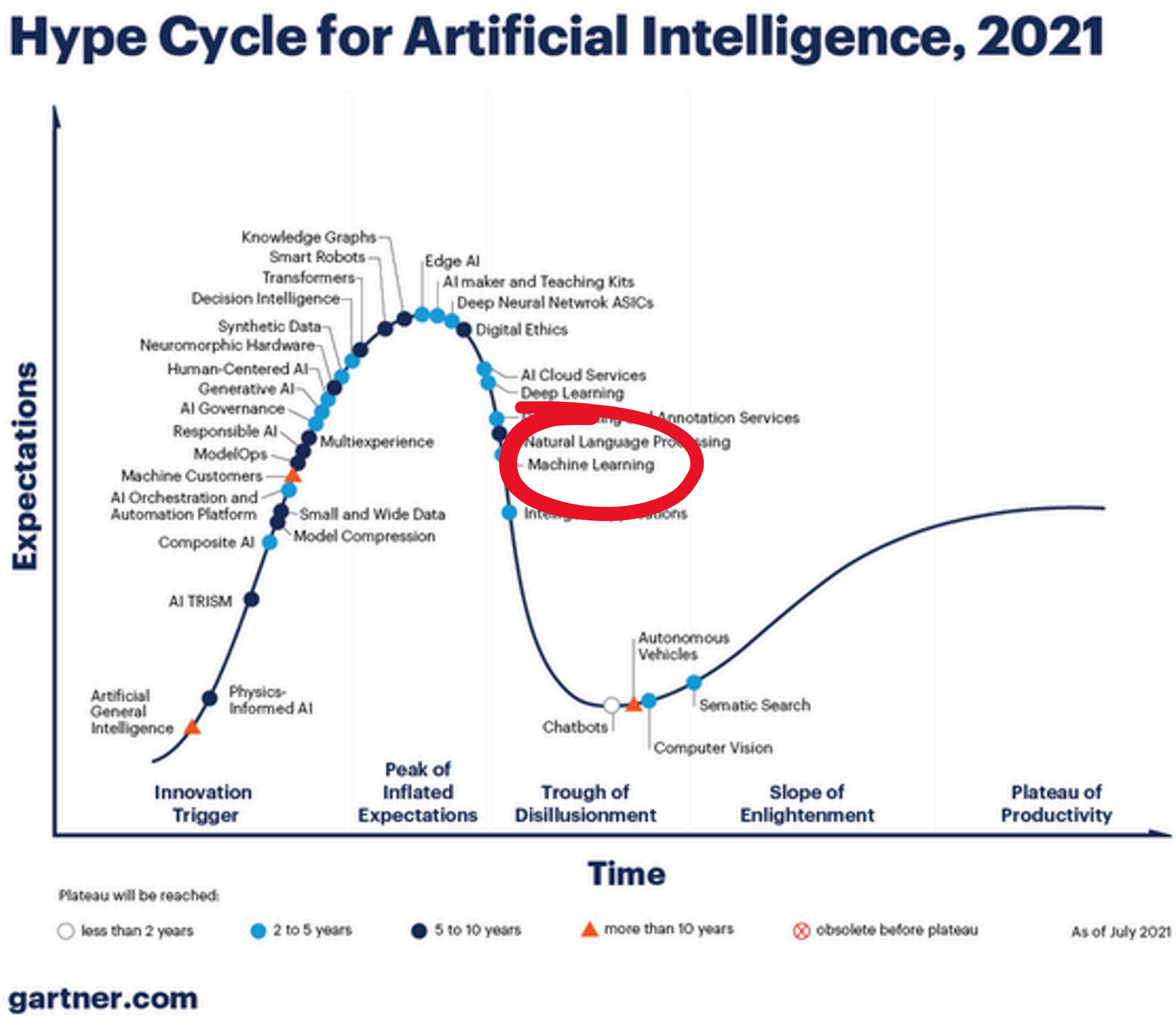
1. There is a fast, greedy learning algorithm that can find a fairly good set of parameters quickly, even in deep networks with millions of parameters and many hidden layers.
2. The learning algorithm is unsupervised but can be applied to labeled data by learning a model that generates both the label and the data.
3. There is a fine-tuning algorithm that learns an excellent generative model which outperforms discriminative methods on the MNIST database of hand-written digits.
4. The generative model makes it easy to interpret the distributed representations in the deep hidden layers.



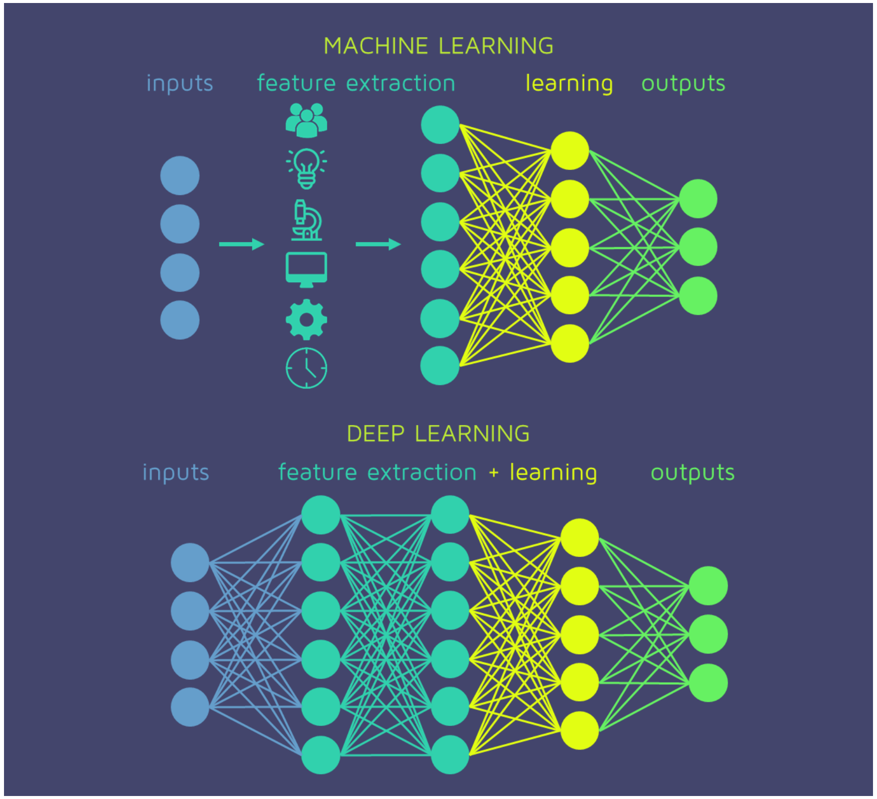
Crédito: https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

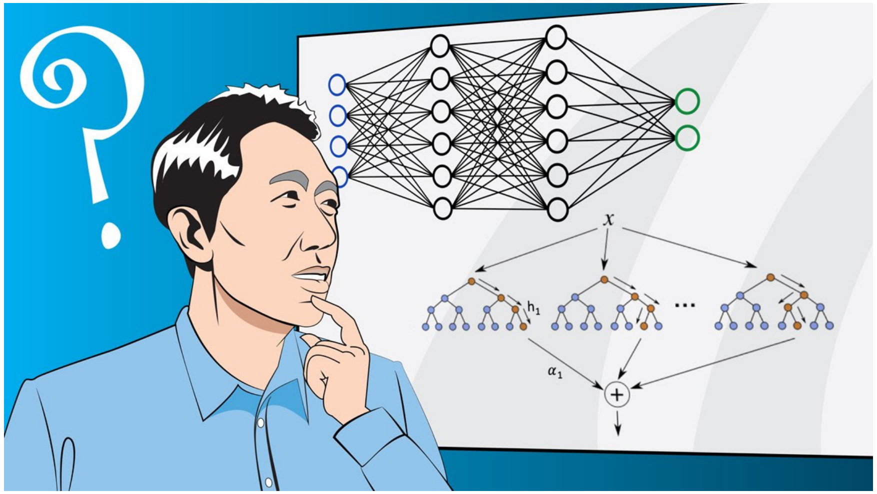
🤑 💰

The Hype


Hype Cycle for Artificial Intelligence, 2021
Expectations
Knowledge Graphs
Smart Robots
Transformers
Decision Intelligence
Synthetic Data
Neuromorphic Hardware
Human-Centered AI
Generative AI
AI Governance
Responsible AI
ModelOps
Machine Customers
AI Orchestration and Automation Platform
Composite AI
AI TRISM
Artificial General Intelligence
Physics-Informed AI
Multiexperience
Small and Wide Data
Model Compression
Edge AI
AI maker and Teaching Kits
Deep Neural Netwrok ASICs
Digital Ethics
AI Cloud Services
Deep Learning
Annotation Services
Natural Language Processing
Machine Learning
Autonomous Vehicles
Chatbots
Sematic Search
Computer Vision
Innovation Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateau of Productivity
Time
Plateau will be reached:
less than 2 years
2 to 5 years
5 to 10 years
more than 10 years
obsolete before plateau
As of July 2021
gartner.com

So, what about Deep Learning?





If I had refreshed my familiarity with boosted trees, I would have made a better decision. – ***Andrew Ng*** *(The Batch, 2022)*

Course roadmap

- ML Foundations

- Classic learning
  - Supervised learning
  - Unsupervised learning

- Deep learning
  - Supervised learning
  - Unsupervised learning

- Advanced topics (fly-by)

(+ 4 special sessions)

Objetivo de ML Fundamentals

- Introducción al aprendizaje automático
  - *Peppered* crash course
- Target: estudiantes IB y oyentes*
- Objetivo práctico: aplicar correctamente herramientas existentes y poder interpretar material de terceros
- Foco en **conceptos**, pero con inevitable dosis de programación

Qué es MLFundamentals, y qué no

- Intro a ML en general
  - No es (solo) Deep Learning
- Sin correlativas
  - Pero un poco de prob & stat habrá
- Se dará via Python
  - pero será **agnóstica** en el lenguaje
  - y no será una materia de programación
- Se asume **proacción**

Fechas y horarios

- Martes y viernes de 9 a 13
  - Vamos a necesitar un par de jueves, a definir
- Del 16/08 al 07/10
- Sin clases la última semana de septiembre (*semanita/RAFA107*)
- La última semana va a ser para evaluación

Evaluación

- Será mediante dos informes:
  - uno a media materia: *informe escrito*
    - tarea y datos dados por la cátedra
  - otro al final: *póster con presentación*
    - tema y datos a definir
    - pdf del póster
    - presentación de 5'-10'
    - código
  - énfasis en manejo de conceptos, no tanto en el código

Repo de GitHub

git@github.com:ML-Fundamentals/MLFundamentals.git

- Slides de las clases
  - con audio, pero sin garantías
- Discusión
- Demos
- Prácticas
- Bibliografía:

- 2006 - Bishop - Pattern Recognition And Machine Learning
- 2012 - Murphy - Machine Learning: A Probabilistic Perspective
- 2012 - Domingos - A Few Useful Things to Know about Machine Learning
- 2013 - Hastie - The Elements of Statistical Learning: Data Mining, Inference, and Prediction
- 2013 - James - An Introduction to Statistical Learning with Applications in R
- 2014 - Shalev-Shwartz - Understanding machine learning: From theory to algorithms
- 2016 - Goodfellow - Deep learning
- 2017 - Muller - Introduction to Machine Learning with Python
- 2018 - Chollet - Deep Learning with Python
- 2019 - Géron - Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow_ Concepts, Tools, and Techniques to Build Intelligent Systems
- 2020 - Chacón – Pro Git

Inscripción

- Hasta principios de septiembre
  - No pueden ser evaluados sin estar inscriptos
  - Mándennos un mail confirmando inscripción
- Están incluidos doctorandos IB
- Vocacionales vs. Oyentes
- Externos?

