Biomimetic Neuroscience

Deep Learning

Rodrigo Ramele
Lab Neurotrónica
Departamento de Ingeniería Informática

Neuroscience
Control
Mechanics
Informatics
Electronics

The Brain

Frenología

Herculano-Souzel (2009)

USC Mark and Mary Stevens
Neuroimaging and Informatics Institute
Play
Keck Schoo
Medicine of

Thanks Phineas![60]

Divisions and Broadmann Mappings

Primary Areas

Primary Areas

Brain Ventricles and Cerebrospinal Fluid

Hemispheres and Corpus Callosum - Network

Subcortical - firmware

Occipital - Human GPU

Frontal Area - Brocca

Natural Language Processing

Frontal - Language

Frontal - Neocortex

The prefrontal cortex as a controller

- Collection of interconnected neocortical areas in the frontal lobe
- Performs **executive functions** that are important for the **conscious control of thought and action** in accordance with **internal goals**
- Dorsolateral PFC: abstract reasoning & problem solving
- Orbitofrontal and medial cortex: affective & motivational functions

http://pubs.niaaa.nih.gov/publications/arh313/215-230.htm

Temporal - Pre, Motor and Somatosensorial

Temporal - Hippocampus and Neurogenesis (RAM)

Krzisch, Temprana, Mongiat 2014

Temporal - Tempo parietal junction

Neuron

Neuron

Dendrites
Microtubule
Neurofibrils
Neurotransmitter
Receptor
Synapse
Synaptic vesicles
Synapse (Axoaxonic)
Synaptic cleft
Axonal terminal
Rough ER
(Nissl body)
Polyribosomes
Ribosomes
Golgi apparatus
Node of Ranvier
Synapse
(Axosomatic)
Myelin Sheath
(Schwann cell)
Axon hillock
Nucleus
Nucleolus
Membrane
Microtubule
Mitochondrion
Smooth ER
Synapse
(Axodendritic)
Nucleus
(Schwann cell)
Microfilament
Microtubule
Axon
Dendrites

Hodgkin-Huxley Model

$$I = -(C_m \frac{dV}{dt}) + \frac{Ie}{A}$$

$$I = C_m \frac{dV_m}{dt} + g_K(V_m - V_K) + g_{Na}(V_m - V_{Na}) + g_l(V_m - V_l)$$

Hodgkin 1952, *Abott* "Theoretical Neuroscience"

Plasmatic Membrane - Na+/K+ Pumps

+Mantiene el volumen de la célula.

+Puede funcionar en reversa.

+Canal de K^+ selectivo de cationes mediante afinidad de terminaciones con grupos hidroxilos con el K^+ deshidratado.

+Sensible al voltaje

+Determina la permeabilidad.

Membrane Potential

- El Na^+ por los Canales de fuga **positiviza** el interior al entrar.
- El K^+ por los Canales de fuga **negativiza** el interior al salir.
- La bomba de Na^+-K^+ saca 3 contra 2 (todos positivos) con lo que **negativiza** el interior.

Ecuación Nerst para cada ion.

$$V = -61 \cdot \log\left(\frac{CNa_i^+ PNa^+ + CK_i^+ PK^+ + CCl_o^- PCl^-}{CNa_o^+ PNa^+ + CK_o^+ PK^+ + CCl_i^- PCl^-}\right)$$

$$+61mV - 94mV = -86mV - 4mv = -90mv$$

Membrane Potential

Equilibrio para cada ion

$$Na^+ \quad 62 \ \log \frac{145}{5} = 90mV$$

$$62 \ \log \frac{145}{15} = 61mV$$

$$K^+ \quad 62 \ \log \frac{5}{140} = -90mV$$

$$Cl^- \ -62 \ \log \frac{110}{4} = -89mV$$

$$Ca^{2+} \quad 31 \ \log \frac{2.5}{10^{-4}} = 136mV$$

$$31 \ \log \frac{5}{10^{-4}} = 146mV$$

Ecuación Nerst para cada ion.

$$V = -61 \cdot \log\left(\frac{CNa_i^+ PNa^+ + CK_i^+ PK^+ + CCl_o^- PCl^-}{CNa_o^+ PNa^+ + CK_o^+ PK^+ + CCl_i^- PCl^-}\right)$$

$$+61mV - 94mV = -86mV - 4mv = -90mv$$

Action Potential

Action Potential

Bomba de Na-K

Reposo

Despolarizacion

Repolarizacion

Canales de Fuga

Saltatory Conduction

+Vainas de mielina (oligodendrocitos, células de Schwann)

+Hasta 100 m/s (sin mielina 0.25 m/s)

Synapse

- Canales por ligando (40 neurotransmisores)
- Iónico:
  - Excitatorio (Na^+)
  - Inhibitorio(Cl^-, K^+)
- Activación de Segundo Mensajero:
  - Proteinas G cAMP, cGMP
  - Ccambios enzimáticos, transcripción, activación de canales K^+
- One-Way

