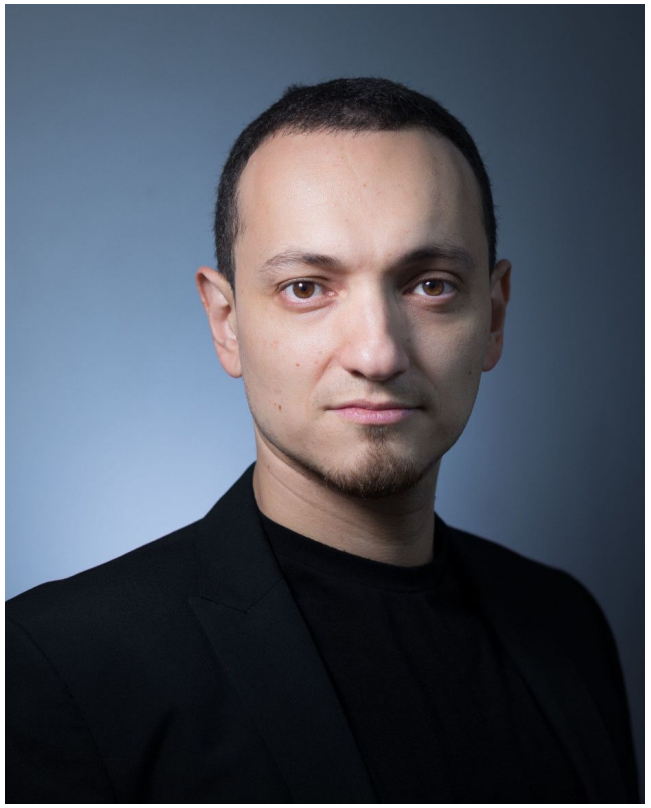
TECNOLÓGICO DE MONTERREY®

A brief history of AI

POSGRADOS | MAESTRÍAS DIGITALES

Meet the Prof.



Dr. Jose Antonio Cantoral Ceballos
JoseAntonio.Cantoral@tec.mx
Research Professor in Advanced Artificial Intelligence
Data Science Hub
School of Engineering and Sciences
Tecnologico de Monterrey
NVIDIA University Ambassador / NVIDIA Certified Instructor

https://mx.linkedin.com/in/jacantoral

https://www.youtube.com/c/PepeCantoralPhD

https://github.com/JACantoral/

AI is new, right?

BnF Museum, Public domain, via Wikimedia Commons

Talos in Greek Mythology

One of the earliest examples comes from ancient Greek mythology. Talos was a giant bronze automaton created by Hephaestus, the god of metalworking, to protect the island of Crete. Talos is an early example of an artificial being created to perform specific tasks.

Golem in Jewish Folklore

In Jewish folklore, there is the concept of the Golem, an animated anthropomorphic being magically created from clay or mud. The Golem was often created to serve its creator or protect a community.

Thander, Public domain, via Wikimedia Commons

AI is new, right?

Photograph Theatre Guild touring company's 1928–1929 production of *R.U.R.* by Karel Čapek.

"R.U.R" by Karel Čapek

The term "robot" was first introduced in the 1920 play "R.U.R." or "Rossum's Universal Robots" by Czech playwright Karel Čapek. The play explored the creation of artificial beings used for labor and eventually revolt against their creators.

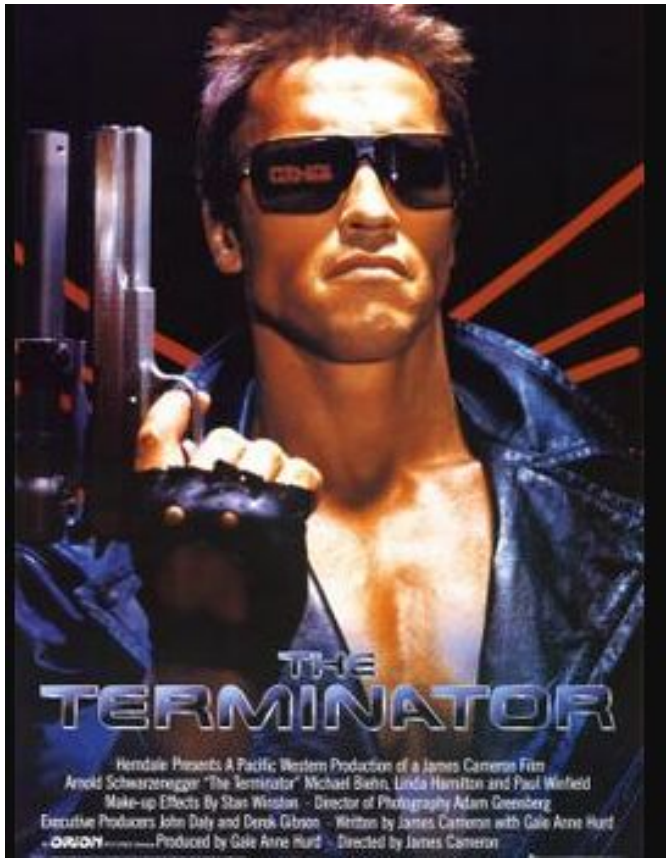
AI is new, right?