Preguntas?

HELP!

- Python's (R, etc.) own documentation (*RTFM* 👉)
- Books! 📚
- Scikit Learn's (Keras, Github, etc.) website
- Google: blogs, articles, homepages, etc.
- Github course discussions/issues
- Stackoverflow and sister sites (mind your asking)
- Profes (last resource!! ;)

Today

- *Motivation*
  - *What's ML?*
  - *ML as programming paradigm change*
- What is learning
- ML paradigms
- End-to-end ML Worflow

Say again, what's ML?

- **Machine learning** is the study of algorithms that turn data into programs: *computers programming themselves*
- **Machine learning** as the inverse of programming (Domingos, The Master Algorithm 2015)

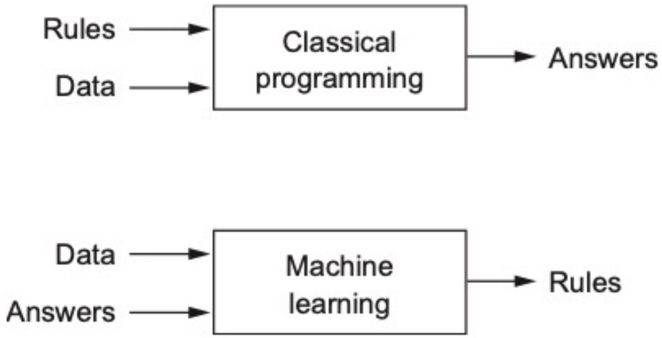
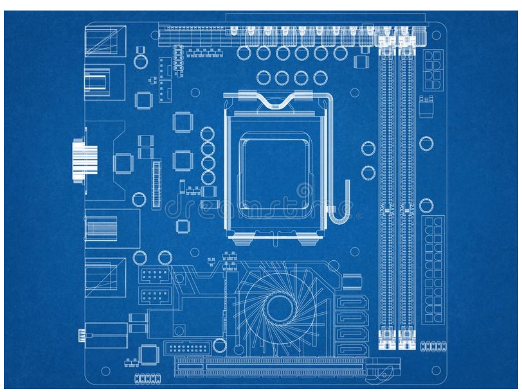
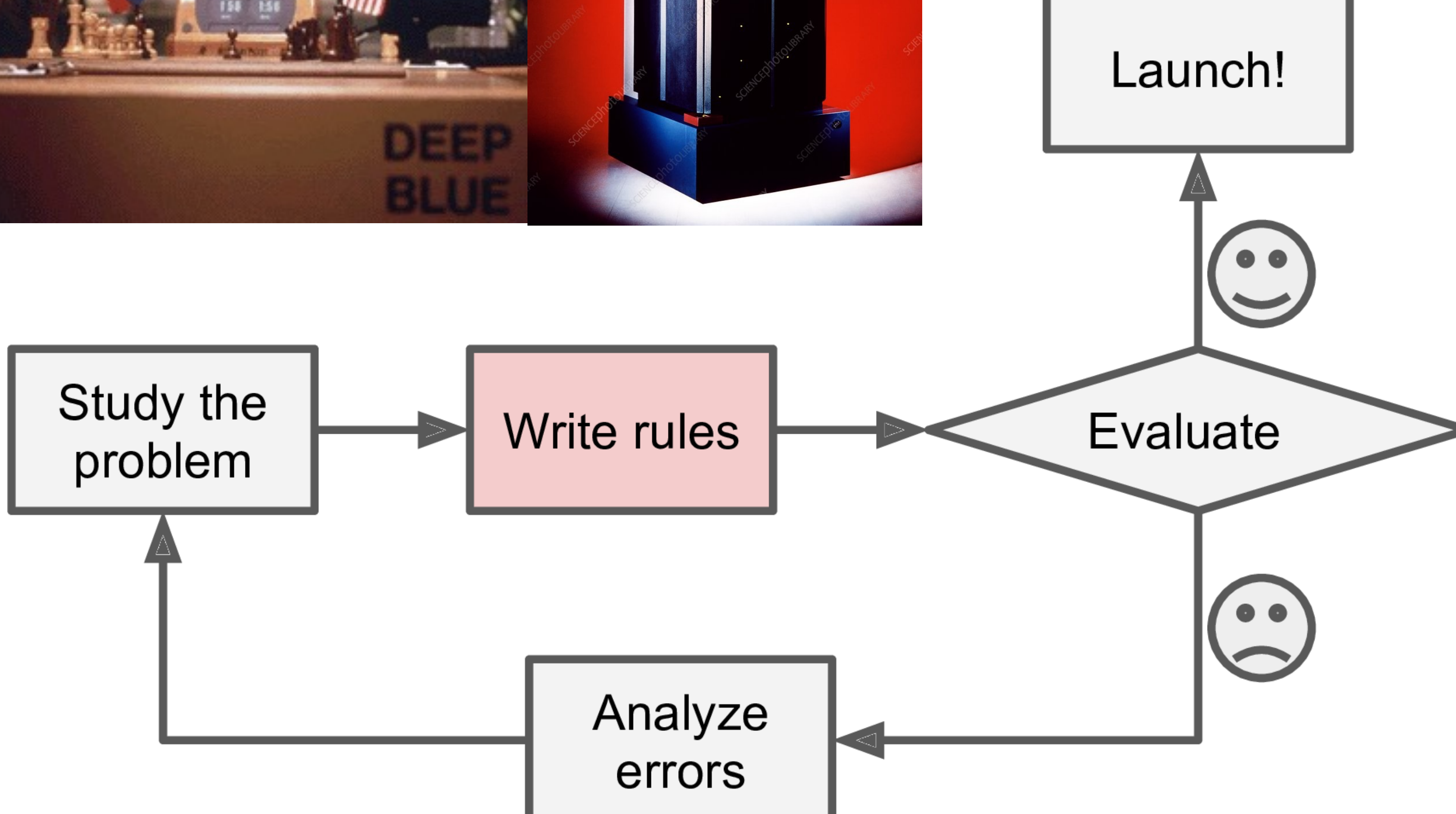


Figure 1.2 Machine learning: a new programming paradigm
Chollet p5



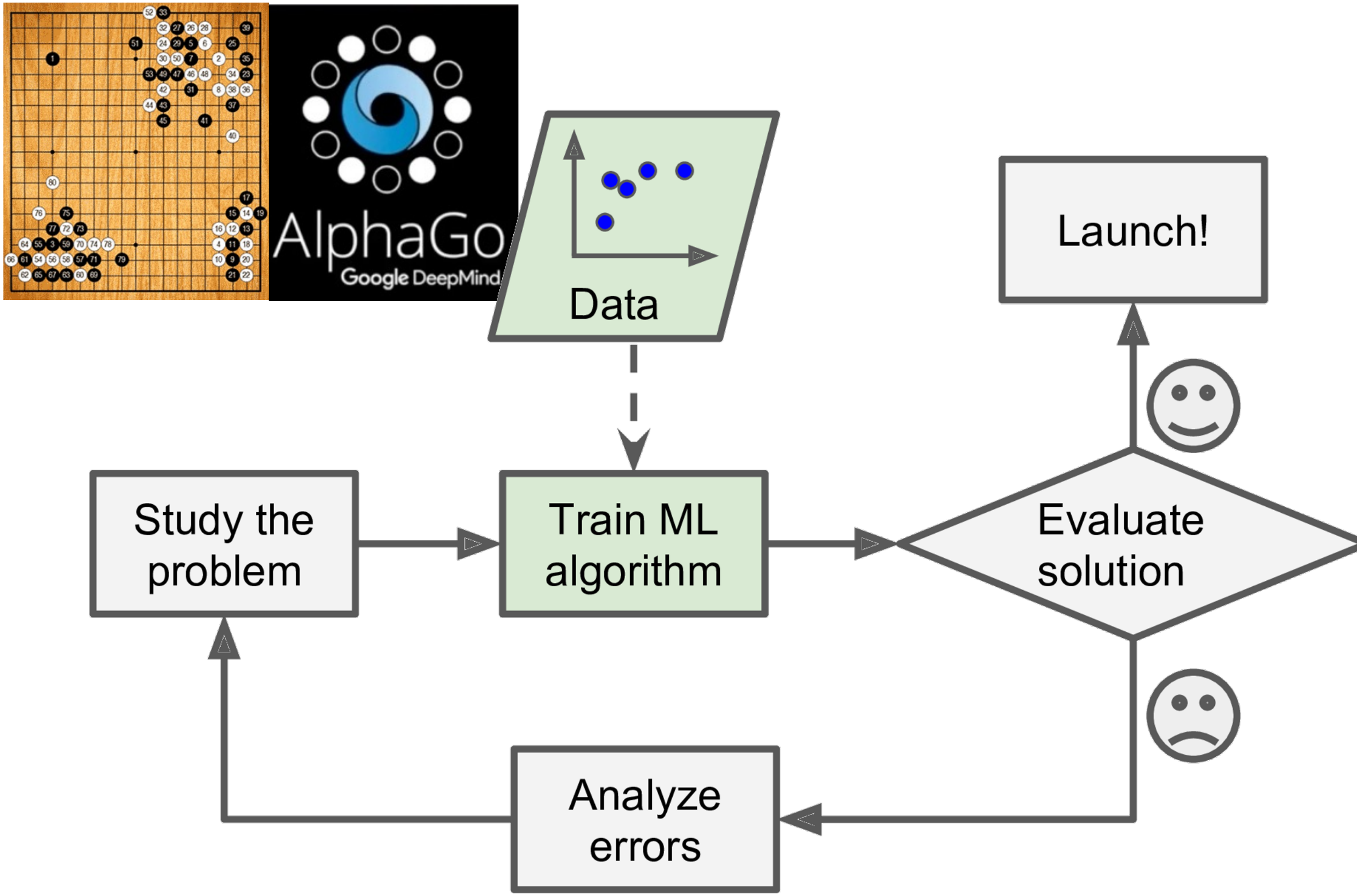
Traditional programming – rule-based

?

