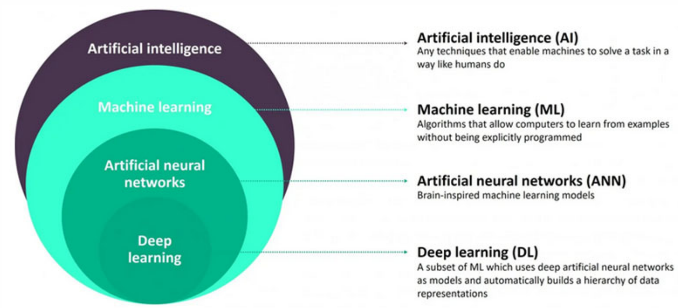
COMP9444: Neural Networks and Deep Learning

Week 1c. Neuroanatomy

Raymond Louie
School of Computer Science and Engineering

Feb 19, 2025

Definitions of AI, machine learning and neural networks



Source: HPE

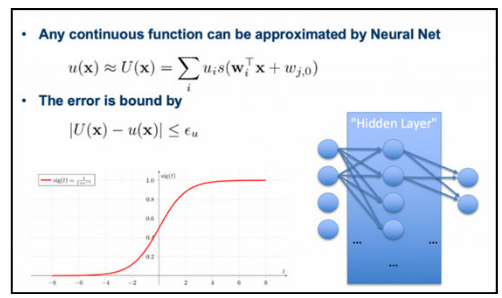
What is a Neural Network?

- massively parallel distributed processor made up of simple processing units
- knowledge acquired from environment through a learning process
- knowledge stored in the form of synaptic weights

Why Neural Networks?

- biologically inspired
- good learning properties
- continuous, nonlinear
- well adapted to certain tasks
- fault tolerant
- graceful degradation
- universal approximation

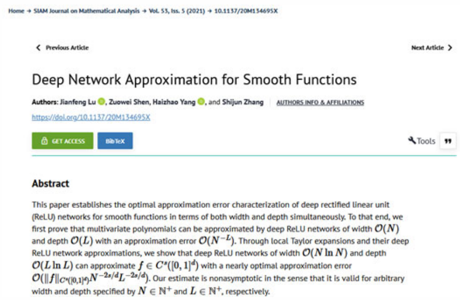
Universal approximation



We can find a one hidden layer representation that approximates any continuous function u(**x**) with an approximation U(**x**) and it is supposed to be very close

https://lme.tf.fau.de/lecture-notes/lecture-notes-dl/lecture-notes-in-deep-learning-known-operator-learning-part-2/

Universal approximation


Home → SIAM Journal on Mathematical Analysis → Vol. 53, Iss. 5 (2021) → 10.1137/20M134695X

Previous Article | Next Article

Deep Network Approximation for Smooth Functions

Authors: Jianfeng Lu, Zuowei Shen, Haizhao Yang, and Shijun Zhang | AUTHORS INFO & AFFILIATIONS

https://doi.org/10.1137/20M134695X

GET ACCESS | BibTeX | Tools

Abstract

This paper establishes the optimal approximation error characterization of deep rectified linear unit (ReLU) networks for smooth functions in terms of both width and depth simultaneously. To that end, we first prove that multivariate polynomials can be approximated by deep ReLU networks of width $\mathcal{O}(N)$ and depth $\mathcal{O}(L)$ with an approximation error $\mathcal{O}(N^{-L})$. Through local Taylor expansions and their deep ReLU network approximations, we show that deep ReLU networks of width $\mathcal{O}(N \ln N)$ and depth $\mathcal{O}(L \ln L)$ can approximate $f \in C^s([0,1]^d)$ with a nearly optimal approximation error $\mathcal{O}(\|f\|_{C^s([0,1]^d)} N^{-2s/d} L^{-2s/d})$. Our estimate is nonasymptotic in the sense that it is valid for arbitrary width and depth specified by $N \in \mathbb{N}^+$ and $L \in \mathbb{N}^+$, respectively.

Depth: number of hidden layers
Width: number of neurons in each layer

Outline

1. **History of AI and neural networks**
2. Biological inspiration of neural networks
3. Recent examples of neural networks

Question

How do we build an intelligent machine which can think like humans?

