

AI in Culture and Arts – Workshop

Introduction to Artificial Intelligence and Interactive Machine Learning

Dr. Benedikt Zönnchen
22nd of October 2026

MUC.DAI
Munich Center for
Digital Sciences and AI



myt Hochschule
für Musik und Theater
München

1. How to Model Intelligence?

2. When is Learning Possible?

3. How Do Machines Learn?

4. How Do Humans Train Machines?

5. How Do Humans Teach Machines?

1. How to Model Intelligence?

2. When is Learning Possible?

3. How Do Machines Learn?

4. How Do Humans Train Machines?

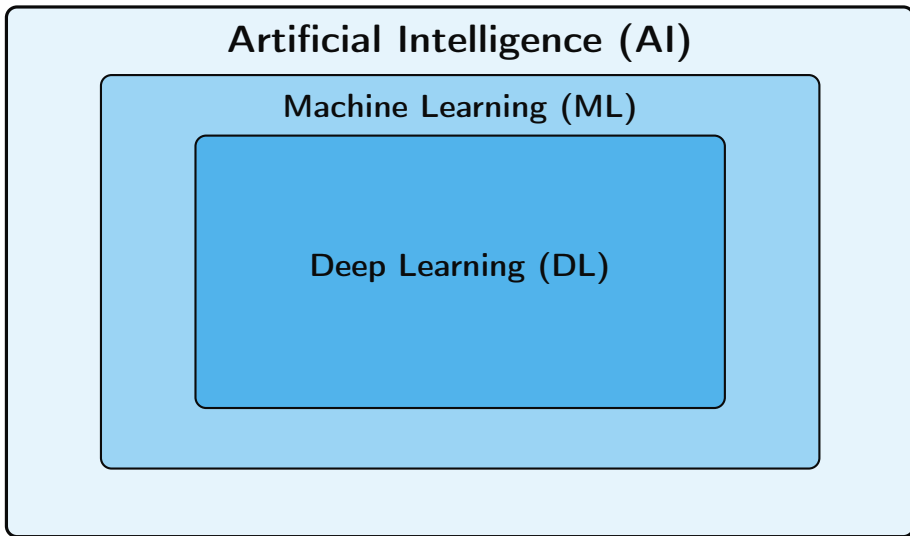
5. How Do Humans Teach Machines?

“The ability of an agent to achieve goals in a wide range of environments.”
– (Russell, 2019)

Let E be the space of all computable reward summable environmental measures with respect to the reference machine $\mathcal{U}$, and let K be the Kolmogorov complexity function. The expected performance of agent π with respect to the universal distribution $2^{-K(\mu)}$ over the space of all environments E is given by,

$$\Psi(\pi) := \sum_{\mu \in E} 2^{-K(\mu)} V_{\mu}^{\pi}.$$

We call this the **universal intelligence** of agent π (Legg & Hutter, 2007).



Most of the hype is based on **Deep Learning**.

Van Rooij et al. describe how the usage of the term “AI” shifted from *AI-as-Psychology* to *AI-as-Engineering* (van Rooij et al., 2024).

Historical (1970s–80s)	Contemporary (Today)
<i>AI-as-Psychology</i> Goal: Understand the mind. AI as a theoretical tool .	<i>AI-as-Engineering</i> Goal: Technical progress. AI as a product .

Symbolism

Exploits explicit, rule-based symbolic manipulation, logic, and structured reasoning to represent knowledge and solve problems.