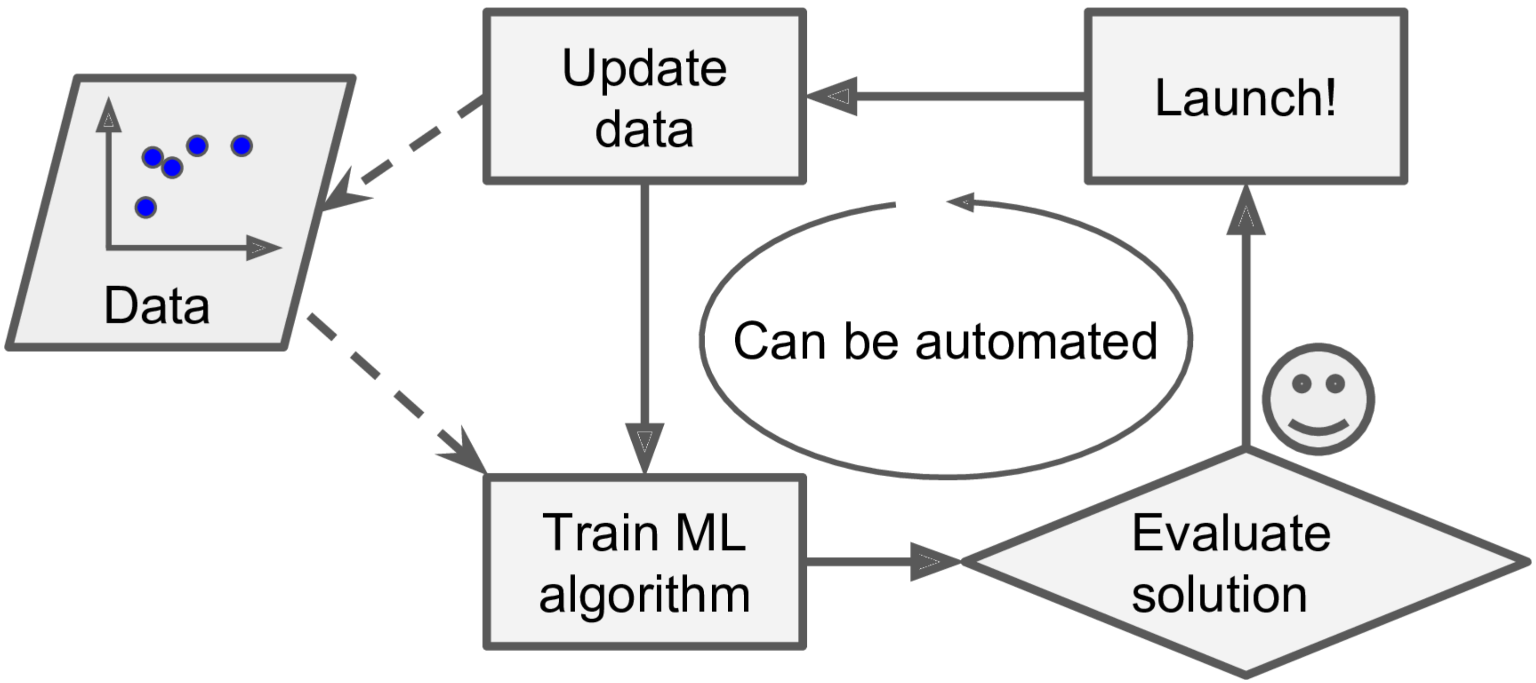


ML programming – data based

?



Workflow ML


Update
data
Launch!
Data
Can be automated
Train ML
algorithm
Evaluate
solution

What is learning

- "Learning is just reducing error"
  2019 - Book - Trask - Grokking Deep Learning-Manning Publications

- "Learning is optimization" (extremization of a cost function)
  Neil D. Lawrence Slide 199

- Machine Learning equates to:
  – Representation
  – Evaluation
  – Optimization

Pedro Domingos, A Few Useful Things to Know about Machine Learning, CACM 55(10), 2012 , in lecture_slides-week6-048_intro_machine_learning_2.pdf

What's ML?

A little deeper...

- ML = ***representation*** + evaluation + optimization

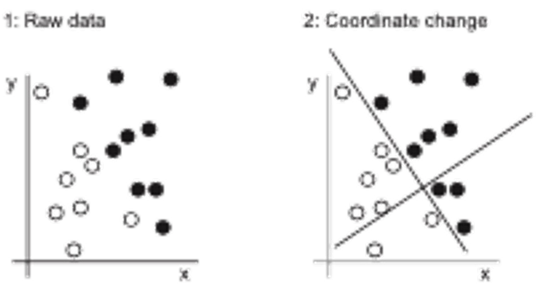
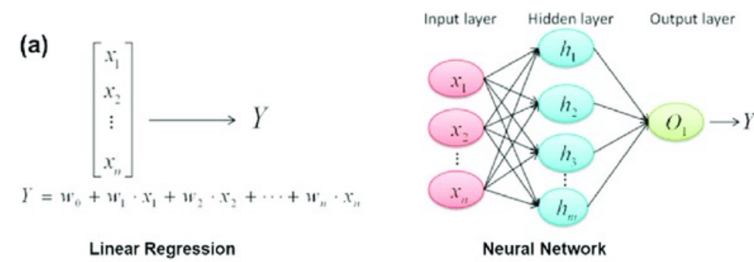


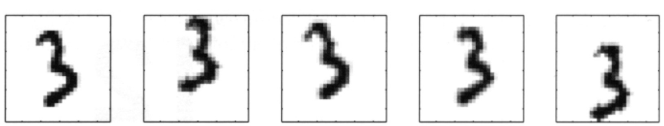
Figure 1.4 Coordinate change

2018 - Chollet - Deep Learning with Python



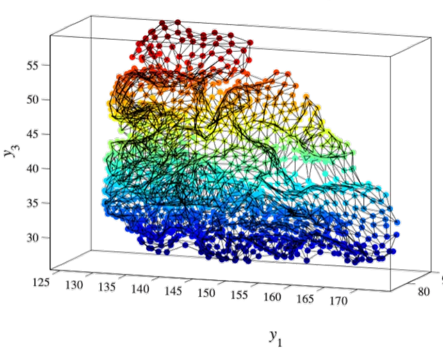
feature representation

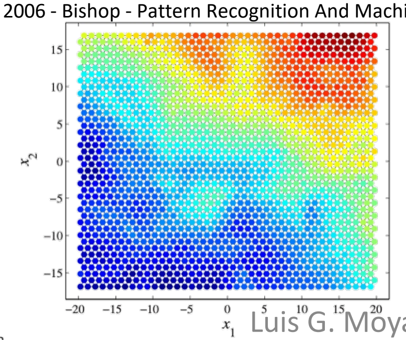
learner space



manifold learning

2006 - Bishop - Pattern Recognition And Machine Learning





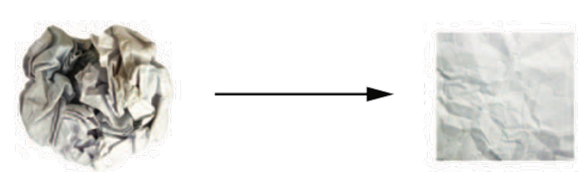


Figure 2.9 Uncrumpling a complicated manifold of data

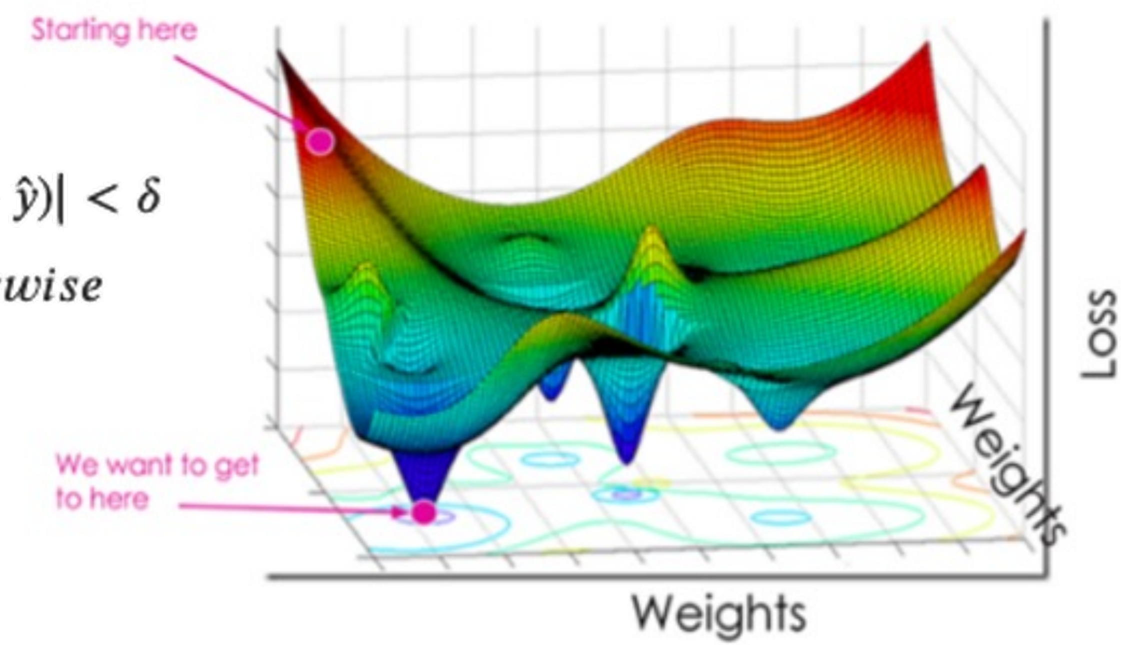
2007 - Lee - Nonlinear Dimensionality Reduction

What's ML?

A little deeper...