Metropolis

The 1927 German expressionist film "Metropolis" directed by Fritz Lang features a robot called "Maria." It's one of the earliest portrayals of a robot in film and deals with issues of industrialization and class struggle.

Illustration by Heinz Schulz-Neudamm.
Distributed by UFA.

Newer takes …


THE
TERMINATOR
Hemdale Presents A Pacific Western Production of a James Cameron Film
Arnold Schwarzenegger "The Terminator" Michael Biehn, Linda Hamilton and Paul Winfield
Make-up Effects By Stan Winston - Director of Photography Adam Greenberg
Executive Producers John Daly and Derek Gibson - Written by James Cameron with Gale Anne Hurd
ORION Produced by Gale Anne Hurd - Directed by James Cameron


28 DAYS LATER
EX_MACHINA
THERE IS NOTHING MORE HUMAN
THAN THE WILL TO SURVIVE
COMING SOON

What is AI - A brief history

The gestation (1943 - 1950)

Warren McCulloch and Walter Pitts (1943). Introduce the concept of artificial neuron, with ‘on’ and ‘off’ states. They suggested that carefully designed networks could learn.

In 1949, Donald Hebb demonstrated the idea of adjustable connection strengths.

In 1950, Marvin Minsky and Dean Edmons, built the first NN computer (the SNARC).

Alan Turing publishes “Computing Machinery and Intelligence” in 1950, where he introduced the concepts of Turing test, machine learning, genetic algorithms, and reinforcement learning.

Thinking humanly and the Turing Test (1950)

In 1950 Turing posed the question 'Can machines think?'

He proposed a test, known as the Turing Test, to determine if a machine exhibits human-like intelligence.

A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer.

Physical simulation of a person is unnecessary for intelligence.

The birth - Darmouth Summit (1956)

In the summer 1956, John McCarthy, Marvin Minsky, Claude Shannon, and Nathaniel Rochester organised a two-month workshop at Dartmouth, where the term AI was coined by McCarthy.

They proposed that during this 2 month period, 10 scientist studying AI would attempt to make machines use language, form abstractions, concepts, solve problems that were reserved for humans and improve themselves.

Allen Newell and Herbert Simon delivered the Logic Theorist, which was capable of proving theorems.

Promising childhood (1952 - 1969)

General Problem Solver, Physical Symbol System, Geometry theorem Prover, checkers.

Lisp Programming Language by John McCarthy,

Using knowledge to find solutions to problems, explicit representation of the world to be manipulated with deductive processes. In 1965, Robinson introduced the resolution method theorem proving algorithm.

Minsky works on limited domains, microworlds.

Rosenblatt introduces the Perceptron in 1957, whose convergence theorem states that learngin algorithms can adjust the connection strengths to match any input data

Reality hit (1966 - 1973)

General over-promising by AI community.

Systems with early successes on simple problems, failed miserably on wider selection of problems.

US government cancels funding for machine translation problems.

British government cancels funding for all but two Universities.

Minsky and Papert publish “Perceptrons” book in 1969, which proved their representation limitations, even when they could learn what they could represent.

Winter 1

Some problems during this time

Computational power was limited

Problem formulation relied in exponential search, so no matter the compute power, the amount of information to search would not be enough.

Some contributions during this time

Lisp

Garbage collection

Knowledge-based systems (1969 – 1979)

DENDRAL

One of the earliest expert systems, aimed at inferring possible molecular structures from mass spectrometry data. It was initially developed by Edward Feigenbaum and others at Stanford University. DENDRAL stands for "Deductive ENumerator of Dihedral ALternatives."

MYCIN

Designed at Stanford University by Edward Shortliffe under the guidance of Bruce Buchanan and Edward Feigenbaum, MYCIN was an expert system for diagnosing bacterial infections and recommending antibiotic treatments. Despite its efficacy, it was never adopted in clinical settings mainly due to lack of a user-friendly interface and concerns about liability.

Expert systems (1980s)

R1 is the first commercial expert system, to configure orders for new computer systems.