1. History of intelligence theory (origin: 380BC)
2. History of machines (origin: 1642)
3. History of "symbolic" intelligent machines (origin: 1956)
4. History of "neural networks" intelligent machines (origin: 1943)

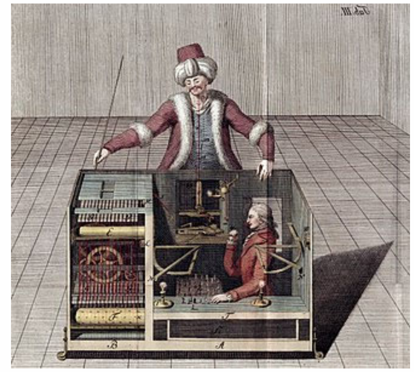
Theories about Intelligence

Intelligence theory	Artificial Intelligence
380BC Plato (Rationalism - innateness)	→ Not innate, but rationale
330BC Aristotle (Empricism - experience)	→ Training on datasets
1641 Descartes (mind-body Dualism)	→ Intelligence = consciousness?
1781 Kant (Critique of Pure Reason)	→ Combine both rationalism and empiricism
1899 Sigmund Freud (Psychology)	→ ? (unconscious processes, childhood experiences)
1953 B.F. Skinner (Behaviourism)	→ Reinforcement learning

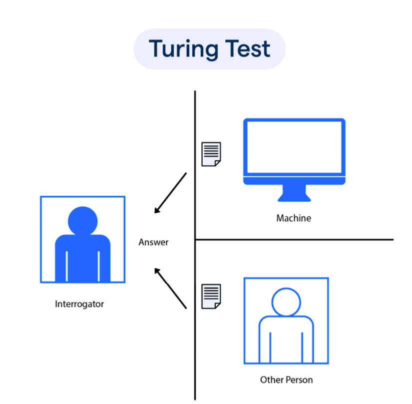
Machine and Artificial Intelligence Origins

- 1642 Blaise Pascal (mechanical adding machine) → Adding/subtracting
- 1694 Gottfried Leibniz (mechanical calculator) → Multiplication/division
- 1769 Wolfgang von Kempelen (Mechanical Turk) → Chess playing machine (hoax)!
- 1837 Babbage & Lovelace (Difference Engine) → Polynomial functions
- 1848 George Boole (the Calculus of Logic) → Logic theory
- 1879 Gottlob Frege (Predicate Logic) → Logic theory
- 1950 Turing Test (imitation game) → Machine/human distinguishability
- 1956 Dartmouth conference → Birth of AI

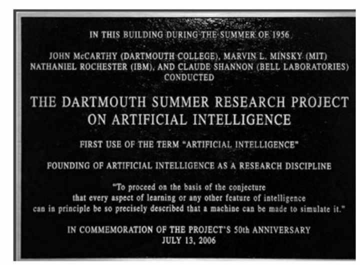
Mechanical Turk



Turing Test



Dartmouth Conference



Two categories of Artificial Intelligence

Serial Symbolic (rule-based AI, classic AI, good old-fashioned AI (GOFA))

- Using a set of rules based on prior knowledge to achieve a task
- Hand-coded
- Example: Following a cooking recipe

Connectionist (neural network)

- Learning associations from data with little or no prior knowledge (not hand-coded)
- Example: Recognizing a friend in a crowd

Serial Symbolic AI

- 1956 Newell & Simon (Logic Theorist) → AI system which prove theorems
- 1959 John McCarthy (Lisp) → Programming language
- 1959 Arthur Samuel (Checkers) → AI checkers
- 1965 Joseph Weizenbaum (ELIZA) → Conversation simulation
- 1974 Edward Shortliffe (Mycin) → Medical diagnosis

Problems with serial-symbolic

- 1970's and early 1980's, AI research focused on symbolic processing, Expert Systems
- Some commercial success, but ran into difficulties:
  - → combinatorial explosion in search spaces
  - → difficulty of formalising everyday knowledge as well as expert knowledge

Neural Network Origins

- 1943 McCulloch & Pitts (neuron models) → First model -> implement logic
- 1948 Norbert Wiener (Cybernetics) → Feedback and adaptive systems
- 1948 Alan Turing (B-Type Networks) → Adaptive Neural network with weights
- 1955 Oliver Selfridge (Pattern Recognition) → How visual images are processed?
- 1962 Hubel and Wiesel (visual cortex) → How neurons respond to visual stimuli?
- 1962 Frank Rosenblatt (Perceptron) → Binary classification tasks