- ML = *representation* + ***evaluation*** + optimization

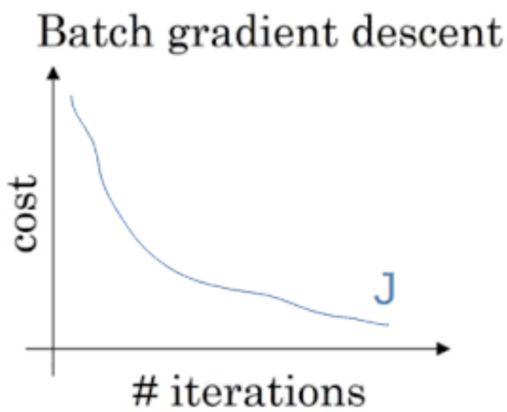
$$L_\delta = \begin{cases} \frac{1}{2}(y-\hat{y})^2 & if\ |(y-\hat{y})| < \delta \\ \delta((y-\hat{y}) - \frac{1}{2}\delta) & otherwise \end{cases}$$

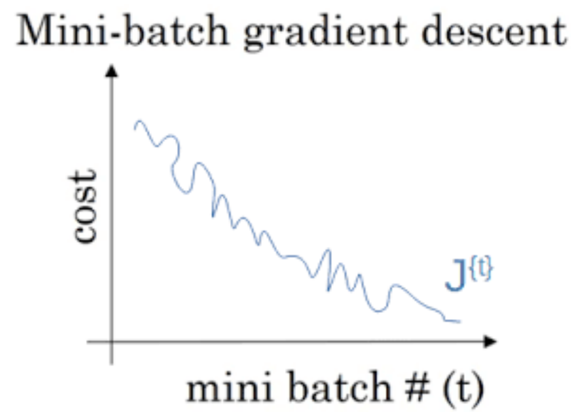


What's ML?

A little deeper...

- ML = *representation* + evaluation + **optimization**





Cost Function

$$J\left(\Theta_0, \Theta_1\right) = \frac{1}{2m} \sum_{i=1}^{m} [h_\Theta(x_i) - y_i]^2$$

True Value

Predicted Value

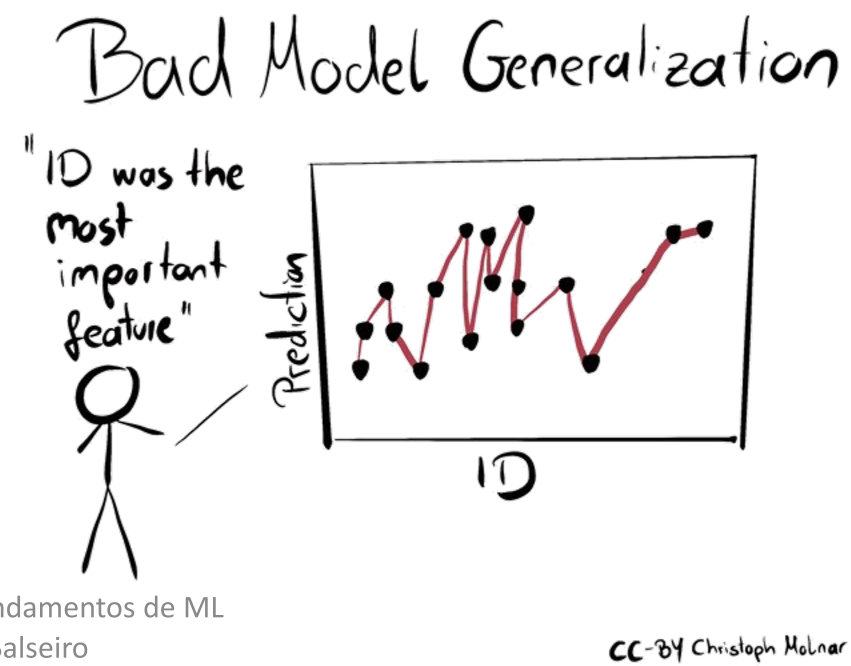
Gradient Descent

$$\boxed{\Theta_j = \Theta_j - \alpha \frac{\partial}{\partial \Theta_j} J\left(\Theta_0, \Theta_1\right)}$$

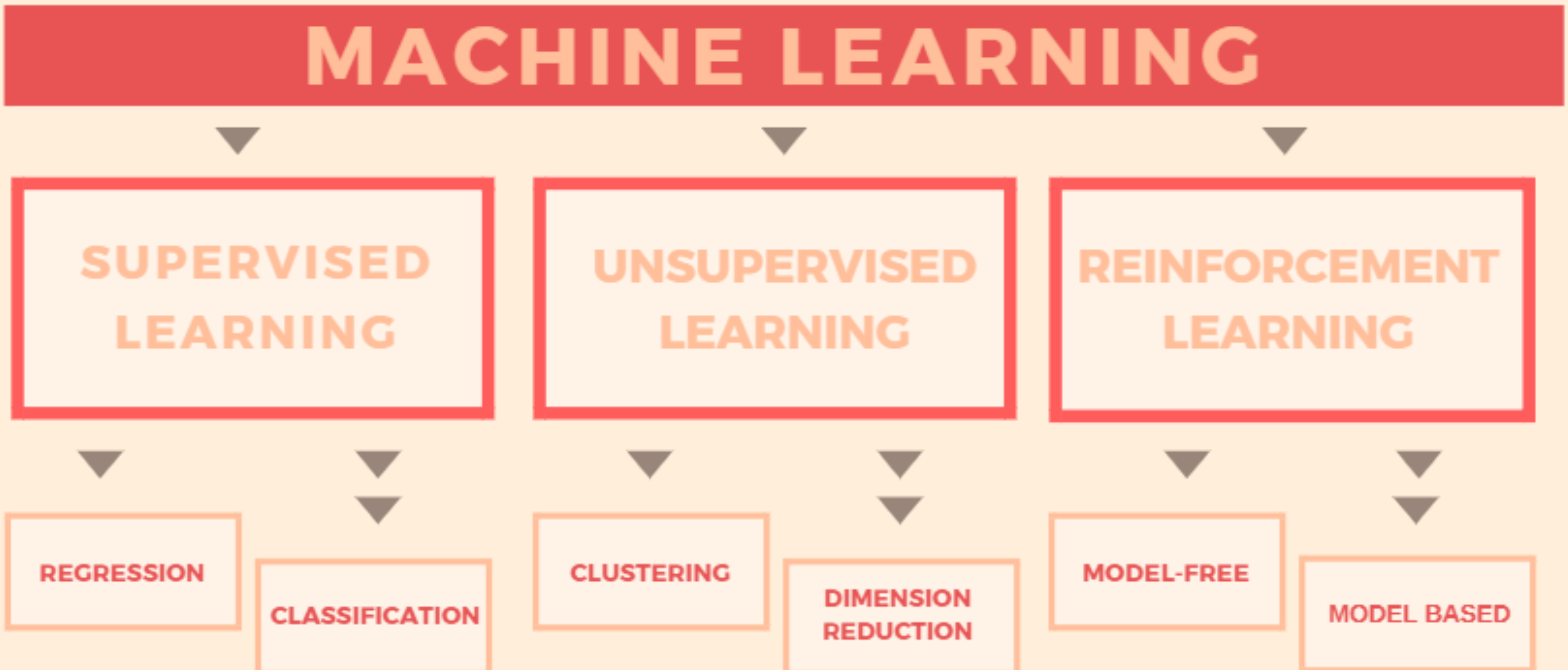
Learning Rate

ML's main objective: to generalize

- In machine learning, the goal is to achieve models that generalize—**that perform well on never-before-seen data**—
- Most of the time, failiure to generalize is due to *overfitting,* i.e. learning traits specific to a given set of data.



Paradigmas de ML y sus modelos



SUPERVISED LEARNING: Developer teaches machine that there are different variables.

UNSUPERVISED LEARNING: Machine randomly clusters and defines variables, sans developer.

REINFORCEMENT LEARNING: Machine is taught which variables are correct through a reward system.

Tareas y algoritmos frecuentes

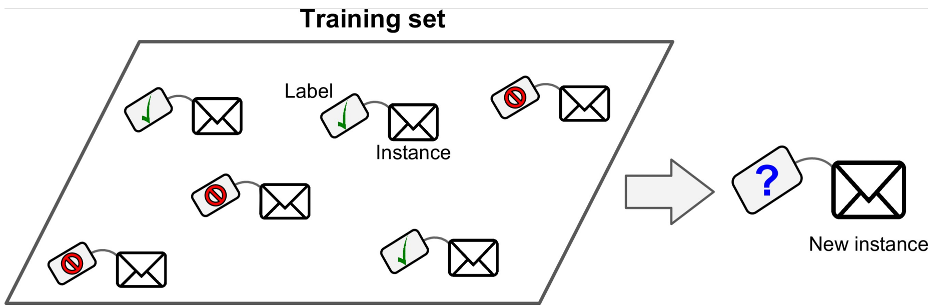
ML Tasks *Broad Categories*	***Supervised***	***Unsupervised***
Discrete	Classification Computer vision | Image Classification Speech, handwriting recognition Drug discovery	Clustering K-means, mean-shift Large-scale clustering problem Hierarchical clustering, GMM
Continuous	Regression Computer vision | Object Detection Linear, logistic regression	Reduction of Dimensionality PCA, LDA (Kernel) Density Estimation

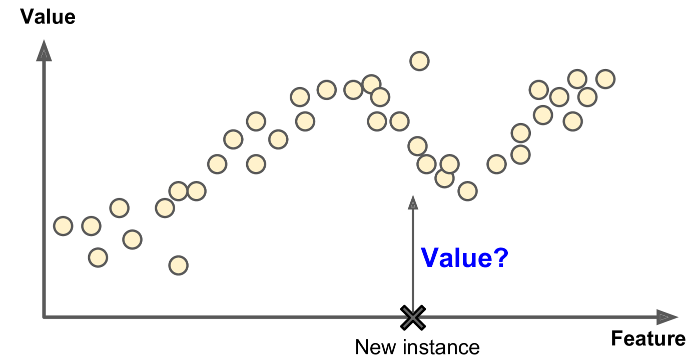
https://pt.slideshare.net/datascienceth/machine-learning-in-image-processing?ref=&smtNoRedir=1
https://eleg5491.github.io/ml_basics/

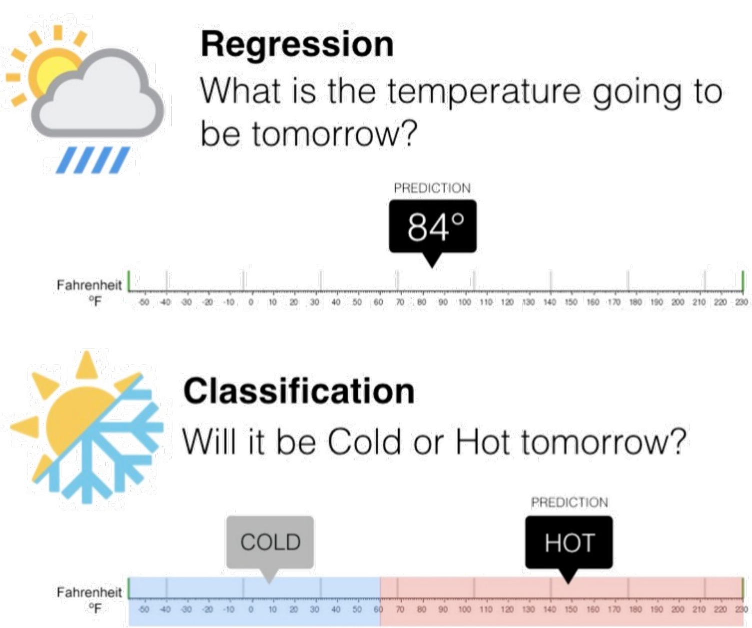
Classification: SPAM or HAM?