Synaptic Weights

Synapse

https://twitter.com/i/status/1502259198603808774

Synaptic Weights

Synapse

Sumation and Propagation

Sumatoria Espacial

Sumatoria Temporal

(+) Gap Junction
(+) Ephatic Coupling
(+) Mechanical

Corriente eléctrica ionica

Segmento inicial del axon (mayor presencia de canales)

Neural Networks

Sumation and Propagation

Hyperpolarized pulse of current

Refractory Period

Rebound Spike

Temporal Summation

Cortical Networks

Barral J & Reyes AD (2016)

Brain Imaging

Medical Imaging

Invasive Scanning

Utah Grid Microarray

Figure 3: The robotic electrode inserter; enlarged view of the inserter-head shown in the inset. **A.** Loaded needle puncher cartridge. **B.** Low-force contact brain position sensor. **C.** Light modules with multiple independent wavelengths. **D.** Needle motor. **E.** One of four cameras focused on the needle during insertion. **F.** Camera with wide angle view of surgical field. **G.** Stereoscopic cameras.

Elon Musks' Neuralink

Deep Brain Stimulation

PMT Corporation, Chanhassen, MN, USA

Non-invasive Scanning

- EEG *Electroencefalográma*
- ECoG *Electrocorticography*
- MEG *Magnetoencefalografía*
- fMRI *Resonancia Magnética Funcional*
- PET *Tomografía por Emisión de Positrones*
- fTCD *Functional TransCranial Doppler*
- fNIR *functional Near InfraRed*

Non-invasive Scanning - EEG

Non-invasive Scanning - fNIR MEG

University of Nottingham, 2018

Drexel University, 2016

Diffusion Tensor Trachtography

The tools for the Connectome

Brain GPS

Moser, 2008

The materialization of Statistical Mathematics

Machine Learning

"Learning" computers

Machine Learning Model

Engineering Inverse Problem

What is “learning”?

Model Free Approach
To determine the mapping between some input X and some output Y by “adjusting” free parameters.

These free parameters are adjusted based on an Optimization Process

What is “learning”?

Optimization Problems with Constraints

(1) A company wants to manufacture cylindrical aluminum cans with a volume of $1000cm^3$ (1 liter). What should the radius and height of the can be to minimize the amount of aluminum used?

Volume of a Cylinder: $V = \pi r^2 h$

Surface Area of a Cylinder: $S = 2\pi r^2 + 2\pi r h$

We wish to optimize (minimize) the surface area $S = 2\pi r^2 + 2\pi r h$ with respect to the constraint $\pi r^2 h = 1000$. From this latter condition we have $h = \frac{1000}{\pi r^2}$, which gives

$$S = 2\pi r^2 + 2\pi r\left(\frac{1000}{\pi r^2}\right) = 2\pi r^2 + \frac{2000}{r}.$$

Differentiating,

$$S' = 4\pi r - \frac{2000}{r^2}.$$

Observe that $r = 0$ is a critical point. We solve the equation $S' = 0$, giving

$$\begin{aligned} S' &= 0 \\ 4\pi r - \frac{2000}{r^2} &= 0 \\ 4\pi r &= \frac{2000}{r^2} \\ 4\pi r^3 &= 2000 \\ r^3 &= \frac{500}{\pi} \end{aligned}$$

$$r = \left(\frac{500}{\pi}\right)^{1/3}.$$

The derivative S' is negative for r in the interval $(0, \left(\frac{500}{\pi}\right)^{1/3})$ and positive for $r > \left(\frac{500}{\pi}\right)^{1/3}$, so we see that $r = \left(\frac{500}{\pi}\right)^{1/3}$ is a minimum. For this value of r the equation $h = \frac{1000}{\pi r^2}$ gives $h = \frac{1000}{\pi\left(\frac{500}{\pi}\right)^{1/3}}$.

Terminology

DL Map

START HERE
Cove of
Actuarial Science
Gulf of
Descriptive Statistics
Data Visualization
Experimentation
Sea of
Probability & Statistics
Business Intelligence
Biostatistics
Operations Research
Causal Inference
Inferential
Hypothesis Testing
Bayesian Inference
Statistical Learning
Econometrics
Financial Quantitative Analysis
Unsupervised Learning
Recurrent NN
Spark
Deep Learning
Convolutional NN
SVM
Hadoop
Kafka
Transformers
Paint
Machine Learning
Big Data Technologies
Feed Forward NN
Reinforcement Learning
Ensembled Decision Trees
NoSQL
Pig
Computer Vision
CircleCI
MLops
Docker
Data Engineering
ETL
DVC
Computer Graphics
Database Admin/Mgmt
Data Structures
SQL
Data architecture
Computer Networks
Bay of
Computer Science
Algorithms
mare nullius
Programming