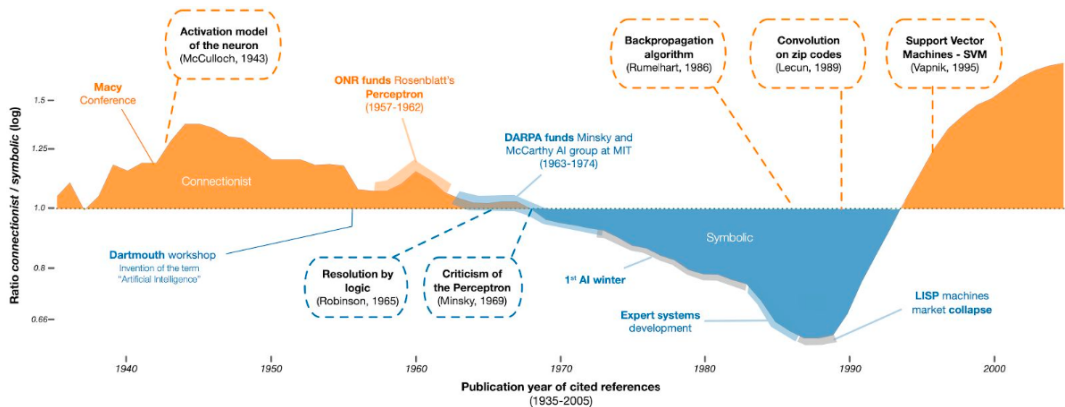
- **Assumption:** Intelligence uses high-level, human-readable symbols to represent problems and logic to solve them.
- **Motivation:** Model the **mind** (as a rational unit)!

Connectionism

Exploits artificial neural networks & statistics, emphasizing learning from patterns, distributed representations, and emergent behaviors.

- **Assumption:** Intelligence emerges from the interaction of simple and low-level units, i. e. biological neurons.
- **Motivation:** Model the **brain** (and the mind will emerge)!

How to Model Intelligence?



Source: (Cardon et al., 2018)

Assumptions of AI Research of the Time

- **The biological assumption:** The brain processes information in discrete operations by way of some biological equivalent of on/off switches.
- **The psychological assumption:** The mind can be viewed as a device operating on bits of information according to formal rules.
- **The epistemological assumption:** All knowledge can be formalized.
- **The ontological assumption:** The world consists of independent facts that can be represented by independent symbols.

"[Dreyfus'] derisiveness has been so provoking that he has estranged anyone he might have enlightened. And that's a pity." – Pamela McCorduck

Video

- **2006–2012: Early Breakthroughs**
 - Revival of deep neural networks combining backpropagation and gradient descent (Leibniz, 1676; Rosenblatt, 1962; Pontryagin, 1950; Dreyfus, 1962; Amari, 1967; Linnainmaa, 1970; Werbos, 1982; Rumelhart, 1982; LeCun, 1987).
 - Major success: ImageNet competition (AlexNet, 2012)
- **2013–2017: Rapid Expansion of Deep Learning**
 - CNN dominance in image recognition
 - Reinforcement Learning breakthroughs (AlphaGo, 2016)
- **2018–Present: Transformers and Large Language Models**
 - Natural Language Processing revolution (BERT, GPT)
 - Generative AI models (GPT-3, ChatGPT, Stable Diffusion)
- **Modern Symbolic Revival and Hybrid AI (Ongoing)**
 - Neurosymbolic AI: Combining deep learning and symbolic reasoning
 - Explainable AI (XAI): Interpretable, trustworthy systems
 - Importance for ethics, transparency, and safety

1. Computational Power

- Massive increase in GPU computing capacity.
- Parallelization of neural network training.

2. Big Data

- Availability of huge, labeled datasets.
- Online data collection (ImageNet, web scraping).

3. Algorithmic Advances

- Improved training algorithms (e. g., backpropagation, Adam optimizer).
- New network architectures (CNNs, RNNs, Transformers).

4. Open-source Ecosystem

- Accessible tools: TensorFlow, PyTorch.
- Community-driven knowledge sharing.

Video

1. How to Model Intelligence?

2. When is Learning Possible?

3. How Do Machines Learn?

4. How Do Humans Train Machines?

5. How Do Humans Teach Machines?

“The organism feeds on negative entropy.” – (Schrödinger, 1944)

0101010101010101010101010101

110011010101110100011100101111

Let's say you see the following two 30-bits:

$$x_1 = 010101010101010101010101010101$$

$$x_2 = 110011010101110100011100101111$$

Question: Now guess the 31st bit! Now someone tells you the 31st bit. What bit informed you “more”?