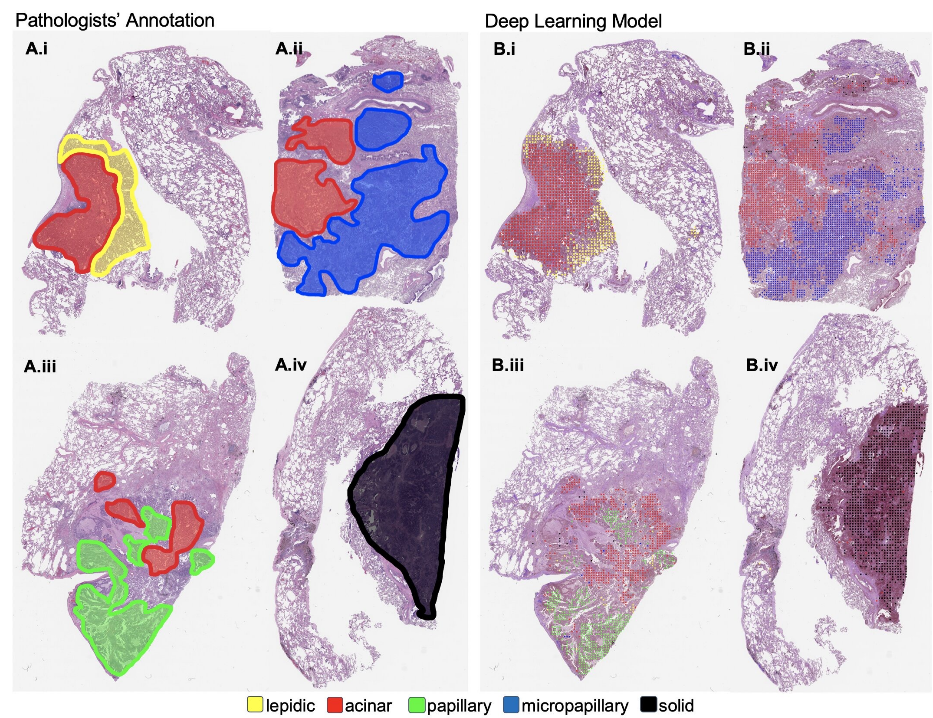
Supervised learning

Generalizing from known examples



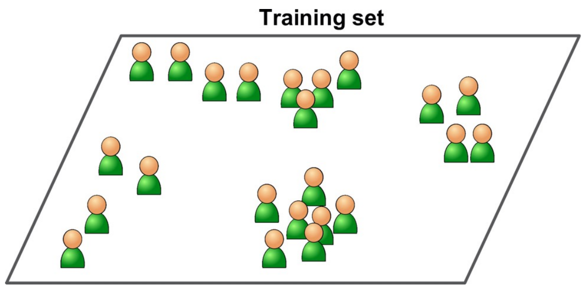


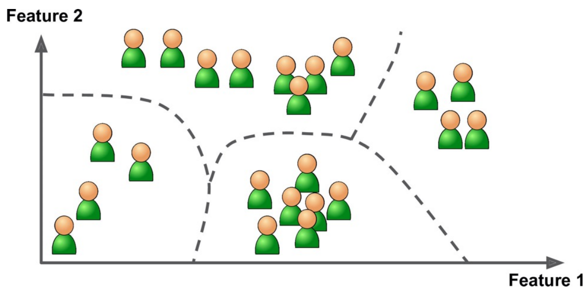


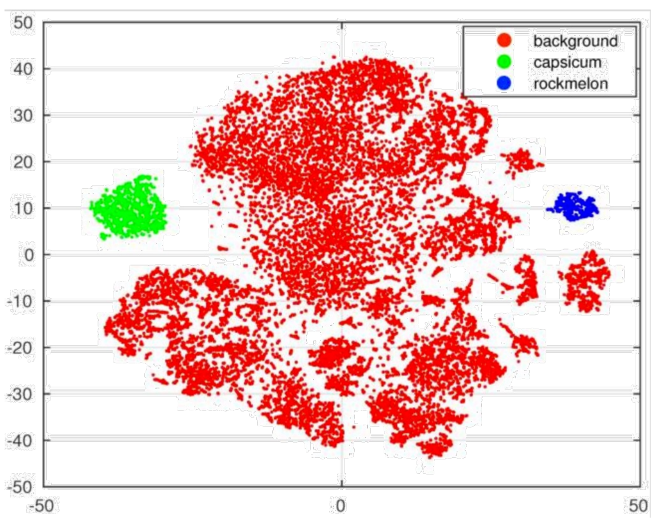


Unsupervised learning

Only attributes are known, no labels are given

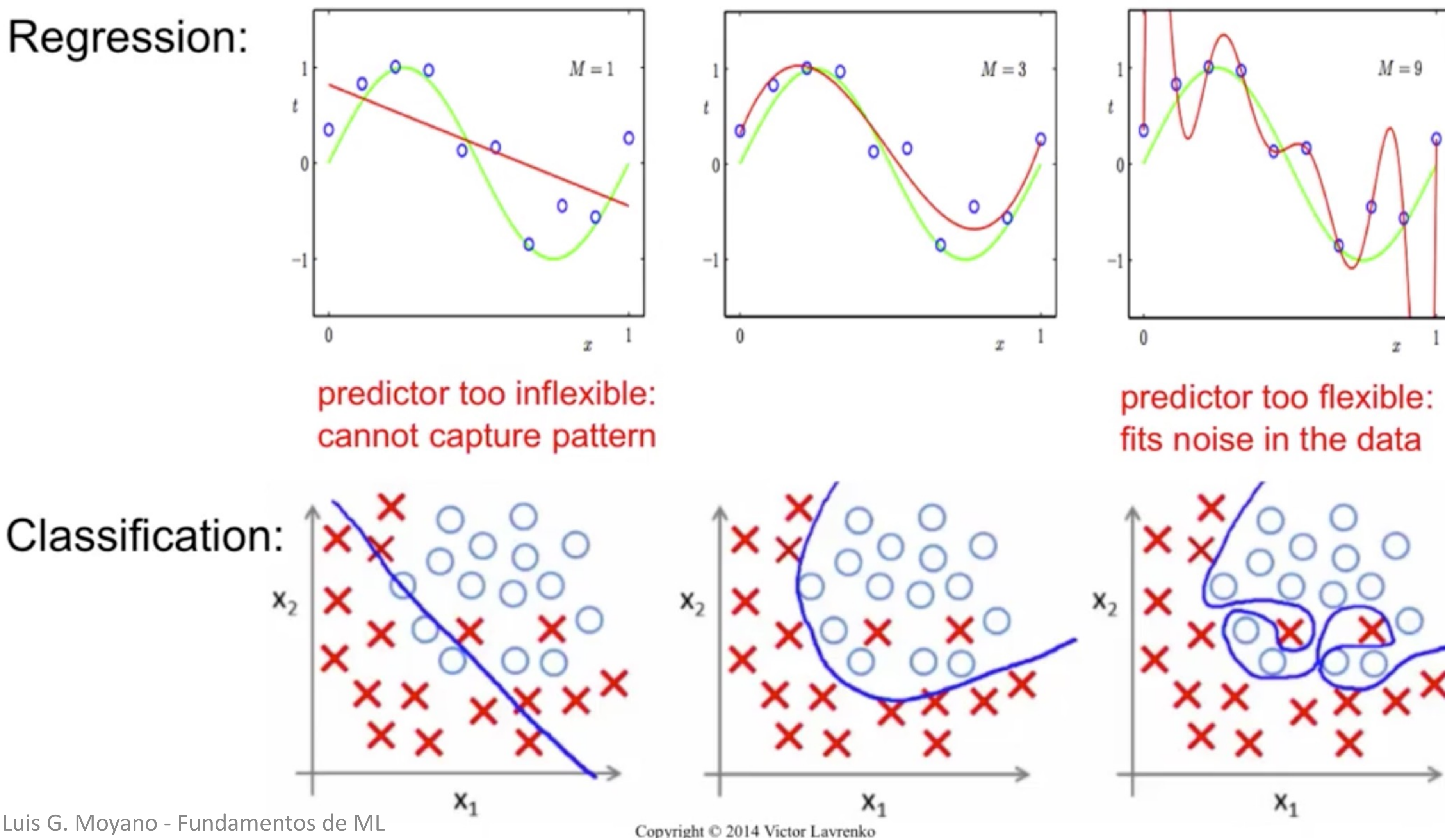






Sensors (Basel). 2016 Aug; 16(8): 1222.
Published online 2016 Aug 3. doi: 10.3390/s16081222

Overfitting/Underfitting the Training Data



Supervised/Unsupervised

A blurry definition

- Unsupervised learning and supervised learning are not formally defined terms.
  - Unsupervised from supervised:

$$p(\mathbf{x}) = \prod_{i=1}^{n} p(\mathrm{x}_i \mid \mathrm{x}_1, \ldots, \mathrm{x}_{i-1}).$$

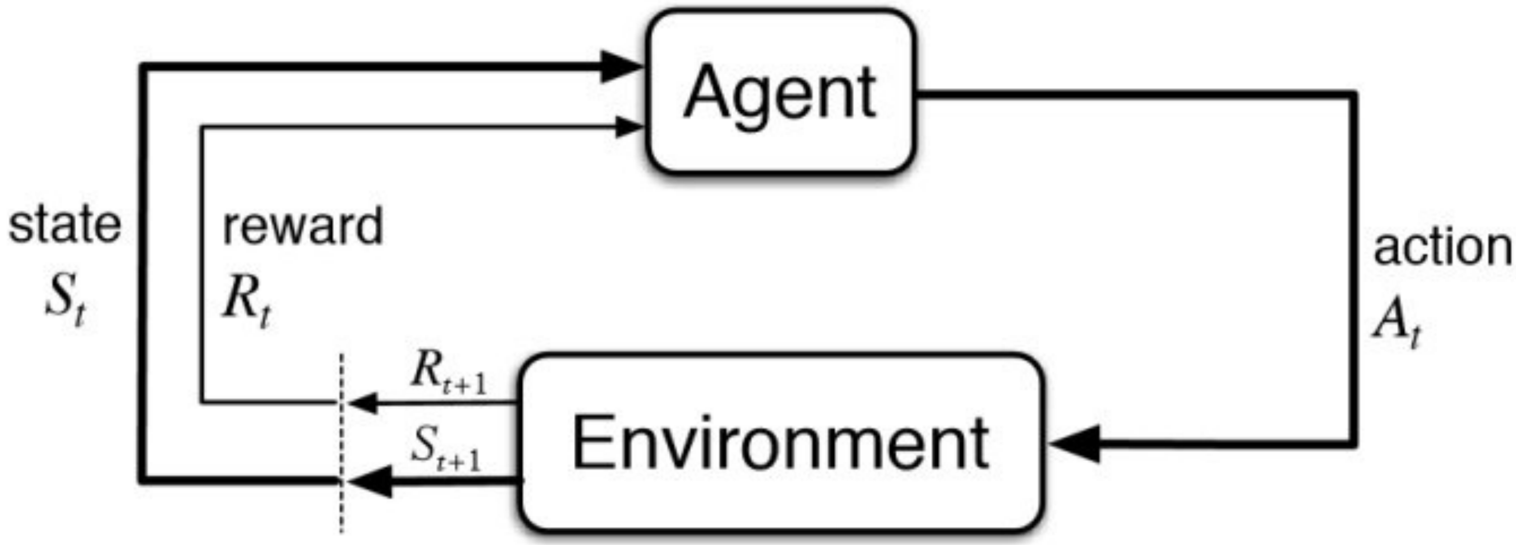
  - Supervised from unsupervised:

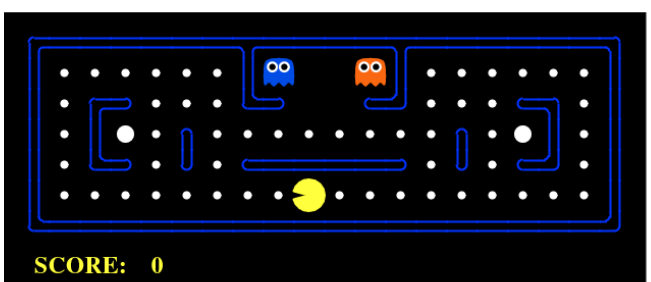
$$p(y \mid \mathbf{x}) = \frac{p(\mathbf{x}, y)}{\sum_{y'} p(\mathbf{x}, y')}.$$

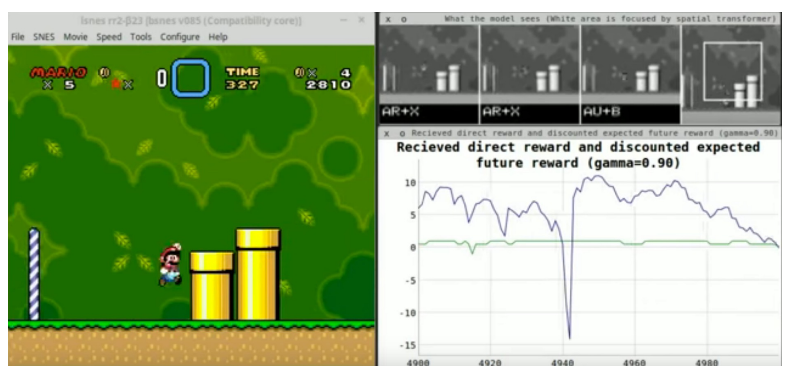
In 2016 - Goodfellow - Deep learning, p105

Reinforcement Learning

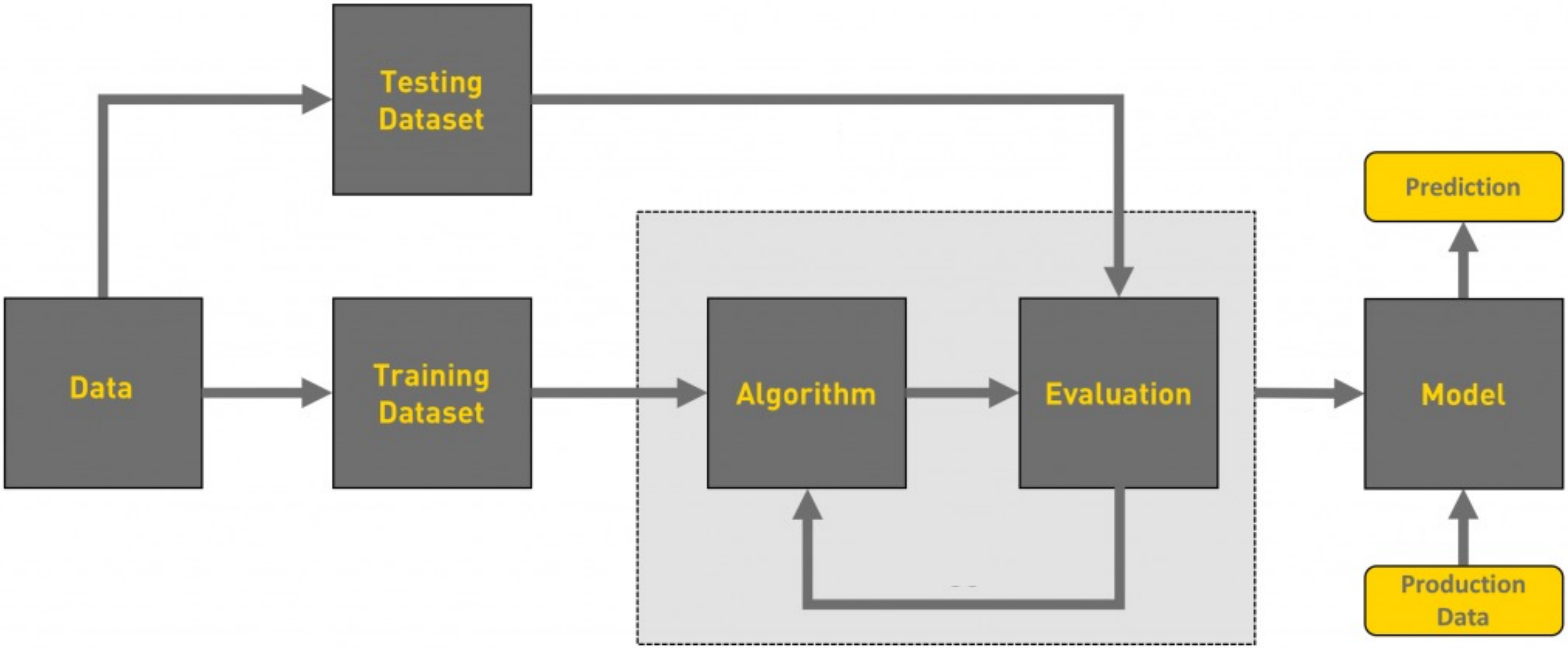
How to make good decisions, by trial and error



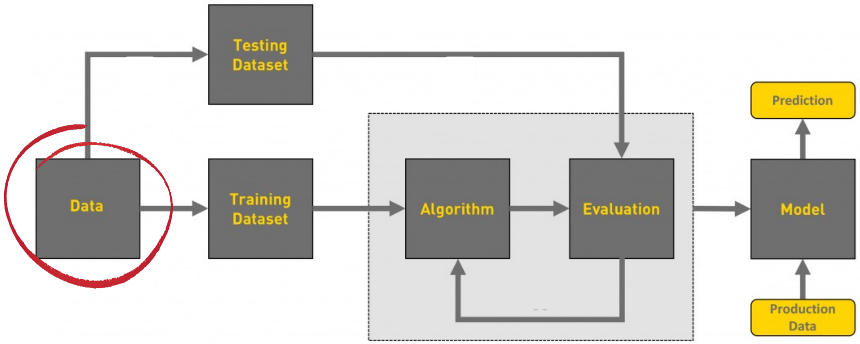




A canonical ML workflow, end-to-end

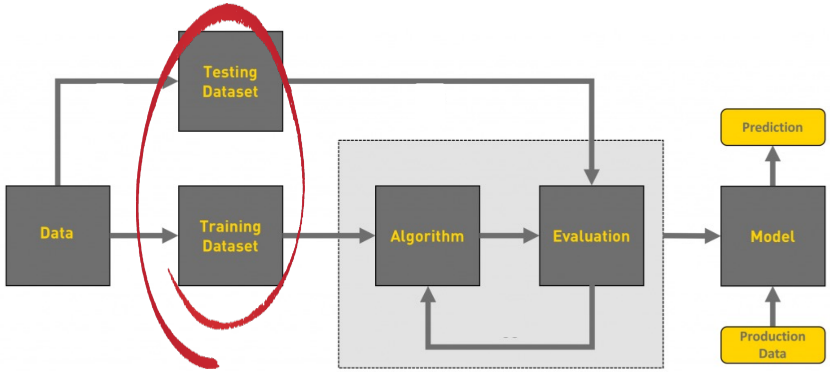


Data preprocessing

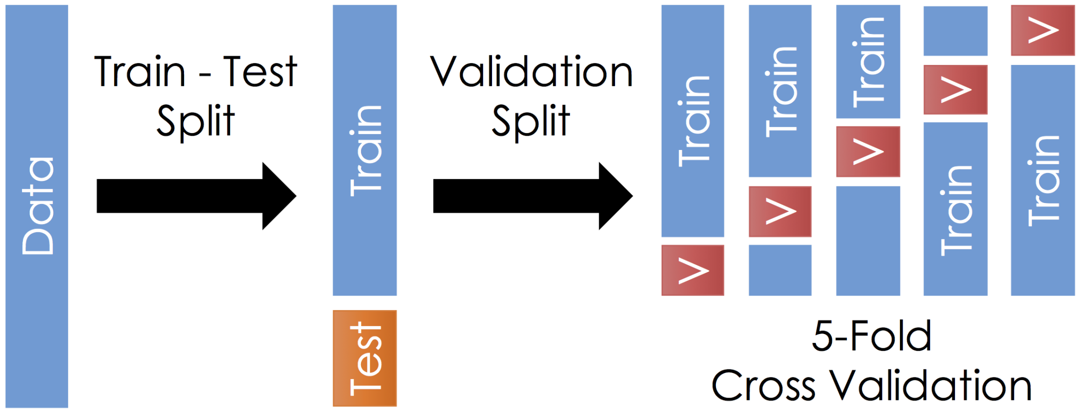


- Data integration
- Data cleaning
- Feature analysis (correlations, importance)
- Feature engineering
- Near zero variance,, rescaling, other transformations