Machine Learning
Supervised Learning
Regression
Linear Regression
Polynomial Regression
Ridge/Lasso Regression
Classification
Logistic Regression
SVM
Decision Trees
Naive Bayes
K-NN
Unsupervised Learning
Clustering
DBSCAN
K-Means
Agglomerative
Mean Shift
Fuzzy C-Means
Association Rule Learning
FP Growth
Euclat
Apriori
Dimensionality Reduction
t-SNE
PCA
LSA
SVD
LDA
Reinforcement Learning
Q-Learning
DQN
SARSA
Genetic Algorithm
A3C
Ensemble Learning
Stacking
Bagging
Random Forest
Boosting
XGBoost
LightGBM
AdaBoost
CatBoost
Neural Networks and Deep Learning
Convolutional Neural Networks (CNN)
DCNN
Recurrent Neural Networks (RNN)
LSM
GRU
LSTM
Generative Adversarial Networks (GAN)
Autoencoders
seq2seq
Perceptrons

Roles

Data Engineer

Data Scientist

Data Analyst

Evolving Artificial Intelligence

Emerging Technologies

- GANs
- Explanaible AI
- Edge AI
- AI PaaS
- Adaptive ML
- Emotion AI

Learning by Data

In ML there are three approaches to force a machine to “learn”

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

Semi Supervised Learning

Self Supervised Learning

- Classification
- Regression
- Generation
- Optimize

Machine Learning Techniques

SVM

LDA

LogReg

kNN

Autoregressive

Machine Learning Techniques

Watching Dreams?

Nishimoto,Cell 2011, Nytimes

Watching Thoughts?

High-resolution image reconstruction with latent diffusion models from human brain activity, Takagi, Nishimoto 2022

Biomimetic neuroscience

Deep Brain

The Revolution of Deep Learning

Neural Networks - McCulloch and Pitts

Neural Networks - MLP Multilayer Perceptron

$$o = \sigma(p + \sigma(b + \bar{x}\,W)V)$$

$$\bar{o} = \sigma(p + \bar{h}V)$$

$$\bar{h} = \sigma(b + \bar{x}\,W)$$

Deep Learning

- Cybernetics (McCulloch and Pits, Hebb 1949, Rosenblatt 1958)
- Connectionism (Backprop: Rumelhart and Hinton^, Hopfield^, 1986)
- Dechter 1986
- Aizenberg 2000
- Krizhevsky AlexNet 2012

} Origins

- Deep Learning (Hinton, Bengio and LeCun* 2006, Ranzato 2007)
- ChatGPT (OpenAI)

- Minsky and Papert 1969
- Olshausen and Field 2005
- Fukushima Neocognitron 1980
- Distributed Representation 1986
- Beating ImageNet Krizhevsky 2012
- Playing Atari Games, Mnih 2015

(*) Turing Award 2019

(^) Physics Nobel Prize 2024

What Deep Learning is NOT?

“Los algoritmos de IA no se programan.”

“Los robots vienen a reemplazarnos.”

“Deep Learning permite reemplazar a los seres humanos.”

“La tecnología empezó a reemplazarnos, y más nos vale sacarle provecho que quedarnos amargados por la herida que nos provoque."

"Los algoritmos tienen además la ventaja que no se cansan y pueden interactuar con el contexto a velocidades supersónicas en comparación con las personas"

“Nada será igual” M.Tetaz

“Trolley Train Problem.”

“AGI Artificial General Intelligence.”

“Make decisions based on data.”

Laura Accion, et al , “Desmistificando la Inteligencia Artificial” en “Inteligencia artificial, una mirada interdisciplinaria” , 2020

Why the AI is so dumb, IEEE Spectrum Nov 2021

What Deep Learning is NOT?

What Deep Learning is NOT?

- El desarrollo científico no es cosa de jóvenes: el 51% de las personas menores de 30 años no lo ubica entre las urgencias del país. Las generaciones mayores lo jerarquizan más.

La Inteligencia Artificial y el horizonte de lo humano: el 46% cree que la IA puede superar a las personas, mientras que el 52% lo descarta. Miedo, fascinación y futuro incierto.

https://pulsar.uba.ar/creencias-sociales-2024-informe-6-los-argentinos-y-la-ciencia/

What Deep Learning is NOT?

Esto muestra **una sociedad escéptica ante la idea de que la IA supere al ser humano, aunque no de manera tajante. Los más jóvenes (18-29 años) son el único grupo donde predomina la creencia en su superioridad futura** (55% está muy o algo de acuerdo). Esto puede vincularse con una mayor familiaridad tecnológica, exposición a narrativas digitales y una socialización en un mundo donde la IA ya opera cotidianamente. En cambio, a medida que se avanza en edad crece el escepticismo. Entre los mayores de 50 años de edad, la mayoría rechaza esa posibilidad (55% está poco o nada de acuerdo), quizás como defensa simbólica del lugar humano o por menor exposición y comprensión al avance de estas tecnologías.

https://pulsar.uba.ar/creencias-sociales-2024-informe-6-los-argentinos-y-la-ciencia/

What Deep Learning is?

Revolutionary Technology to move ahead the digital boundary

Computer Vision

Natural Language Processing

Speech Recognition

"Artificial Intelligence and Life in 2030", Sept 2016

Digital Boundary

Only that ?

Computer Vision

Natural Language Processing

Blueprints Thinking Machine

Speech Recognition

https://whatisintelligence.antikythera.org/

What Deep Learning is?