Answer: We say the **process generating** x_2 has a higher **entropy** (Shannon, 1948) than the one generating x_1 . Because its pattern is unpredictable, we are more “often” **surprised** when reading the bits of x_2 from left to right.

Information entropy measures uncertainty / surprise. It is the **expected value of surprise**.

- High entropy: highly unpredictable, many possible outcomes, each similarly likely
- Low entropy: predictable, only a few likely outcomes

Learning requires

$$\text{entropy} + \text{structure} \approx \text{complexity}.$$

We have to be surprised but also be able to exploit a structure to compress observations.

A physical formula like

$$F = m \cdot a$$

can be thought of as a **highly compressed representation** of some aspects of physical reality!

In other words, $F = m \cdot a$ reveals the structure of motion, thus the informational entropy cannot be too high.

Some neuroscientists (K. Friston, Kilner, & Harrison, 2006; K. J. Friston, 2011) think that organisms try to **minimize surprise** by

1. adjusting expectations (perception, learning) and
2. realize expectations (acting).

According to this school of neuroscience,

organisms are their own existence proof.

By acting to keep themselves alive, they continuously generate sensory inputs that confirm their continued existence.

1. How to Model Intelligence?

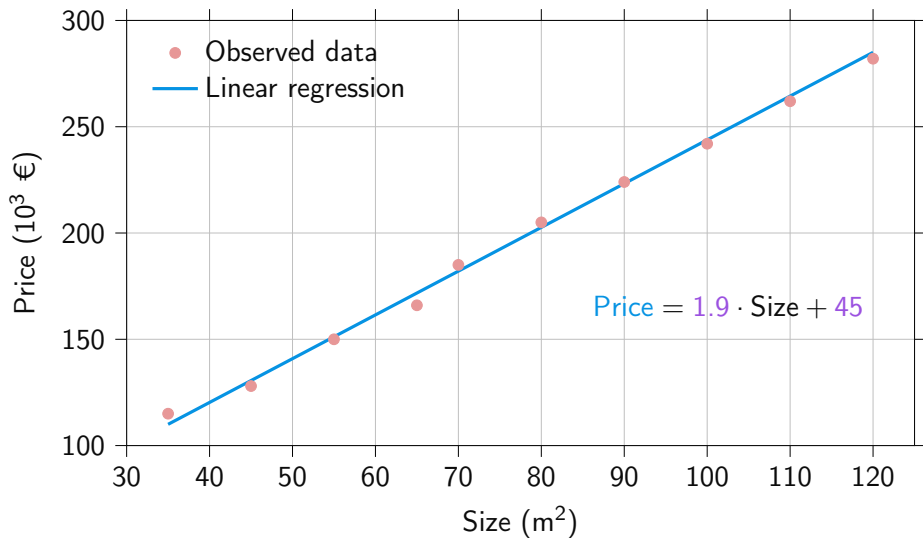
2. When is Learning Possible?

3. How Do Machines Learn?

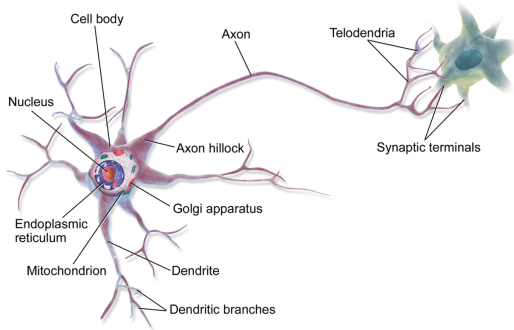
4. How Do Humans Train Machines?