Strong funding from Japan, USA, and Britain.

AI industry boomed to billions of dollars by 1988. Expert systems, vision systems, robots, and specialized software and hardware.

Complex and expensive systems, along with failure to deliver extravagant promises, lead to AI winter 2

Expert systems (1980s)

IBM Deep Blue beats Gary Kasparov in 1997.

Hard-coded knowledge about the world in formal languages to create a knowledge base.

Abstracting human knowledge is not an easy task.
Expensive.

Difficult to update.
Mistakes when given unusual data

The recent times

Neural Networks Milestones

1943 – Artificial Neural Networks
1957 – The perceptron
1969 – ‘Perceptrons’ by Minsky and Papert
1970’s – 1980’s Backpropagation (Winter number 1)
1980’s – 1990’s – CNN, MNIST, RNN, LSTM (Winter number 2)
2006 - ‘Deep learning’, Restricted Boltzman machine, DBNs
2009 - ImageNet
2012 – AlexNet!

Rise of the machine

Support Vector Machines (SVM)

Early 1990s
Vapnik and Cortes introduced the concept of Support Vector Machines, which became a popular method for classification tasks. SVMs are particularly known for their effectiveness in high-dimensional spaces and are widely used in various domains such as text classification, image recognition, and bioinformatics.

Decision Trees and Random Forests

1980s, but popularized in the 1990s
Decision Trees like ID3 and C4.5 were developed in the 1980s but gained widespread use in the 1990s. The Random Forest algorithm, introduced by Leo Breiman in 2001, improved upon basic decision tree algorithms by creating an ensemble of trees that vote on the final classification or regression result.

Neural Networks and Backpropagation

Renaissance 1990s
Although the concept of neural networks and the backpropagation algorithm were known earlier, they gained renewed interest in the 1990s thanks to increased computational capabilities and some improvements in the algorithms, like the introduction of better activation functions.

Rise of the machine

Natural Language Processing (NLP)

Late 1990s - 2000s
Techniques like Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA), and others enabled better text understanding and classification. This set the stage for more advanced NLP algorithms and the use of machine learning in language technologies.

Recommender Systems

Late 1990s (Amazon's Item-to-Item Collaborative Filtering)
Recommender systems started becoming essential for online commerce, most notably with Amazon's recommendation engine. These systems used machine learning algorithms to suggest products to customers based on their past behavior.

Deep Learning

Late 2000s
A key breakthrough came in 2012 when AlexNet won the ImageNet competition, marking the beginning of the deep learning era. This period also saw the development of frameworks like TensorFlow and PyTorch that have since accelerated machine learning research and applications.

ImageNet

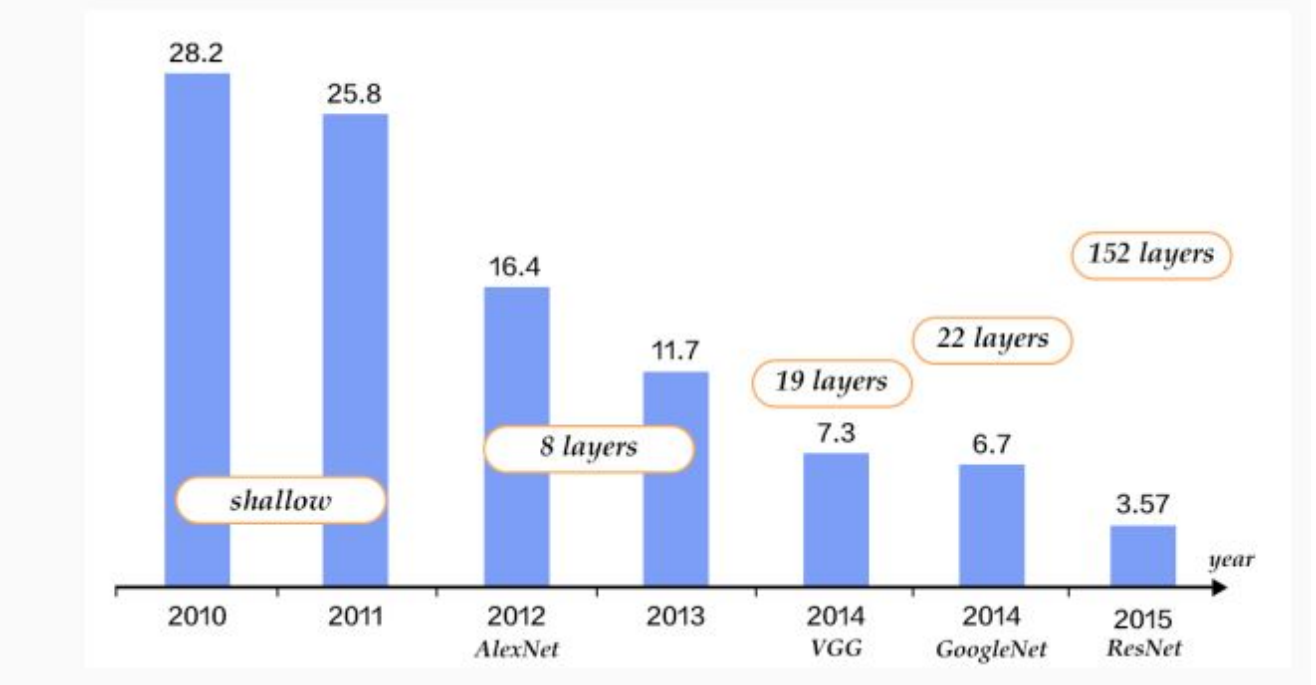
Over 14 million hand-annotate images, categorized into over 20,000 classes

