Artificial Intelligence
Advanced Topics in AI & ML
Deep Learning and Neural Nets. History of AI

Aleksandr Petiushko

ML Research

SOFIA UNIVERSITY

AP

Content

Content

Content

What is Artificial Intelligence?

Natural Intelligence (human)

- Able to perceive the information, analyze it, make decisions based on this analysis

What is Artificial Intelligence?

Natural Intelligence (human)

- Able to perceive the information, analyze it, make decisions based on this analysis

Artificial Intelligence

- (Strong) The same as natural intelligence, but computer is instead of human

What is Artificial Intelligence?

Natural Intelligence (human)

- Able to perceive the information, analyze it, make decisions based on this analysis

Artificial Intelligence

- (Strong) The same as natural intelligence, but computer is instead of human
- (**Weak**) Algorithm which is able to train using the input data in order to do tasks afterward — instead of human

What is Machine Learning

In 1959 Arthur Samuel introduced the term “machine learning” into scientific use.

General definition

Machine Learning — the process leading computers to gain ability to show the behavior that wasn’t explicitly programmed.

What is Machine Learning

In 1959 Arthur Samuel introduced the term “machine learning” into scientific use.

General definition

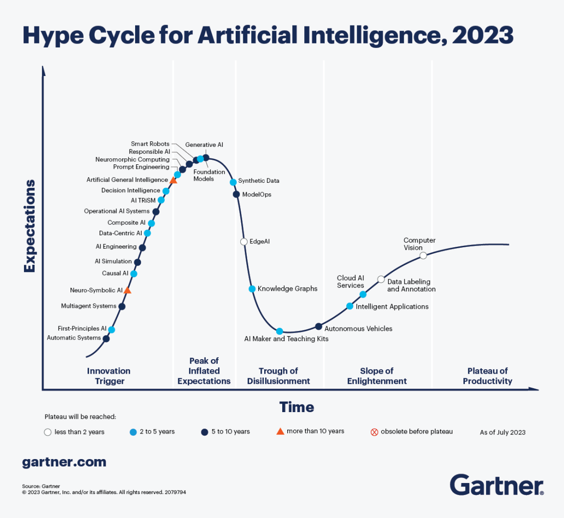
Machine Learning — the process leading computers to gain ability to show the behavior that wasn’t explicitly programmed.

In 1997 Tom M. Mitchell introduced more formal definition of a machine learning algorithm.

Formal definition

A **computer program** is said **to learn** from examples E for some set of problems T and a quality metric P if its performance on problems from T, as measured by P, is improved by using examples E.

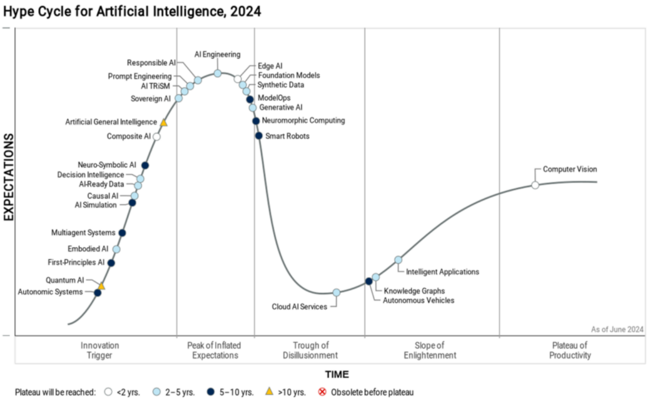
Broad concepts: AI Hype Cycle 2023[1]



[1] www.gartner.com

Broad concepts: AI Hype Cycle 2024[2]

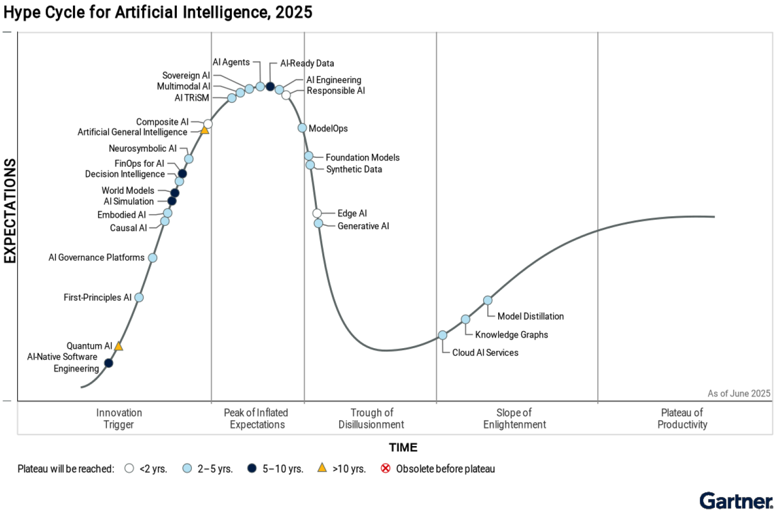
Figure 1: Hype Cycle for Artificial Intelligence, 2024