Neural Network "Dark Ages"

- In 1969, Minsky and Papert published a book highlighting the limitations of Perceptrons, and lobbied various funding agencies to redirect funding away from neural network research, preferring instead logic-based methods such as expert systems.
- It was known as far back as the 1960's that any given logical function could be implemented in a 2-layer neural network with step function activations. But, the the question of how to learn the weights of a multi-layer neural network based on training examples remained an open problem. The solution, which we describe in the next section, was found in 1976 by Paul Werbos, but did not become widely known until it was rediscovered in 1986 by Rumelhart, Hinton and Williams.
- from 1969 to 1985, very little work in neural networks or machine learning.
  - a few exceptions, e.g. Stephen Grossberg, Teuvo Kohonen, Paul Werbos

Neural Network Renaissance

- 1986 Rumelhart, Hinton & Williams (multi-layer, backprop) → Algorithm to efficiently train neural networks
- 1989 Dean Pomerleau (ALVINN) → Self driving car
- late 1980's renewed enthusiasm, hype
- 1990's more principled approaches → Rigorous statistical methods
- 2010's deep learning networks, GPU's → Image recognition, NLP and gaming
- 2020's transformers, stable diffusion → High quality image generation and text-to-image generation

Applications of Deep Learning

- Image Processing
  - → Classification
  - → Segmentation
  - → Image Generation
- Language Processing
  - → Translation
  - → Sentiment Analysis
  - → Text Generation
  - → Chatbots
- Game Playing
  - → Deep Q-Learning
  - → AlphaGo
  - → AlphaStar (StarCraft II)
- Combining Images and Language
  - → Automatic Captioning
  - → Generating Images from Text
  - → Video Generation
  - → Multimodal AI

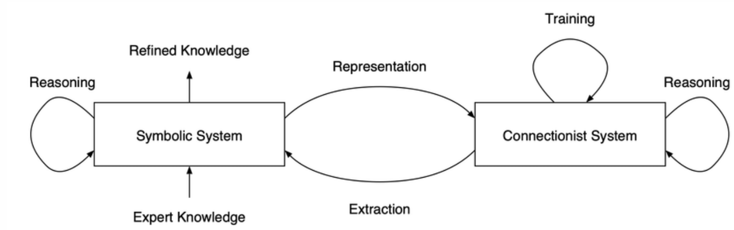
Reading material: History of Deep Learning

Two perspectives on the history of Deep Learning

Viewpoint 1: Focusing on recent work (after 2012)
https://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf

Viewpoint 2: Focusing on earlier work (before 2012)
http://people.idsia.ch/~juergen/deep-learning-overview.html

Neuro symbolic AI



Outline

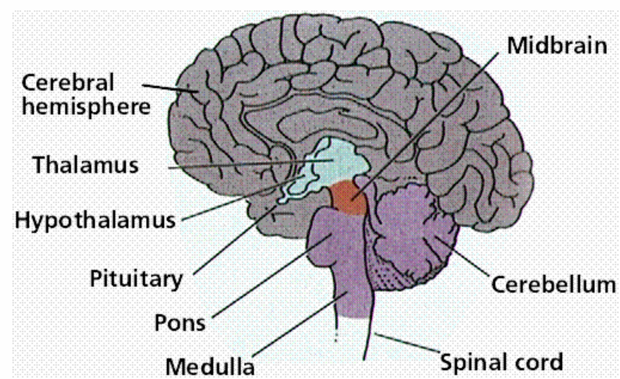
1. History of AI and neural networks
2. **Biological inspiration of neural networks**
3. Recent examples of neural networks

Biological motivation of neural networks

Neural networks are designed as simplified models of how the human brain processes information