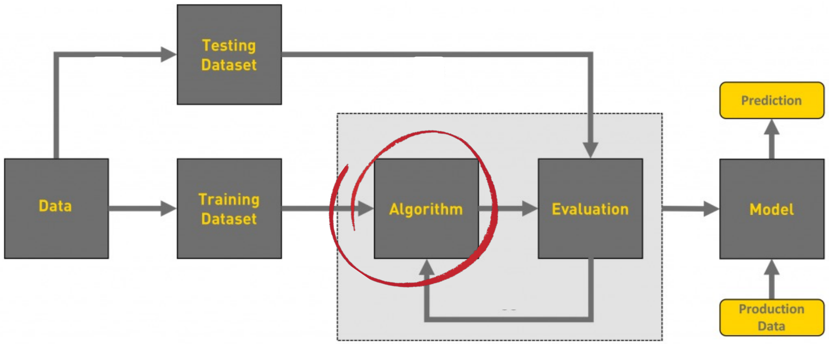
Error estimation



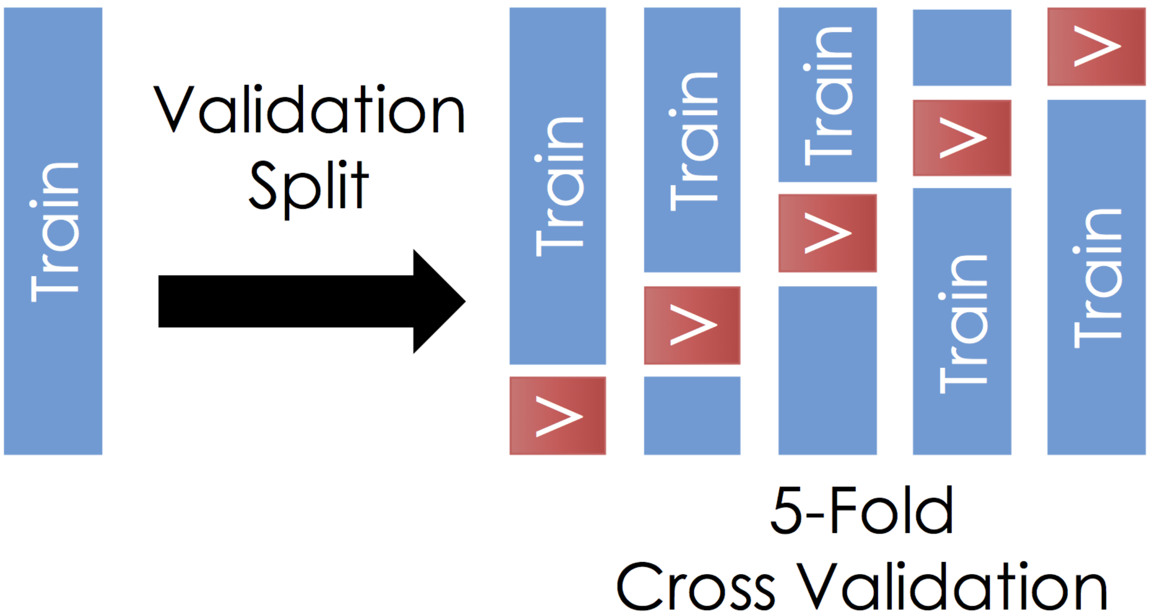
- Model selection: estimating the performance of different models in order to choose the best one.
- Model assessment: having chosen a final model, estimating its prediction error (generalization error) on new data.

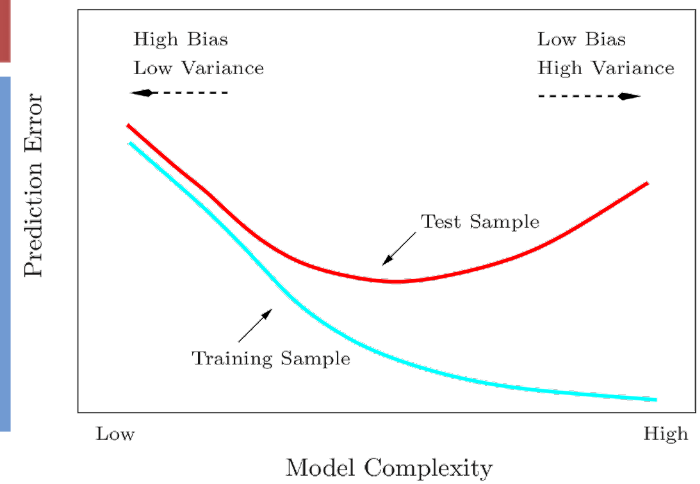


Model selection

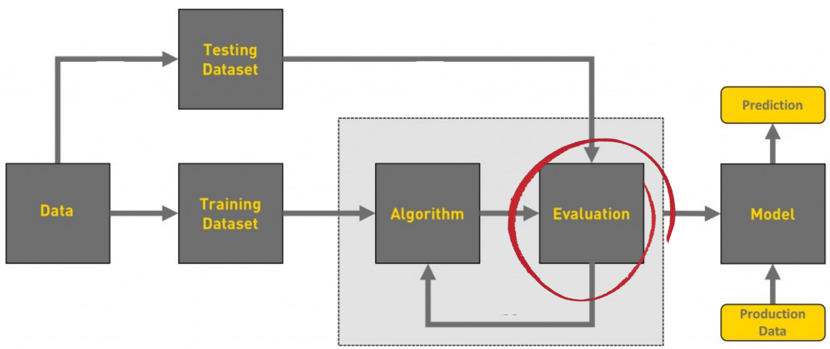


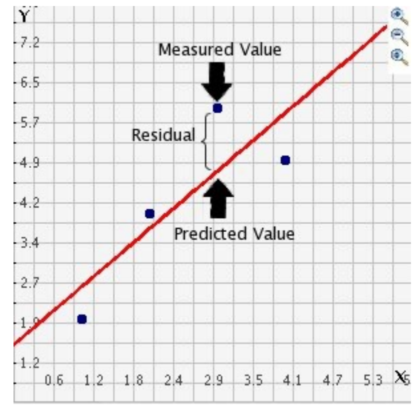
- ML task (classification, regression, unsup.)
- Model (Linear regression? Deep learning?)
- Model hyperparameters (resampling)

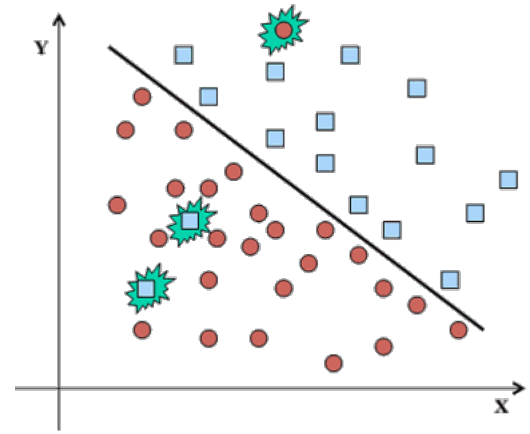


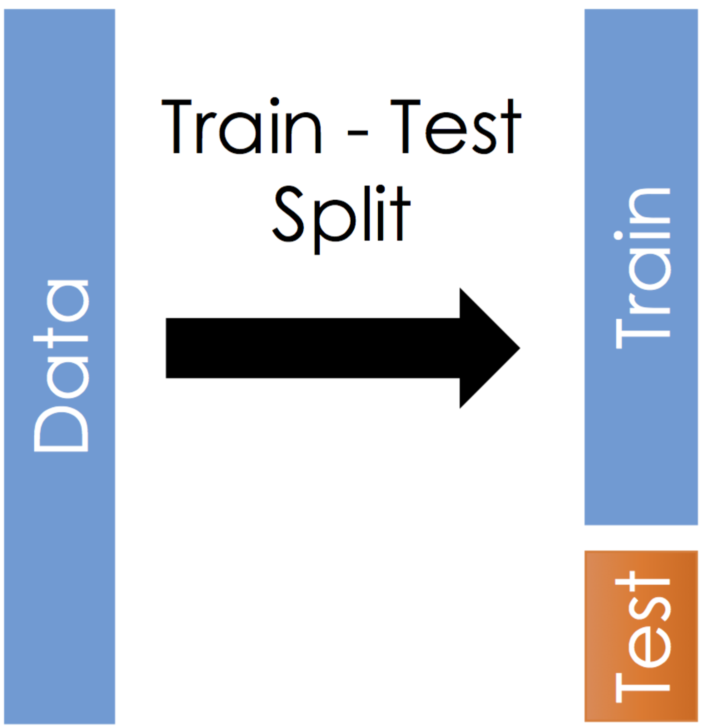


Model evaluation









Equation 2-1. Root Mean Square Error (RMSE)

$$\mathrm{RMSE}(\mathbf{X}, h) = \sqrt{\frac{1}{m} \sum_{i=1}^{m} \left(h\left(\mathbf{x}^{(i)}\right) - y^{(i)}\right)^2}$$