Gartner

[2] www.jaggaer.com

Broad concepts: AI Hype Cycle 2025[3]



[3] www.mrak.at

Course content

- Deep Learning and Neural Nets
- Generative AI: Generative Adversarial Networks
- Generative AI: Diffusion
- Transformers: encoders and decoders
- LLMs: BERT, GPT
- Applications: Computer Vision
- Applications: Speech Recognition
- Multi-tasking
- Multi-modality
- Interpretability and Explainability
- Embodied AI: Self-Driving
- AI Ethics
- Robust ML

Deep Learning[4] and Neural Nets

- Neural Net (NN): a (usually!) non-linear function mapping a (usually) multi-dimensional input to some output (which can be of the same dimension, or a bigger/smaller one)

[4] Deep Learning Classical Book

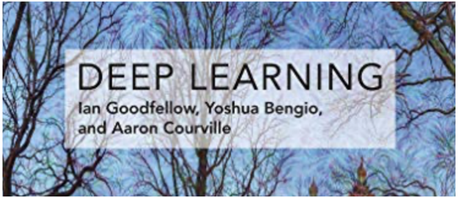
Deep Learning[4] and Neural Nets

- Neural Net (NN): a (usually!) non-linear function mapping a (usually) multi-dimensional input to some output (which can be of the same dimension, or a bigger/smaller one)
- Most common NN atomic operations: addition, multiplication, scalar non-linearity, aggregation/normalization

[4] Deep Learning Classical Book

Deep Learning[4] and Neural Nets

- Neural Net (NN): a (usually!) non-linear function mapping a (usually) multi-dimensional input to some output (which can be of the same dimension, or a bigger/smaller one)
- Most common NN atomic operations: addition, multiplication, scalar non-linearity, aggregation/normalization
- Deep Learning: a NN consisting of more 2 layers of atomic operations (that's why deep) and the corresponding procedure of the training ("learning") it weights using back propagation process



[4] Deep Learning Classical Book

Feature engineering

- Initially the Machine Learning machinery was targeted at **feature engineering** (or extraction, or discovery)

Feature engineering

- Initially the Machine Learning machinery was targeted at **feature engineering** (or extraction, or discovery)
  - Mapping f from arbitrary object $o \in O$ to its unified feature space representation $x \in X \subseteq \mathbb{R}^D$: $f : O \to X$

Feature engineering

- Initially the Machine Learning machinery was targeted at **feature engineering** (or extraction, or discovery)
  - Mapping f from arbitrary object $o \in O$ to its unified feature space representation $x \in X \subseteq \mathbb{R}^D$: $f : O \to X$
- Suppose we have m objects $\Rightarrow$ can construct training dataset X^m

Feature engineering

- Initially the Machine Learning machinery was targeted at **feature engineering** (or extraction, or discovery)
  - Mapping f from arbitrary object $o \in O$ to its unified feature space representation $x \in X \subseteq \mathbb{R}^D$: $f: O \to X$
- Suppose we have m objects $\Rightarrow$ can construct training dataset X^m
- On top of that representation (usually a *linear*) a classifier is learned (e.g., SVM): $a(w, X^m) = \operatorname{sign} g(x, w) = \operatorname{sign}\langle x, w\rangle$

Representation Learning

- Main goal: to have objects representations that are rich and descriptive:
  - Representations of objects from the same class are better *to be close* (according to some metric, e.g., Euclidean) in the feature space

Representation Learning

- Main goal: to have objects representations that are rich and descriptive:
  - Representations of objects from the same class are better *to be close* (according to some metric, e.g., Euclidean) in the feature space
  - Representation of objects from different classes are better to *lie far* from each other (according to some metric)

Representation Learning

- Main goal: to have objects representations that are rich and descriptive:
  - Representations of objects from the same class are better *to be close* (according to some metric, e.g., Euclidean) in the feature space
  - Representation of objects from different classes are better to *lie far* from each other (according to some metric)
- Sometimes the initial features $x \in X$ are not suited enough for this task