1. How does the brain process information?
2. **What is the function and structure of the brain?**

Brain Regions



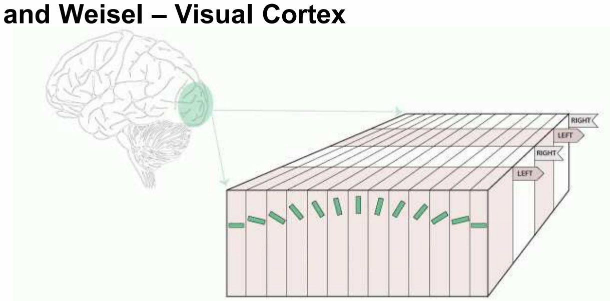
Brain Regions

Brain region	Function	Example AI application
Cerebral Hemisphere	Higher-order thinking, reasoning	Decision-making systems, natural language processing
Thalamus	Sensory relay and integration	Multimodal data integration in robotics
Hypothalamus	Regulating emotions, hunger, and body temperature	Emotion-aware AI, smart home climate control
Pituitary	Hormonal regulation	Health monitoring systems (e.g., blood glucose tracking)
Midbrain	Vision, eye movement, and body movement	Image recognition, autonomous vehicles
Cerebellum	Balance, posture, and movement	Robotics, drone stabilization
Medulla	Controls autonomic functions like heart rate	Wearable health devices

Key messages

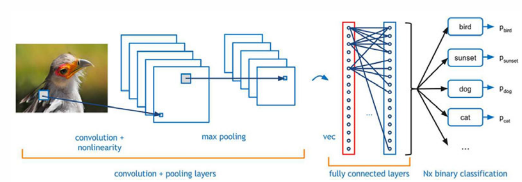
Biology	Artificial
Brain has different functions	Neural networks can implement different functions
Each function in different location	Different neural networks for each function

Hubel and Weisel – Visual Cortex



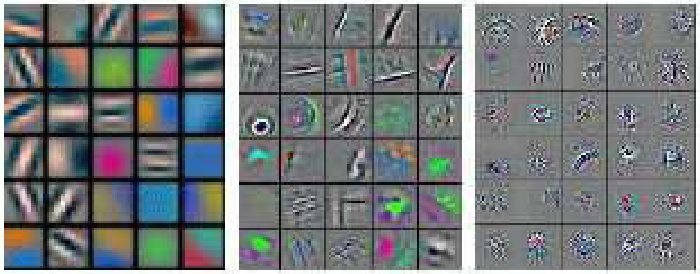
- Cells in V1 (primary visual cortex) respond to lines at different angles
- Cells in V2 (secondary cortex) respond to more sophisticated visual features
- Convolutional Neural Networks are inspired by this neuroanatomy
- CNN's can now be simulated with massive parallelism, using GPU's

Convolutional Networks



- Suppose we want to classify an image as a bird, sunset, dog, cat, etc.
- If we can identify features such as feather, eye, or beak which provide useful information in one part of the image, then those features are likely to also be relevant in another part of the image.
- We can exploit this regularity by using a convolution layer which applies the same weights to different parts of the image.

Convolutional Filters



First Layer Second Layer Third Layer

How far to take biological inspiration?

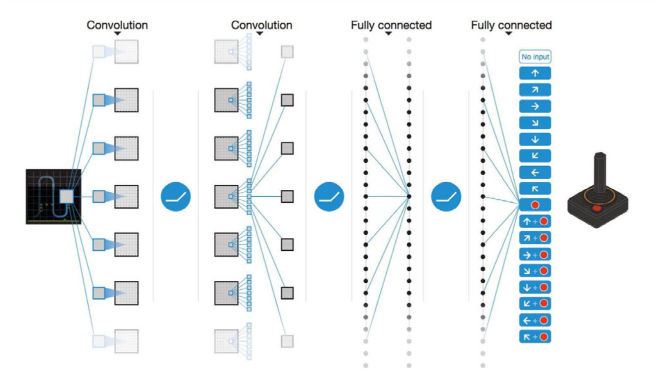
Outline

1. History of AI and neural networks
2. Biological inspiration of neural networks
3. **Recent examples of neural networks**