Magnetic control of tokamak plasmas through deep reinforcement learning

Jonas Degrave, Federico Felici, ... Martin Riedmiller + Show authors

How a Plucky Robot Found the Long-Lost *Endurance* Shipwreck

Over a century ago, two dozen men were stranded in Antarctica. Here's how a robot dove 10,000 feet to glimpse their lost ship for the first time.

AlphaFold AI predicts SARS-CoV-2 omicron variant might not evade antibody neutralization

AlphaFold: a solution to a 50-year-old grand challenge in biology

Discovering faster matrix multiplication algorithms with reinforcement learning

Alhussein Fawzi, Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes,

Competitive Coder

DeepMind's StarCraft-playing AI beats 99.8 per cent of human gamers

Nature,Wired, The Batch, Newscientist

What Deep Learning is?

A window to our imagination

Where Deep Learning may have something to do?

Where Deep Learning may have something to do?

Agility Robotics

Where Deep Learning may have something to do?

Boston Dynamics

Realistic Applications

きつい・汚い・危険

- Aburrido

- Descuidado

- Peligroso

Every successful deployment has either one of two expedients:

- It has a person somewhere in the loop, or
- The cost of failure, should the system blunder, is very low

Rodney Brooks, “An inconvenient truth About AI”, IEEE Spectrum, Sept 2021

Deep Learning

NEW NAVY DEVICE LEARNS BY DOING

Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser

WASHINGTON, July 7 (UPI) —The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.

The embryo—the Weather Bureau's $2,000,000 "704" computer—learned to differentiate between right and left after fifty attempts in the Navy's demonstration for newsmen.

The service said it would use this principle to build the first of its Perceptron thinking machines that will be able to read and write. It is expected to be finished in about a year at a cost of $100,000.

Dr. Frank Rosenblatt, designer of the Perceptron, conducted the demonstration. He said the machine would be the first device to think as the human brain. As do human beings, Perceptron will make mistakes at first, but will grow wiser as it gains experience, he said.

Dr. Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said Perceptrons might be fired to the planets as mechanical space explorers.

Without Human Controls

The Navy said the perceptron would be the first non-living mechanism "capable of receiving, recognizing and identifying its surroundings without any human training or control."

The "brain" is designed to remember images and information it has perceived itself. Ordinary computers remember only what is fed into them on punch cards or magnetic tape.

Later Perceptrons will be able to recognize people and call out their names and instantly translate speech in one language to speech or writing in another language, it was predicted.

Mr. Rosenblatt said in principle it would be possible to build brains that could reproduce themselves on an assembly line and which would be conscious of their existence.

1958 New York Times…

In today's demonstration, the "704" was fed two cards, one with squares marked on the left side and the other with squares on the right side.

Learns by Doing

In the first fifty trials, the machine made no distinction between them. It then started registering a "Q" for the left squares and "O" for the right squares.

Dr. Rosenblatt said he could explain why the machine learned only in highly technical terms. But he said the computer had undergone a "self-induced change in the wiring diagram."

The first Perceptron will have about 1,000 electronic "association cells" receiving electrical impulses from an eye-like scanning device with 400 photo-cells. The human brain has 10,000,000,000 responsive cells, including 100,000,000 connections with the eyes.

IEEE Spectrum Nov 2021, Why AI is harder than we think?[49]

Deep Learning

1956 Dartmouth AI Conference

Why AI is harder than we think?[49]

The Four “Deep Learning” pillars

- Multiple Level of Composition
- Representational Learning
- Massive and standardized datasets
- GPU

Why deep?

Limitations of Machine Learning on world problems

Schmidhuber 2014

Hochreiter 1991, Hinton 2006

ML Problem - The Curse of Dimensionality

ML Problem - The Curse of Dimensionality

ML Problem - The Curse of Dimensionality

Dessert: High-Dimensional Intuition Failure

- In $\mathbb{R}^2$:
  - 4 unit-radius circles inside $[-2,2]^2$
  - circle of radius r centred at the origin, confined by the unit-radius circles
  - distance from origin to centers: $\sqrt{2}$
  - $r = \sqrt{2} - 1$
- In $\mathbb{R}^d$:
  - 2^d unit-radius spheres inside $[-2,2]^d$
  - radius-r sphere centred at the origin, confined by the unit-radius spheres
  - $r = \sqrt{d} - 1$
  - for $d \geq 10$, we have $r > 2$

- J. Michael Steele, *The Cauchy-Schwarz Master Class*, Cambridge, 2004.

If $d \geq 10$, the sphere confined by the spheres in the box, goes outside the box!

Why deep?

Universal Approximation Theorem

A feedforward neural network with a single hidden layer containing a finite number of neurons can approximate any continuous function on a compact subset of R, given appropriate activation functions.

"Approximation by superpositions of a sigmoidal function", Cybenko 1989

Why deep?

Limitations of ANN on real world problems

Shallow Network → UAT ← **Deep Network**

Difficult to train
Difficult to find the optimal
Difficult to generalize well

Compositional Structure { Shift Invariance / Locality at different scales

"Many natural questions on images corresponds to algorithms that are compositional"

"The unreasonable effectiveness of deep learning in artificial intelligence", Sejinowski 2019
"When and why..", Mhaskar 2017

Why deep?