5. How Do Humans Teach Machines?

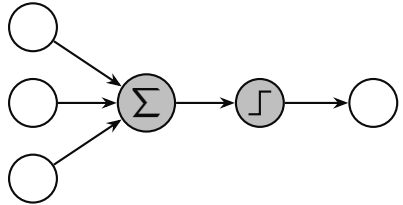
Learning a model of “the world”



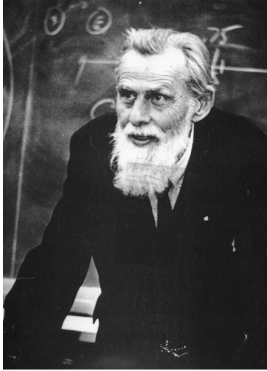
Neuron



Artificial neuron

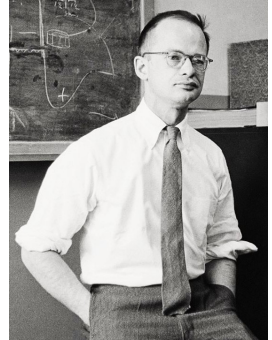
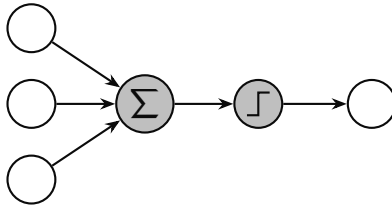


The Mathematical Neuron



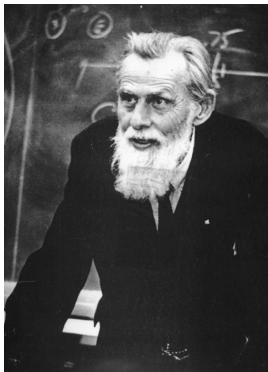
Warren S. McCulloch

"Because of the 'all-or-none' character of nervous activity, neural events and the relations among them can be treated by means of propositional logic."



Walter H. Pitts Jr.

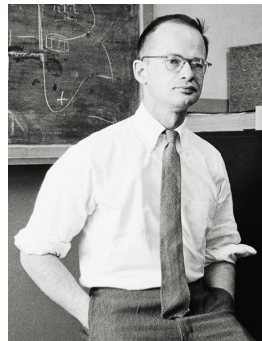
The Mathematical Neuron



Warren S. McCulloch

"Because of the 'all-or-none' character of nervous activity, neural events and the relations among them can be treated by means of propositional logic."

$$f(x_1, \dots, x_n) = \begin{cases} 1 & \text{if } \sum_{k=1}^n x_k > 1 \\ -1 & \text{otherwise.} \end{cases}$$



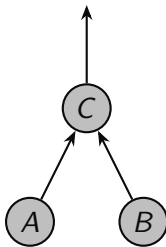
Walter H. Pitts Jr.

The Mathematical Neuron

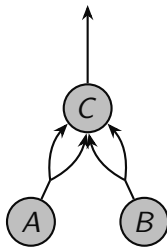
Complex logical operations can be performed using networks of binary neurons.