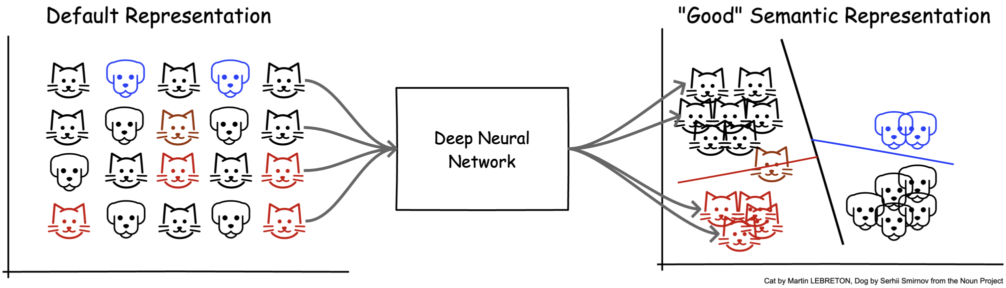
Representation Learning

- Main goal: to have objects representations that are rich and descriptive:
  - Representations of objects from the same class are better *to be close* (according to some metric, e.g., Euclidean) in the feature space
  - Representation of objects from different classes are better to *lie far* from each other (according to some metric)
- Sometimes the initial features $x \in X$ are not suited enough for this task
- For this purpose, additional mapping h from X into some additional representation space $R \subseteq \mathbb{R}^N$: $h : X \to R$

Representation Learning

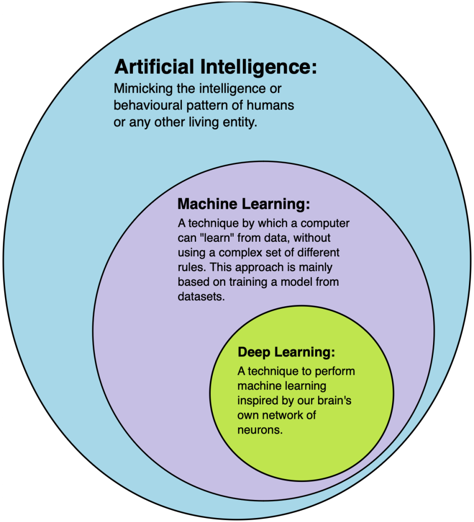
- Main goal: to have objects representations that are rich and descriptive:
  - Representations of objects from the same class are better *to be close* (according to some metric, e.g., Euclidean) in the feature space
  - Representation of objects from different classes are better to *lie far* from each other (according to some metric)
- Sometimes the initial features $x \in X$ are not suited enough for this task
- For this purpose, additional mapping h from X into some additional representation space $R \subseteq \mathbb{R}^N$: $h : X \to R$
- The learning of this additional mapping h is called **Representation Learning**

Representation spaces illustration[5]



[5] Image credit: blog.fastforwardlabs.com

AI vs ML vs DL[6]



[6] Wiki

Deep Learning History I[7]

1943	McCulloch & Pitts	introduced MCP Model, which is considered as the ancestor of Artificial Neural Model.
1949	Donald Hebb	considered as the father of neural networks, introduced Hebbian Learning Rule, which lays the foundation of modern neural network.
1958	Frank Rosenblatt	introduced the first perceptron, which highly resembles modern perceptron.
1974	Paul Werbos	introduced Backpropagation
1980	Teuvo Kohonen	introduced Self Organizing Map
	Kunihiko Fukushima	introduced Neocogitron, which inspired Convolutional Neural Network
1982	John Hopfield	introduced Hopfield Network
1985	Hilton & Sejnowski	introduced Boltzmann Machine
1986	Paul Smolensky	introduced Harmonium, which is later known as Restricted Boltzmann Machine
	Michael I. Jordan	defined and introduced Recurrent Neural Network

[7] On the Origin of Deep Learning

Deep Learning History II[8]

1990	Yann LeCun	introduced LeNet, showed the possibility of deep neural networks in practice
1997	Schuster & Paliwal	introduced Bidirectional Recurrent Neural Network
	Hochreiter & Schmidhuber	introduced LSTM, solved the problem of vanishing gradient in recurrent neural networks
2006	Geoffrey Hinton	introduced Deep Belief Networks, also introduced layer-wise pretraining technique, opened current deep learning era.
2009	Salakhutdinov & Hinton	introduced Deep Boltzmann Machines

- 2011: AlexNet — the first neural net winning the ImageNet challenge
- 2017: Invention of Transformer, the main architecture of LLM
- 2022: Invention of ChatGPT

[8] On the Origin of Deep Learning

More details on DL History and NN Architectures

Please read two links below:

- Deep Learning in a Nutshell: Core Concepts
- Deep Learning in a Nutshell: History and Training

Takeaway notes

1. Please introduce yourself and complete the **Assignment 1**

Takeaway notes

1. Please introduce yourself and complete the **Assignment 1**
2. Deep Learning is responsible for the most of the AI success today!

Takeaway notes

1. Please introduce yourself and complete the **Assignment 1**
2. Deep Learning is responsible for the most of the AI success today!
3. Let's get our journey started!

Thank you!