ImageNet Large Scale Visual Recognition Challenge (ILSVRC) ~ 1200 images for each 1000 categories in 2012

AlexNet on ILSVRC

Over 14 million hand-annotated images, categorized into over 20,000 classes

ImageNet Large Scale Visual Recognition Challenge (ILSVRC) ~ 1200 images for each 1000 categories in 2012



Top-5 error rates on ILSVRC image classification over Time

https://www.cs.toronto.edu/~rgrosse/courses/csc321_2018/tutorials/tut6_slides.pdf

Geoffrey Hinton - The Godfather of Deep Learning

DNNresearch Inc. in 2012

Auction: In a move that felt almost like a scene from a movie, the company was put up for auction, with major tech companies showing interest.

Google's Win: Eventually, Google won the auction and acquired DNNresearch in 2013.

The Deal: Although exact terms are not public, the acquisition signaled Google's deep commitment to leading in AI.

Geoffrey Hinton - The Godfather of Deep Learning

DNNresearch Inc. in 2012

Auction: In a move that felt almost like a scene from a movie, the company was put up for auction, with major tech companies showing interest.

Google's Win: Eventually, Google won the auction and acquired DNNresearch in 2013.

The Deal: Although exact terms are not public, the acquisition signaled Google's deep commitment to leading in AI.

Geoffrey Hinton - The Godfather of Deep Learning

1943 – Artificial Neural Networks
1957 – The perceptron
1969 – ‘Perceptrons’ by Minsky and Papert
1970’s – 1980’s Backpropagation (Winter number 1)
1980’s – 1990’s – CNN, MNIST, RNN, LSTM (Winter number 2)
2006 - ‘Deep learning’, Restricted Boltzman machine, DBNs
2009 - ImageNet
2012 – AlexNet!

How we got here!

1943 – Artificial Neural Networks
1957 – The perceptron
1969 – ‘Perceptrons’ by Minsky and Papert
1970’s – 1980’s Backpropagation (Winter number 1)
1980’s – 1990’s – CNN, MNIST, RNN, LSTM (Winter number 2)
2006 - ‘Deep learning’, Restricted Boltzman machine, DBNs
2009 - ImageNet
2012 – AlexNet!

Since AlexNet

AlphaGo

- **Year**: 2016
- **Organization**: DeepMind
- **Significance**: First computer program to defeat a world champion Go player, a feat considered a major milestone due to the complexity of Go compared to previous games tackled by AI like chess. AlphaGo used deep convolutional neural networks and Monte Carlo Tree Search (MCTS) for its achievement.
-

AlphaFold

- **Year**: 2018, 2020
- **Organization**: DeepMind
- **Significance**: Solved the problem of protein folding, a grand challenge in biology. Its 2020 version outperformed 100 other teams in a biennial protein-folding competition known as CASP. It represents a monumental leap in bioinformatics, with huge implications for drug discovery and understanding diseases.

Since AlexNet

MuZero

- **Year**: 2019
- **Organization**: DeepMind
- **Significance**: Generalized the concept of model-based reinforcement learning. Unlike its predecessor AlphaZero, it operates efficiently without knowing the rules of the environment, making it more versatile for real-world applications.
-

BERT (Bidirectional Encoder Representations from Transformers)

- **Year**: 2018
- **Organization**: Google
- **Significance**: Created a new paradigm for natural language processing tasks. It uses bidirectional training and transformers to understand the context and semantics of words. Google uses BERT to better understand user searches.