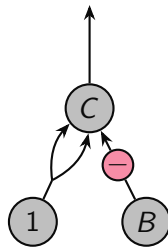
Identity: $C = A$



And: $C = A \wedge B$

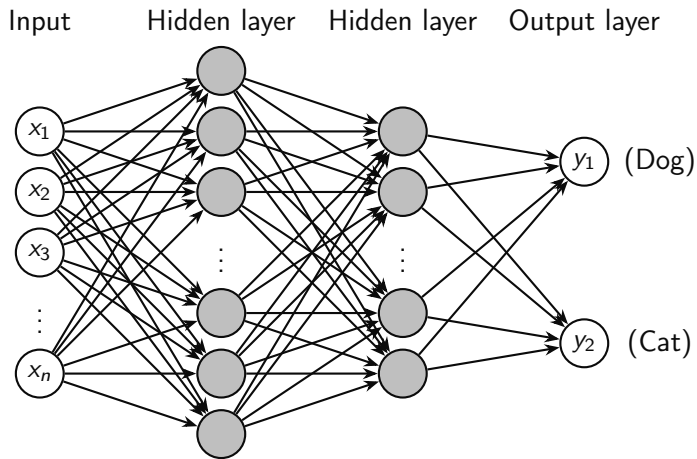


Or: $C = A \vee B$



Negation: $C = \neg B$

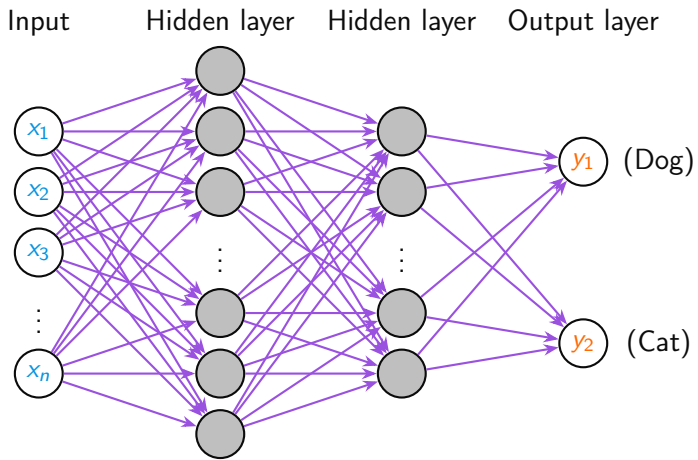
Synaptic Plasticity



$$h_{\theta}(\mathbf{x}) = \mathbf{y},$$

where $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_k)$

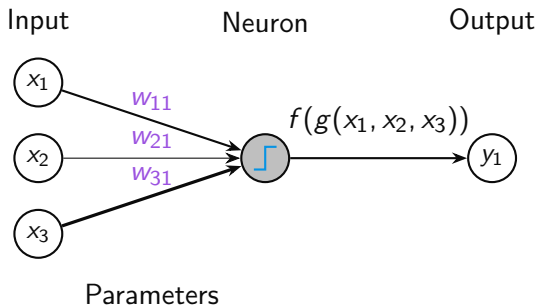
Synaptic Plasticity



$$h_{\theta}(\mathbf{x}) = \mathbf{y},$$

where $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_k)$

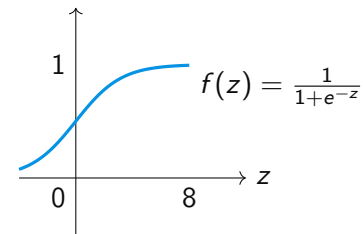
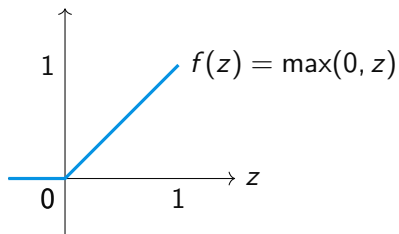
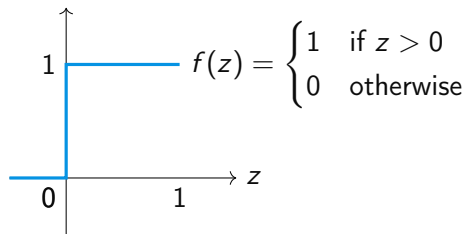
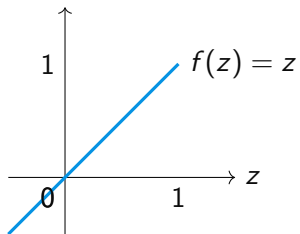
Synaptic Plasticity