Limitations of ANN on real world problems

The Curse of Dimensionality

Local Consistency Assumption

Interests points are Sparse

Limitations of Machine Learning on world problems

Deep

Vanishing or Exploding Gradients

Fundamental Deep Learning Problem

Fundamental Credit Assignment Problem

Credit Assignment Path

"When and why..", Mhaskar 2017

Schmidhuber 2014

Hochreiter 1991, Hinton 2006

Add more layers

Overcoming the FDLP

Limitations of ANN on real world problems

Fundamental Deep Learning Problem

- Unsupervised Pretraining
- Greedy Layer-Wise
- LSTM cells
- Hessian Free optimization, Dropout, Xavier initialization
- Random weight guessing

Schmidhuber 2014
Hochreiter 1991, Hinton 2006

Neural Networks - Convergence Paradox

"The unreasonable effectiveness of deep learning in artificial intelligence", Sejinowski 2019

Representational Learning

Representational Learning

In ML there are three approaches to force a machine to "learn"

Supervised Learning

Unsupervised Learning

Reinforcement Learning

Semi Supervised Learning

Self Supervised Learning

Massive Datasets

MNIST handwritten digit dataset

ImageNet Classification Error

Iris 10^2 , ImageNet 5000 images per category, WMT 10^9 total size,

Mammalian Brain Like?

[9] GoogLeNet Szegedy 2014

[18] MultiGPU ConvNet Krizhevsky 2012

GPU

CUDA

OpenCL

Shaders

GPUs

Google Coral Dev Board

Nvidia Jetson Tk1

Intel Movidius Neural Stick

Nvidia Titan V

Nvidia Jetson Nano. RPi para GPU, Hardware Lottery [44]

Money is hardware

ImageNet Training in 24 Minutes

Yang You, Zhao Zhang, James Demmel, Kurt Keutzer, Cho-Jui Hsieh

(Submitted on 14 Sep 2017)

Finishing 90-epoch ImageNet-1k training with ResNet-50 on a NVIDIA M40 GPU takes 14 days. This training requires 10^18 single precision operations in total. On the other hand, the world's current fastest supercomputer can finish 2 * 10^17 single precision operations per second (Dongarra et al 2017). If we can make full use of the supercomputer for DNN training, we should be able to finish the 90-epoch ResNet-50 training in five seconds. However, the current bottleneck for fast DNN training is in the algorithm level. Specifically, the current batch size (e.g. 512) is too small to make efficient use of many processors

For large-scale DNN training, we focus on using large-batch data-parallelism synchronous SGD without losing accuracy in the fixed epochs. The LARS algorithm (You, Gitman, Ginsburg, 2017) enables us to scale the batch size to extremely large case (e.g. 32K). We finish the 100-epoch ImageNet training with AlexNet in 24 minutes, which is the world record. Same as Facebook's result (Goyal et al 2017), we finish the 90-epoch ImageNet training with ResNet-50 in one hour. However, our hardware budget is only 1.2 million USD, which is 3.4 times lower than Facebook's 4.1 million USD.

Software and Cloud Based Tools

- **Amazon**: Machine Learning, DSSTNE, Orbeus
- **Facebook**: TorchLibrary
- **Google**: Cloud Machine Learning, TensorFlow
- **IBM**: Watson Analytics, IBM SystemsML (AlchemyAPI)
- **Microsoft**: Azure Machine Learning, Computational Network Toolkit (SwiftKey)

- OpenAI
- Theano
- PyLearn2
- Torch
- DistBelief
- Caffe
- MXNet
- Tensorflow
- Keras

https://spectrum.ieee.org/computing/software/now-you-too-can-buy-cloudbased-deep-learning

Keras and TensorFlow

Deep Learning Networks

ConvNet - Convolutional Networks

Representational Learning

- Pooling = downsampling
- ReLU: Rectified Linear Neurons

Fukushima 1980, Krizhevsky 2012

RNN - Recurrent Neural Network

$$h_{t+1} = \tanh(W_{xh}x_t + W_{hh}h_t)$$

$$y_{t+1} = g(h_t)$$

Time unfolding

"On the Computational Power of Neural Nets" Siegelmann 1995, Chen 2016

LSTM - Long Short Term Memory

Hochreiter & Schmidhuber 1997, Chen 2016

DBN - Deep Belief Networks

Hopfield Neural Network

$$E = \sum_i W_{ij} XY$$

Hebian Energy Function

Stacked Boltzmann Machine

"Deep Belief Nets" Hinton 1995

Autoencoders

Dimensionality Reduction
PCA

Sparse Autoencoder
(Kohonnen)

Denoising Autoencoder

Contracting Autoencoder

Variational Autoencoders Generators

https://colab.research.google.com/drive/1Hobi6plfrrUQ9DCiFPZ6vaztOu5Q05ND

GAN - Generative Adversarial Networks

Playing a counterfeit game.

G network is trained to increase the chance of D assuming and output of G being valid.