Since AlexNet

GPT Models

- **Year**: 2018 (GPT), 2019 (GPT-2), 2020 (GPT-3), 2021 (GPT-4, to the best of my knowledge)
- **Organization**: OpenAI
- **Significance**: These models, particularly GPT-3 and the subsequent versions, set new benchmarks for a variety of natural language understanding and generation tasks. They've opened new avenues in machine-human interaction, content creation, programming help, and much more.

DALL·E

- **Year**: 2021
- **Organization**: OpenAI
- **Significance**: An extension of the GPT-3 model trained to generate high-quality images from textual descriptions, representing another step towards multimodal AI systems.

Since AlexNet

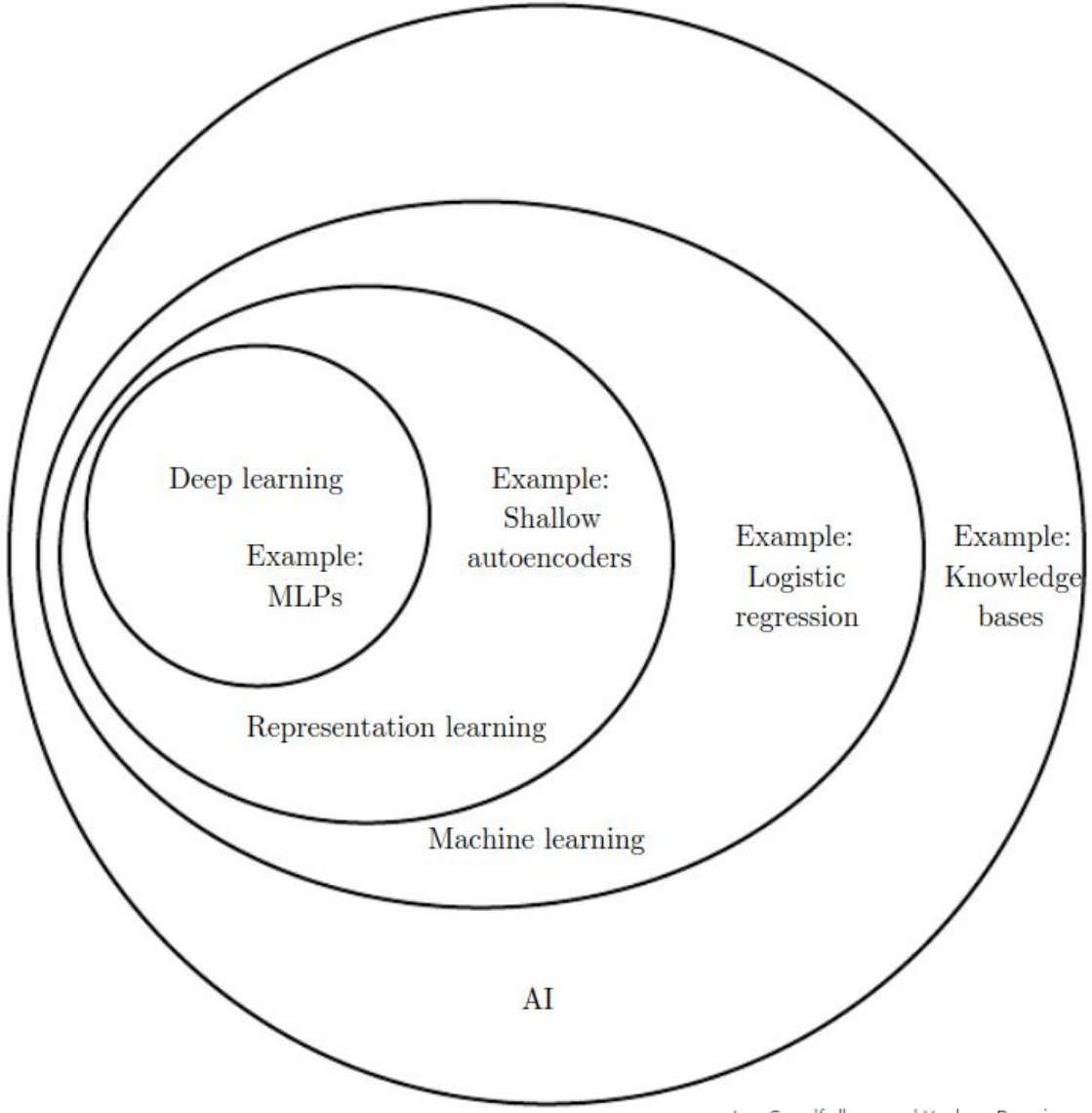
Tesla Autopilot

- **Year**: Ongoing
- **Organization**: Tesla
- **Significance**: Though not without controversy, Tesla's efforts in using deep learning for full self-driving capabilities have set significant milestones in the practical application of AI in everyday life.