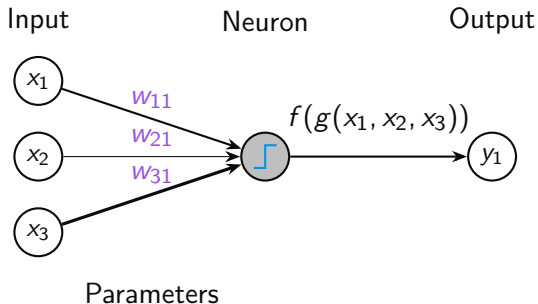
Parameters determine how strong neurons are wired together:

$$g(x_1, x_2, x_3) = x_1 \cdot w_{11} + x_2 \cdot w_{21} + x_3 \cdot w_{31}$$

Activation Functions



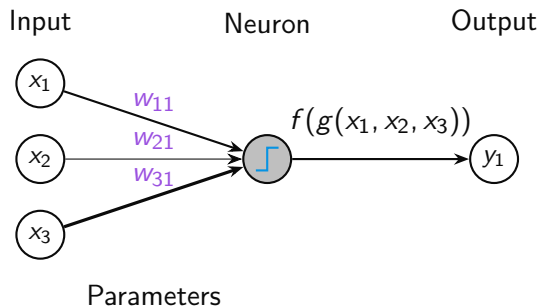
Synaptic Plasticity (Biology)



"Neurons that fire together, wire together."

$$w_{ij} = w_{ij} + \eta \cdot x_i \cdot y_j$$

Decentralized way for a brain to self-organize based on pattern of experience without a built in "goal" or "error correction"!



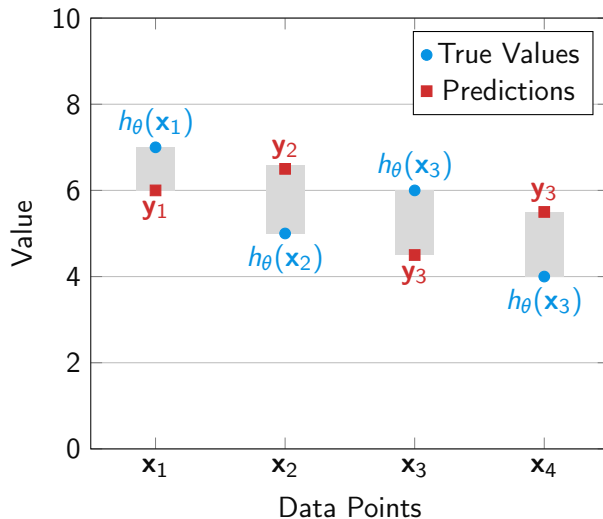
"Trace the error, tune the weight."

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla J(\theta_t) \text{ with } \theta_t = (w_{11}, w_{21}, w_{31})$$

Global error-driven optimization algorithm!

Cost Function (Regression)

Mean Squared Error (MSE):



Mean Squared Error (MSE):

$$J(\theta) = \frac{1}{N} \sum_{i=1}^N (\mathbf{y}_i - h_{\theta}(\mathbf{x}_i))^2$$

where $\mathbf{y}_i$ is the correct label of a data point $\mathbf{x}_i = (x_1, \dots, x_n)$ in our training data.

Example: Let us suppose

$$h_{\theta}(x_1, x_2) = w_{11} \cdot x_1 + w_{21} \cdot x_2$$

and let us assume $w_{11} = 2$, $w_{21} = 0.5$ and we have two data points $\mathbf{x}_1 = (1, 1)$, $y_1 = 1$ and $\mathbf{x}_2 = (-1, -2)$, $y_2 = -3$. Then our mean squared error is:

$$\begin{aligned} J(w_{11}, w_{21}) &= \frac{1}{2} [(1 - (w_{11} \cdot 1 + w_{21} \cdot 1))^2 + (-3 - (w_{11} \cdot (-1) + w_{21} \cdot (-2)))^2] \\ &= \frac{1}{2} [(1 - (2 \cdot 1 + 0.5 \cdot 1))^2 + (-3 - (2 \cdot (-1) + 0.5 \cdot (-2)))^2] \end{aligned}$$

The gradient would be:

$$\nabla J(w_{11}, w_{21}) = \begin{bmatrix} -4 + 2w_{11} + 3w_{21} \\ -7 + 3w_{11} + 5w_{21} \end{bmatrix}$$

Categorical Cross Entropy Cost:

$$J(\theta) = -\frac{1}{N} \sum_{i=1}^N [\mathbf{y}_i \cdot \log(h_{\theta}(\mathbf{x}_i))]$$

where $\mathbf{y}_i$ is interpreted as the probability distribution of categories for $\mathbf{x}_i = (x_1, \dots, x_n)$, i.e. a data point.