D network is trained to decrease the chance of assuming G output to be valid.

$$\min_G \max_D V(D, G) = E[\log D(x)] + E[\log(1 - D(G(z))]$$

Goodfellow 2014

GAN - Generative Adversarial Networks

Mario Klingermann

GAN - Generative Adversarial Networks

Ming-Yu Liu et al https://arxiv.org/pdf/1905.01723.pdf

Computer Art

COMPUTERWORLD September 11, 1968

Computer Art Is Called Genuine New Art Form

Computer art, drawn by a plotter under the direction of a computer, has become something of a fad this year, and at least one expert thinks such art represents a genuine new art form.

The computer has definite potential as an art medium and should be thoroughly explored, according to Dr. Thomas S. Fern, chairman of the Department of Art, University of Notre Dame.

Fern said the computer meets the first essential condition of art: it can serve as a medium for human expression. If computer plotting were merely a mechanical matter in which creativity did not play a role, then it would not be a valid art form, he explained.

Similar Art Forms

There are other forms of art in which the artist doesn't necessarily create a finished product with his own hands, Fern pointed out. For example, an architect visualizes the form and substance of a building but leaves the construction to someone else. The fact that the architect produces blueprints, or instructions to the builder, rather than the building itself, does not diminish his creative role. Logically then, Fern added, a human who instructs a computer also is creative and can be considered an artist.

In Fern's opinion, computer produced abstract drawings are more exciting than representational work.

"Orderliness and precision are the outstanding qualities of computer drawings," he observed. "I personally think it's more interesting to see such qualities used in an abstract way than for a literal representation of a scene."

https://x.com/quasimondo/status/1706947023037927690?s=20

Transformers

Attention is all you need, Vaswani 2017

Transformers

Attention is all you need, Vaswani 2017

Diffusion Models

"a bowl of soup that is a portal to another dimension as digital art"

CLIP Contrastive Learning

GLIDE Diffusion Model

Transformer Model

Diffusion Models

High-Resolution Image Synthesis With Latent Diffusion Models, Rombach 2022

Stable Diffusion

Latent Stable Diffusion

U-Net Network conditioned to a Condition Embedding. It reverses the diffusion process.

VAE Variational Autoencoder to convert to and from the Latent Space.

Condition Encoder CLIP Generate Condition Embedding for the UNET

Residual Neural Networks and U-Nets

U-Net: Convolutional Networks for Biomedical Image Segmentation, Ronneberger 2015

Cross Attention

U-Net: Convolutional Networks for Biomedical Image Segmentation, Ronneberger 2015

Conditional Encoder

U-Net: Convolutional Networks for Biomedical Image Segmentation, Ronneberger 2015

Reinforcement Learning

Figure 1.2: The RL problem.

Mnih 2015,Lavet 2018,Sutton 2018

Deep Reinforcement Learning

- Start at s:state, a:action
- Q(s,a) based on greediest a
- Iterate
  - Apply a
  - Get s′
  - Get r(s′)
  - $Q(s,a) = \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$

Q: Deep Neural Network

Mnih 2015,Lavet 2018,Sutton 2018

Are our jobs been displaced by robotics?

Plenty of work to be done

Let the robot do the more dangerous and boring stuff

Noam Chomsky https://www.youtube.com/watch?v=ojzNM6tRiyc

Privacidad

HIPAA

GDPR

Deep Fake

Neuroderechos

What are current real risks of DL?

Is there anything unbiased?

Biased

Minority Report

Lost Agency

Cambridge Analytical

Privacy

DeepFake

Hacking Autonomous Car

Digital Boundary

Accountability

https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Unsettled discussion

El pato de Vaucanson

Planes don't flap their wings

Principle of multiple realizability:
"different substrates can nonetheless perform the same input-output mapping but they are not the same".

Grace Lindsay[47], Guest 2021 [52]

Functionalism vs Mechanism

if the clockwork clock behaves like the digital clock, then the clockwork clock is a digital clock

The ability to fit

Collective Wisdom:
principles and mechanism [52]

The ability to summarize

A Theory of Abstraction in RL, David Abel
Del Rigor de la Ciencia, Borges

Connectionism vs Symbolism

Looking back, looking ahead: Symbolism vs connections AI, Goel 2021

The Age of Data Manifolds

Robust.AI on Collaborative Productivity

https://www.robust.ai/

IEEE Spectrum DARPA Challenge

Is DL++ good or bad?

make the world more open and connected

Don't be evil

Internet will provide to everyone, access
To humanity's knowledge

Is DL++ good or bad?