Neuralink

- **Year**: Ongoing
- **Organization**: Neuralink Corp.
- **Significance**: Elon Musk's company is working on integrating AI with the human brain, aiming to address neurological issues and eventually facilitate human-AI symbiosis.

Machine Learning vs GOFAI



Ian Goodfellow and Yoshua Bengio and Aaron Courville, Deep Learning, 2016, MIT Press, http://www.deeplearningbook.org

Colloquially, according to Russell and Norvig (2009), AI is the ability in machines that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving".

Machine Learning (ML)

AI systems acquire their own knowledge, by extracting patterns from raw data.

ML enabled tackling problems involving machine knowledge of the real world.

The performance of these simple ML algorithms depends heavily on the representation of the data they are given.

Machine Learning (ML)

AI systems acquire their own knowledge, by extracting patterns from raw data.

ML enabled tackling problems involving machine knowledge of the real world.

The performance of these simple ML algorithms depends heavily on the representation of the data they are given.

Supervised learning

Abstracting patterns from many labelled (annotated) examples.

The algorithm adjusts parameters to learn to produce the outcome we desire, i.e. predict unknown examples labels.

Most of the current achieved landmarks have used some sort of supervision to achieve results, e.g.

- Speech recognition
- Image classification
- Image captioning
- Language translation

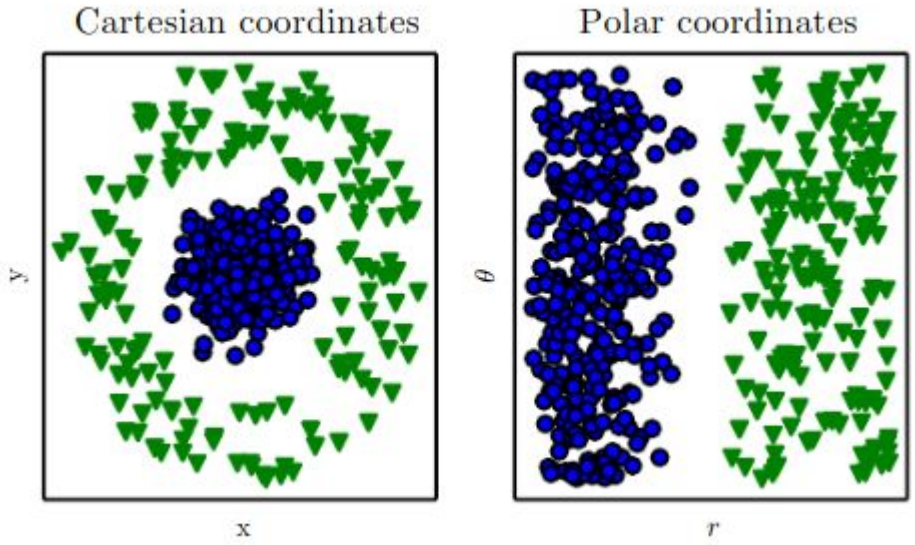
Constraints?

We need lots of labelled data!

We need fairly powerful hardware, although we can achieve a lot with small networks trainable in a laptop.

From Machine Learning to Deep Learning

Why is DL so famous whilst ML is probably not?
Data representation is Key!
We want to separate two categories of data by drawing a line between them in a scatterplot, which one is better?



From Goodfellow, et. al., Deep Learning, 2015

Learned representations often result in much better performance than the one that can be obtained with hand-designed representations.