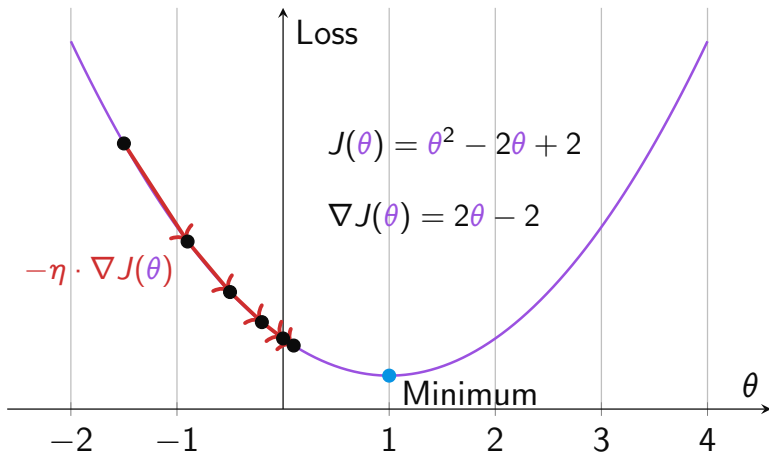
To improve the model's prediction, we try to minimize the cost function. One way to do this is **gradient descent**:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla J(\theta_t)$$

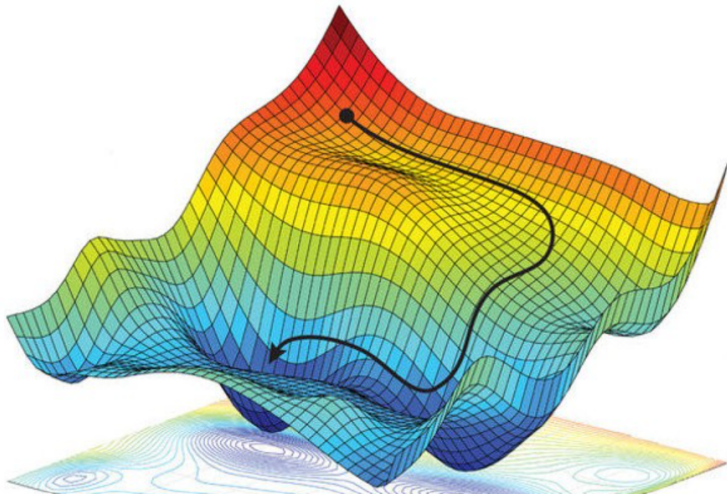
Condition: $\nabla J(\theta_t)$ exists!