Large language models

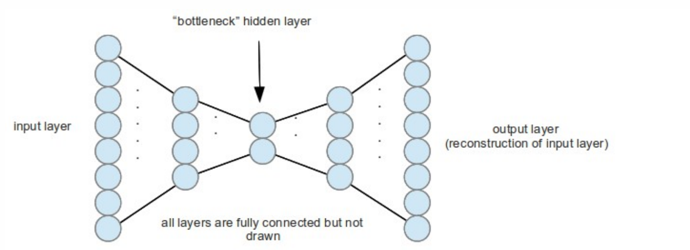
Model	API Pirce / 1M Tokens	
	Input $	Output $
DeepSeek-Coder-V2	0.14	0.28
DeepSeek-V2	0.14	0.28
GPT-4-Turbo-1106	10.00	30.00
GPT-4-0613	30.00	60.00
GPT-3.5	1.50	2.00
Gemini 1.5 Pro	7.00	21.00
Claude 3 Opus	15.00	75.00
Claude 3 Sonnet	3.00	15.00
Claude 3 Haiku	0.25	1.25
abab-6.5（MiniMax）	4.14	4.14
abab-6.5s（MiniMax）	1.38	1.38
ERNIE-4.0（文心一言）	16.56	16.56
GLM-4（智谱清言）	13.80	13.80
Moonshot-v1（月之暗面）	3.32	3.32
Qwen1.5 72B（通义千问）	2.76	2.76
LLaMA 3 70B	3.78	11.34
Mixtral 8x22B	2.00	6.00

Deep Q-Learning



Starcraft

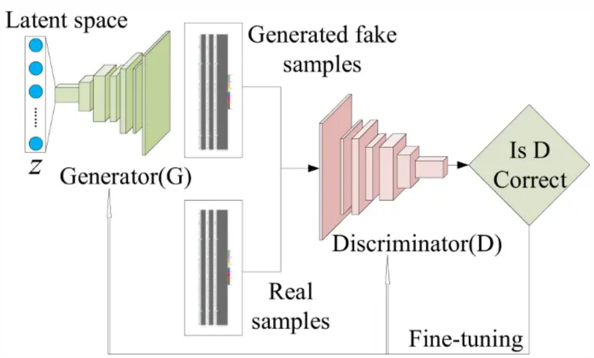
Autoencoder Networks



- output is trained to reproduce the input as closely as possible
- activations normally pass through a bottleneck, so the network is forced to compress the data in some way
- Application: image denoising, image compression, image

Generative Adversarial Networks (image generation)

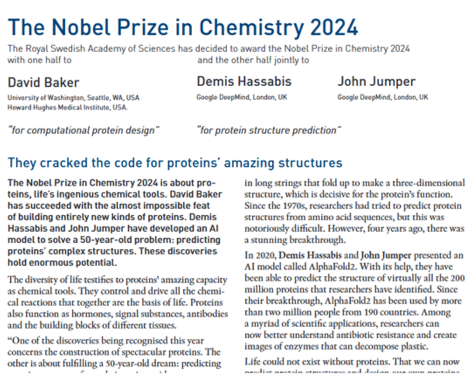
Generative Adversarial Networks (image generation)



Stable Diffusion

- add and remove noise sequentially to create final picture

Alphafold2



The Nobel Prize in Chemistry 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Chemistry 2024 with one half to

David Baker
University of Washington, Seattle, WA, USA
Howard Hughes Medical Institute, USA.

"for computational protein design"

and the other half jointly to

Demis Hassabis
Google DeepMind, London, UK

John Jumper
Google DeepMind, London, UK

"for protein structure prediction"

They cracked the code for proteins' amazing structures

The Nobel Prize in Chemistry 2024 is about proteins, life's ingenious chemical tools. David Baker has succeeded with the almost impossible feat of building entirely new kinds of proteins. Demis Hassabis and John Jumper have developed an AI model to solve a 50-year-old problem: predicting proteins' complex structures. These discoveries hold enormous potential.

The diversity of life testifies to proteins' amazing capacity as chemical tools. They control and drive all the chemical reactions that together are the basis of life. Proteins also function as hormones, signal substances, antibodies and the building blocks of different tissues.

"One of the discoveries being recognised this year concerns the construction of spectacular proteins. The other is about fulfilling a 50-year-old dream: predicting

in long strings that fold up to make a three-dimensional structure, which is decisive for the protein's function. Since the 1970s, researchers had tried to predict protein structures from amino acid sequences, but this was notoriously difficult. However, four years ago, there was a stunning breakthrough.

In 2020, **Demis Hassabis** and **John Jumper** presented an AI model called AlphaFold2. With its help, they have been able to predict the structure of virtually all the 200 million proteins that researchers have identified. Since their breakthrough, AlphaFold2 has been used by more than two million people from 190 countries. Among a myriad of scientific applications, researchers can now better understand antibiotic resistance and create images of enzymes that can decompose plastic.

Life could not exist without proteins. That we can now

Healthcare

Outline

1. History of AI and neural networks
2. Biological inspiration of neural networks
3. Recent examples of neural networks