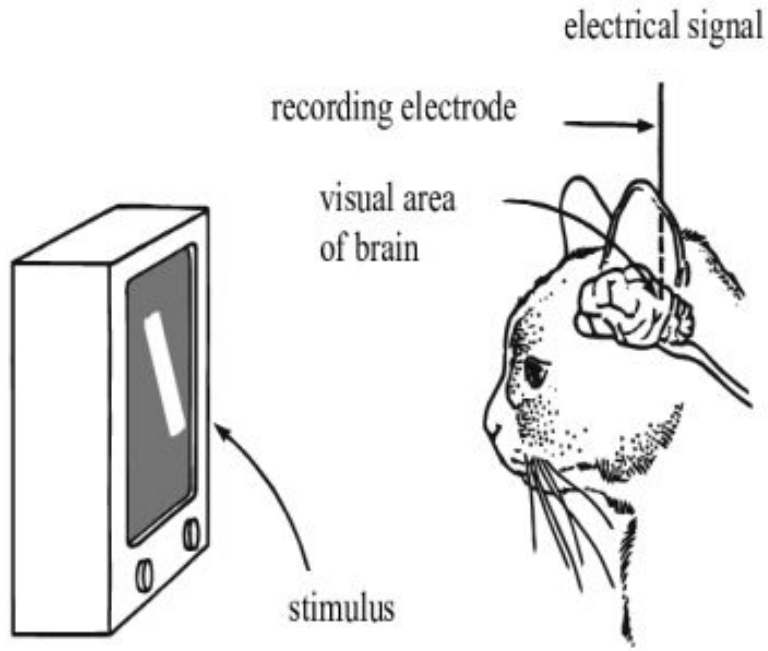
Deep Learning

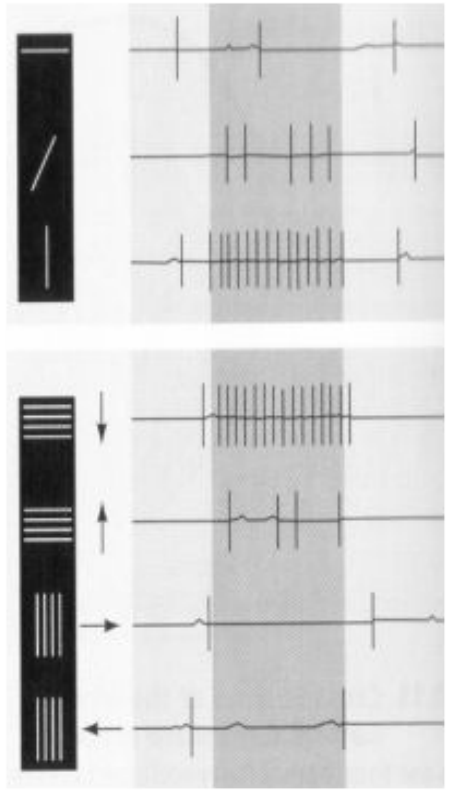
DL achieves great power and flexibility by representing the world as a nested hierarchy of concepts, with each concept defined in relation to simpler concepts, and more abstract representations computed in terms of less abstract ones.

Abstract useful patterns directly from data by allowing the algorithm to build complex concepts out of simpler concepts.

Something interesting about Deep Learning

David Hubel & Torsten Weisel experiments





Hubel and Wiesel,
1959, Nobel prize
1981, edges

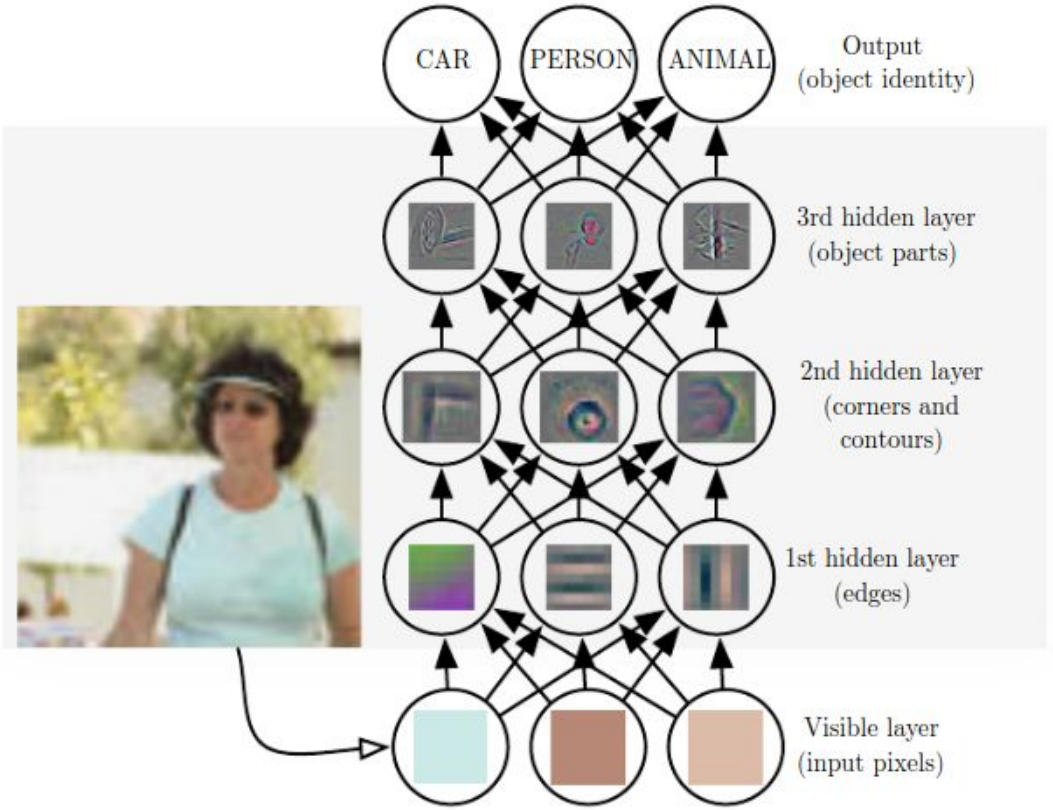
Something interesting about Deep Learning