Technology
Is
Inherently
Good

Allows more DOF

Emerging Technologies

- GANs
- Explanaible AI
- Edge AI
- AI PaaS
- Adaptive ML
- Emotion AI

Emerging Technologies

- Autonomic Systems
- Generative Design AI
- Causal IA
- Metaverse
- Web3/Decentralized Identity
- Data Observability

Applied AI

DS

Liveness Detection

Estado de la fruta
Seguros de Exportaciones

Retail

Customer Satisfaction
NLP

OCR de los Telegrama Electorales

AI ACTION SUMMIT

INSTITUT POLYTECHNIQUE DE PARIS

The Individual versus the Collective

- And then came LLMs…
- Suddenly those large, distributed collections of data and functions and gradients seemed to be acting human-like and have a "mind"—the phrase "AI" was retrieved—and the focus again was on an intelligent agent, for no particularly good reason

Fallacies of Data Science

- Biased Cherry Picking: Selecting results that fit your claim.
- Data Dredging: Results of correlation was due to chance.
- Survivorship bias: Drawing conclusions from an incomplete set of data that survived.
- Perverse incentive: the incentive produces the opposite result (wrong metrics, KPIs).
- False causality: two things occur together, they are related (falsely)
- Garbage in, garbage out.
- Sampling bias (cross validation, bootstrapping)
- Gamblers fallacy: believing that when something happen the chance of happening again is reduced.
- Unbalanced Datasets: medical images
- Simpsons paradox: a trend is present in a group but disappears when data is combined.
- Clever Hans Effect: https://arxiv.org/pdf/2006.10609.pdf

Fallacies of Data Science

It also muddies the origin of certain data sets. This can mean that researchers miss important features that skew the training of their models. Many unwittingly used a data set that contained chest scans of children who did not have covid as their examples of what non-covid cases looked like. But as a result, the AIs learned to identify kids, not covid.

Driggs's group trained its own model using a data set that contained a mix of scans taken when patients were lying down and standing up. Because patients scanned while lying down were more likely to be seriously ill, the AI learned wrongly to predict serious covid risk from a person's position.

In yet other cases, some AIs were found to be picking up on the text font that certain hospitals used to label the scans. As a result, fonts from hospitals with more serious caseloads became predictors of covid risk.

Machine Learning

https://twitter.com/vardi/status/1566070939586027520?s=20&t=MHwWjzTDfRfm2d4emgIImg

Artificial Intelligence

https://twitter.com/ExcelHumor/status/1589098712177217538?s=20&t=w97fO1jJWtTOw8R5kLwJmQ

The age of Data Manifolds

Manifold Learning: Model Reduction in Engineering, Ryckelynck

2025

2050

2075

En resumen

Sutton

http://www.incompleteideas.net/IncIdeas/BitterLesson.html

Rodney Brooks

https://rodneybrooks.com/a-better-lesson/

The bitter lesson

En resumen

IEEE Spectrum March 2020

En resumen

- Focalizado para representar datos que intrínsecamente están estructurados jerárquicamente.
- Las capas intermedias ofrecen contenido que tienen valor semántico (es como una salida intermedia parcial de la red). E.g.: la primer capa me separa arriba y abajo y la segunda derecha e izquierda.
- El "deep" implica que la cantidad de saltos en la dependencia causal es mayor.
- Estos layers adicionales permiten que los features los genere y encuentre la misma red (en vez de ponérselos a mano). La red hace la reducción de la dimensionalidad.
- Algunos ajustes a los algoritmos de arquitectura y aprendizaje que aprovechan estas arquitecturas y el paralelismo.
- No estaríamos hablando de esto sin GPUs y sin Moore´s Law.
- Extraordinarios resultados en tres áreas particulares: image processing, speech recognition y natural language processing.
- Excelentes para implementaciones Cheery-Picking en escenarios donde no existe actualmente Digitalización.
- El componente estocástico que tienen las redes hace muy complejo su parametrización.
- Limitaciones cuando la cantidad de muestras es baja[26].
- Inteligible Property[25]
- Hyperhype: Through of Disillusionment (Gartner Hype Cycle)

どうもありがとう

Gracias !

Centro Inteligencia Computacional
ITBA

Follow up

- Tutorials at Conferences (e.g. ICML, NeurIPS)
- Two Minute Papers - Youtube Channel
- The Batch - Andrew NG newsletter
- Yannic Kilcher - Youtube Channel
- Arxiv Insight - Youtube Channel
- http://monostuff.logdown.com/posts/7835544-aprendizaje-modo-mquina

[1] Alberts, Bruce, et al. *Essential cell biology*. Garland Science, 2013.

[2] Guyton and Hall, “Textbook of Medical Physiology”, 12th edition

[3] Izhikevich, “Dynamical systems in neuroscience··, 1st edition

[4] Abott & Dayan, Theoretical Neuroscience, 2012

[5] Ramele, R., A. J. Villar, and J. M. Santos. "A Brain Computer Interface Classification Method Based on Signal Plots." 4th Winter Conference on Brain Computer Interfaces, Yongpyong, Korea, February 2015. IEEE Signal and Processing, 2016.

[6] Chia-Chu et al, “Slow periodic activity in the longitudinal hippocampal slice can self-propagate non-synaptically by a mechanism consistent with ephaptic coupling”, The Physiological Society Journal, 2018

[7] Belousov et al, “Neuronal gap junctions: making and breaking connections during development and injury”, Trends in Neuroscience, 2012

[8] Khraiche, “Ultrasound induced increase in excitability of single neurons”, 2008

[9] Szegedy, “Going deeper with convolutions”, 2014 https://arxiv.org/abs/1409.4842

[10] Montufar, “On the Number of Linear Regions of Deep Neural Networks”, 2014

[11] Graves, “Neural Turing Machines”, 2014

[12] Schmidhuber, “Deep learning in neural networks: An overview”, 2015

[13] Goodfellow, “Deep Learning”, 2017

[14] Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT press, 2018.