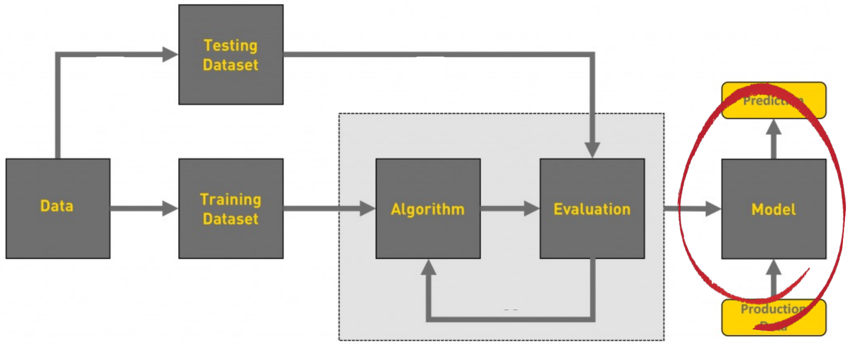
Equation 3-2. Recall

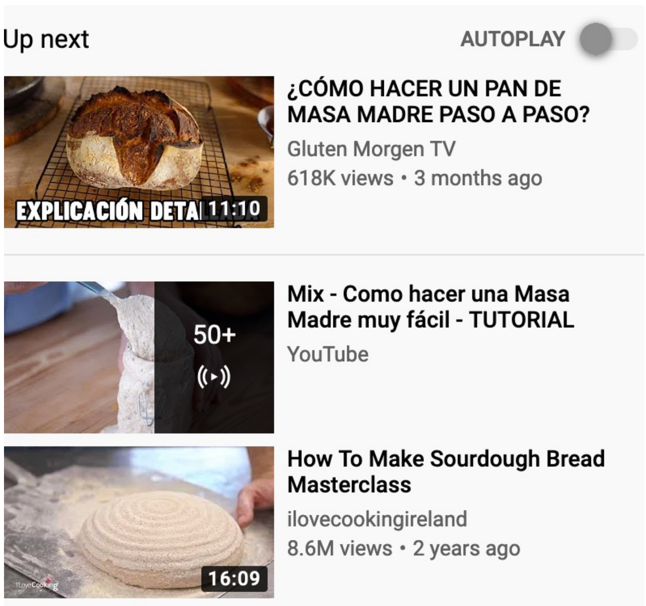
$$\text{recall} = \frac{TP}{TP + FN}$$

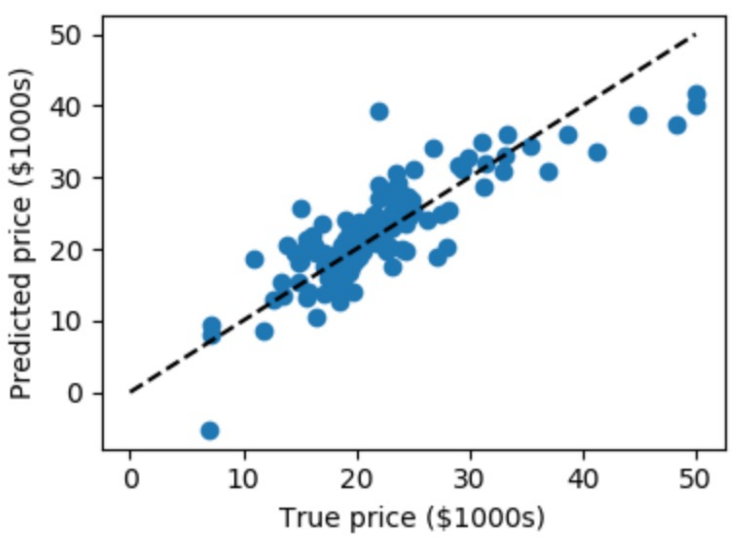
Equation 3-1. Precision

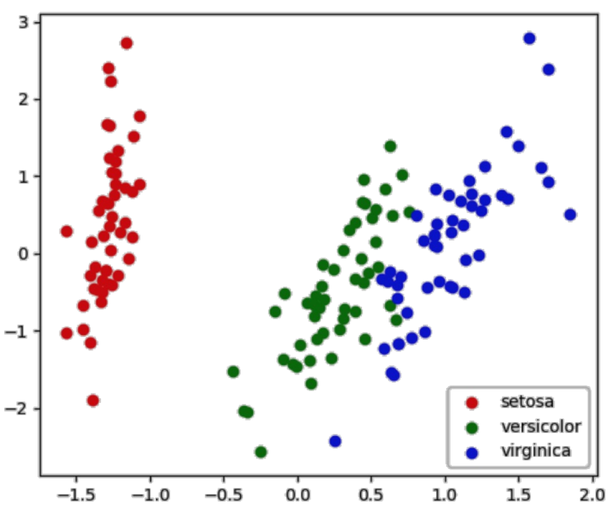
$$\text{precision} = \frac{TP}{TP + FP}$$

Prediction

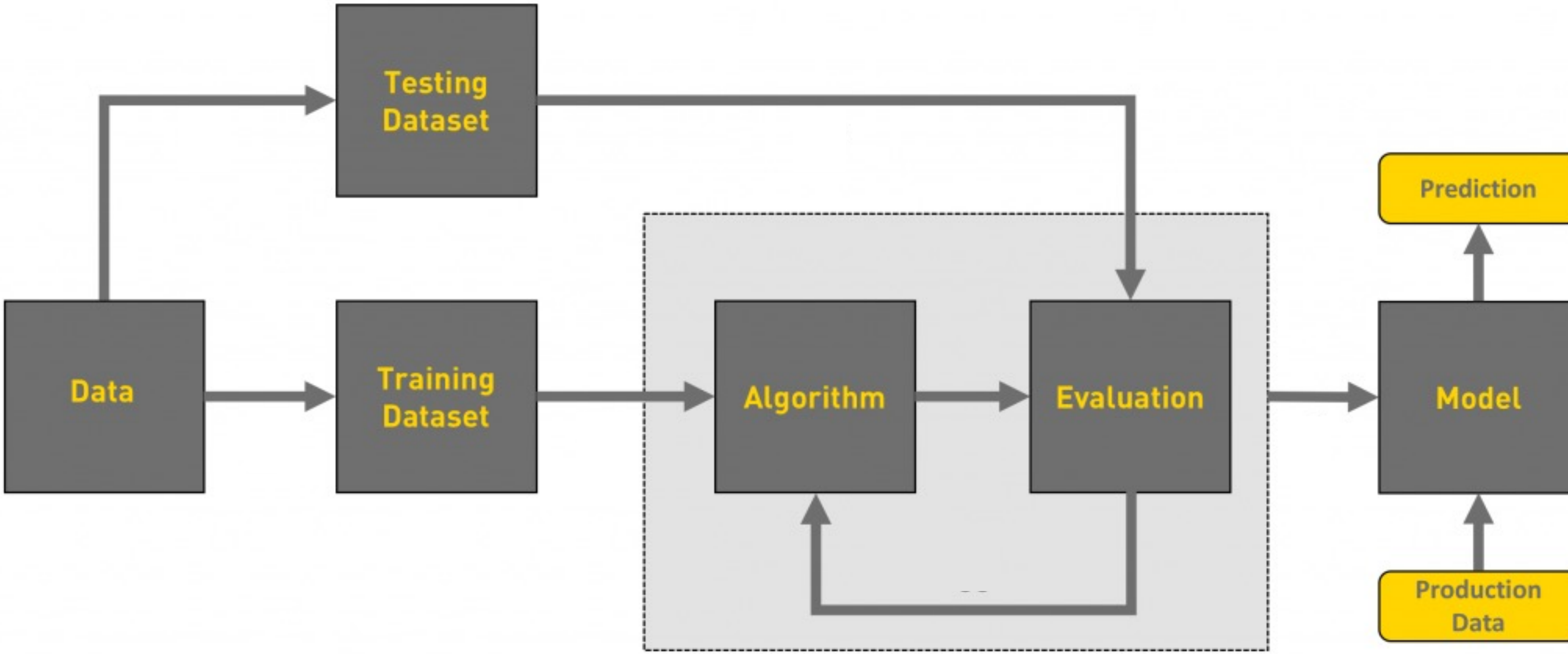

Testing Dataset
Data
Training Dataset
Algorithm
Evaluation
Model


Up next
AUTOPLAY
EXPLICACIÓN DETAl
11:10
¿CÓMO HACER UN PAN DE MASA MADRE PASO A PASO?
Gluten Morgen TV
618K views • 3 months ago
50+
Mix - Como hacer una Masa Madre muy fácil - TUTORIAL
YouTube
16:09
How To Make Sourdough Bread Masterclass
ilovecookingireland
8.6M views • 2 years ago

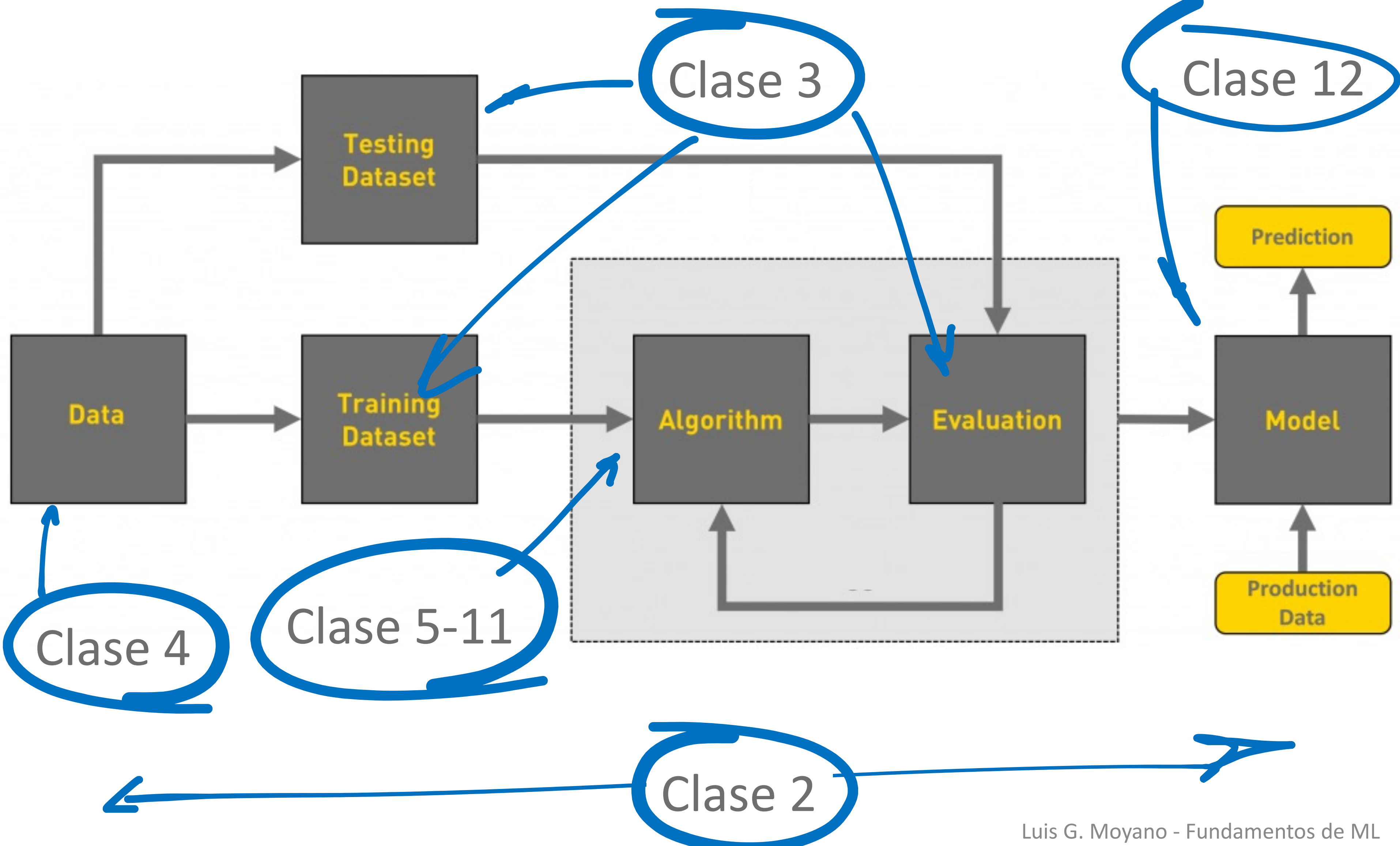

Predicted price ($1000s)
True price ($1000s)


setosa
versicolor
virginica

A canonical ML workflow, end-to-end



A canonical ML workflow, end-to-end



Next

- **Code** - Python and the ML ecosystem
  - Python: bring anaconda/miniconda, & other packages installed
  - R: bring caret/tidiverse installed
- **Special session:** Jupiter Notebooks & Google Collab (Martín Onetto)
  - End-to-end ML workflow special session
- **Practice time:**
  - Git/Github special session
  - Open a Github account
  - Install git, anaconda & ML modules