DL is a type of machine learning, that enables computer systems to improve with experience and data.

DL breaks the desired complicated mapping into a series of nested simple mappings, each described by a different layer of the model.

A series of hidden layers extracts increasingly abstract features from the image.

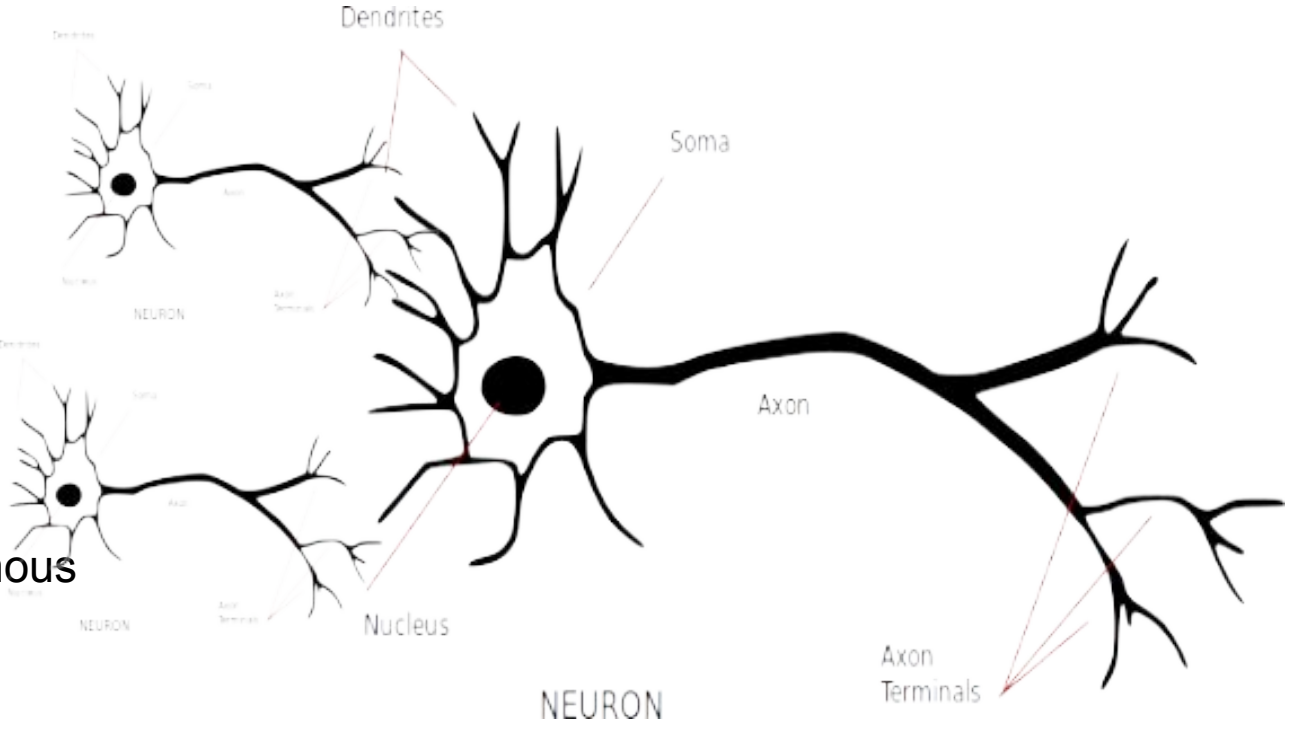
The model must determine which concepts are useful for explaining the relationships in the observed data.



Deep Learning Inspiration- Neural Networks

What is a Neuron?
Biological Inspiration
- Human brain ~ 86 billion neurons
- 1000 trillion synapses (1,000 x 1012)
- Complex Artificial NN ~ +500 billion parameters
- ~ Still orders of magnitude more complex

Common mediatic misconception
- There is some hierarchy but probably asynchronous
- Learning algorithm?
- Power consuption?
- Learning does not end



Thanks!

Tecnológico de Monterrey