[15] http://www.dfki.de/~klusch/DeepLearning-seminar-ss17/Topics/topics.html

[16] Domingos, A Few Useful Things to Know about Machine Learning, 2015

[17] http://deeplearning.net/

[18] Krizhevsky, “ImageNet Classification with Deep Convolutional Neural Networks”, 2012

[19] https://www.acm.org/media-center/2019/march/turing-award-2018

[20] https://www.jeremyjordan.me/autoencoders/

[21] MNIST Dataset http://yann.lecun.com/exdb/mnist/

[22] IRIS Dataset https://archive.ics.uci.edu/ml/datasets/Iris

[23] AUTO MPG Dataset https://archive.ics.uci.edu/ml/datasets/Auto+MPG

[24] Kaggle Datasets

[25] M Bragg, “The Challenge of Crafting Intelligible Intelligence”, ACM Symposium on User Interface Software and Technology, 2018.

[26] F. Lotte et al, “A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update”, Journal of Neural Engineering 15 (2018), no. 3, 031005.

[27] https://medium.freecodecamp.org/how-to-build-your-first-neural-network-to-predict-house-prices-with-keras-f8db83049159

[28] Goodfellow, “Generative Adversarial Nets”, 2014

[29] Coral Google Dev Board https://coral.withgoogle.com/products/dev-board/

[30] Chollet, Deep Learning with Python, 2018

[31] Jetson TK1 Dev https://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html

[32] Langr, GANs in action, Manning 2019

[33] Reitz, The Hitchhiker's Guide to Python!, 2016

[34] A Few Useful Things to Know about Machine Learning, Pedro Domingos, ACM 2010

[35] Data Mining and Analysis, Mohammed J.Zaki, 2010

[36] An Introduction to Multivariate Analysis, Everit ,2012

[37] https://adventuresinmachinelearning.com/keras-lstm-tutorial/

[38] https://www.thestar.com/news/world/2015/04/17/how-a-toronto-professors-research-revolutionized-artificial-intelligence.html

[39] Ben Simon, The dark side of the alpha rhythm: FMRI evidence for induced alpha modulation during complete darkness, 2012

[40] https://medium.com/@jonathan_hui/gan-why-it-is-so-hard-to-train-generative-advisory-networks-819a86b3750b

[41] https://www.tensorflow.org/swift

[42] Gaussia Process https://yugeten.github.io/posts/2019/09/GP/

[43] Hopfield J.J, Computing with neural circuits: a model, 1986

[40] Anthony Zador, A critique of pure learning and what artificial neural networks can learn from animal brains, Nature Communications 2019

[41] https://thegradient.pub/the-limitations-of-visual-deep-learning-and-how-we-might-fix-them/

[42] Chollet, On the Measure of Intelligence

[43] Cybenko, G. (1989) Universal Approximation Theorem

[44] Sara Hooker, The Hardware Lottery, 2020, https://arxiv.org/abs/2009.06489

[45] https://blog.acolyer.org/2016/04/18/deep-learning-in-neural-networks-an-overview

[46] Thompson, R. F. & Kim, J. J. Memory systems in the brain and localization of a memory. Proc. Natl. Acad. Sci. U. S. A. 93, 13438–44 (1996).

[47] Grace Lindsay, https://aeon.co/ideas/planes-dont-flap-their-wings-does-ai-work-like-a-brain

[48] Jordan Guerguiev, Towards deep learning with segregated dendrites, 2017

[49] Melanie Mitchell, Why AI is harder than we think?, 2021, https://arxiv.org/abs/2104.12871

[50] Hinton, “A Fast Learning Algorithm for Deep Belief Nets”, Neural Computation, 2006

[51] https://medium.com/@shridhar743/a-beginners-guide-to-deep-learning-5ee814cf7706

[52] Guest, “On logical inference over brains, behaviour, and artificial neural networks”, 2021

[53] http://myselph.de/hodgkinHuxley.html

[54] https://twitter.com/slava__bobrov/status/1454338973606850563?s=20

[55] https://www.assemblyai.com/blog/diffusion-models-for-machine-learning-introduction

[56] https://www.assemblyai.com/blog/how-dall-e-2-actually-works/

[57] AlphaZero for Matriz Multiplication https://www.nature.com/articles/s41586-022-05172-4

https://www.youtube.com/watch?v=AQrMWH8aC0Q&t=4s

[58] Berghel, Hal. "Malice domestic: The Cambridge analytica dystopia." Computer 51.05 (2018): 84-89.

[59] Sejinowski, The unreasonable effectiveness of deep learning in artificial intelligence (2019)

[60] https://slate.com/technology/2014/05/phineas-gage-neuroscience-case-true-story-of-famous-frontal-lobe-patient-is-better-than-textbook-accounts.html