Interactive Tutorial

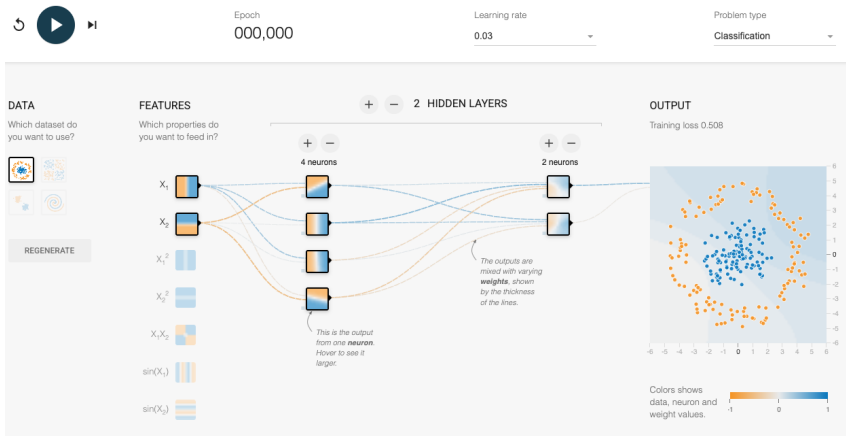
Gradient Descent



Gradient Descent



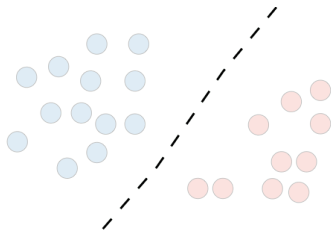
Design and Try Your Perceptron



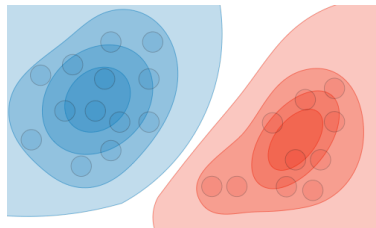
Simplified Tensorflow Playground

Extended Tensorflow Playground

- **Discriminative models:** Learn the boundaries of decisions.
- **Generative models:** Learn the whole distribution of the data.



Discriminative modelling



Generative modelling

Artificial neural networks (ANNs) are the most prominent model types currently in use. However machine learning spans over much more model types such as:

Model type	Best for	Complexity
Linear models	Simple trends, high interpretability	Low
Tree-based	Tabular data, excel-style sheets	Medium
SVM	Clear margins of separation	Medium
Bayesian	Risk assessment, small datasets	Low/Medium
Neural networks	Images, audio, large-scale text	High

Table: Comparison of machine learning model types.

Geoffrey Hinton on AI and consciousness:

Video

Yann LeCun on the need for a paradigm shift:

Video 1

Video 2

1. How to Model Intelligence?

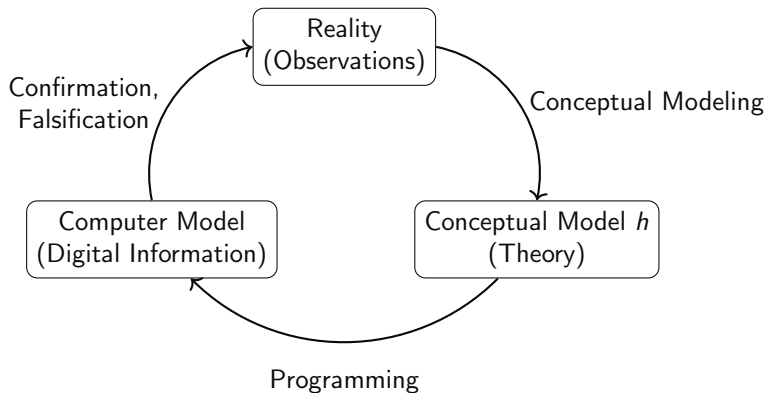
2. When is Learning Possible?

3. How Do Machines Learn?

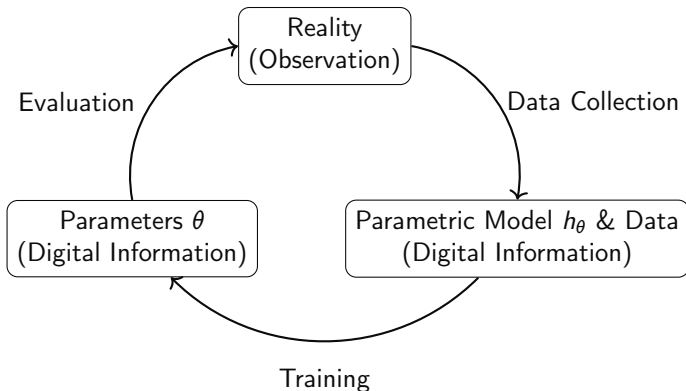
4. How Do Humans Train Machines?

5. How Do Humans Teach Machines?

Theory-driven Modeling

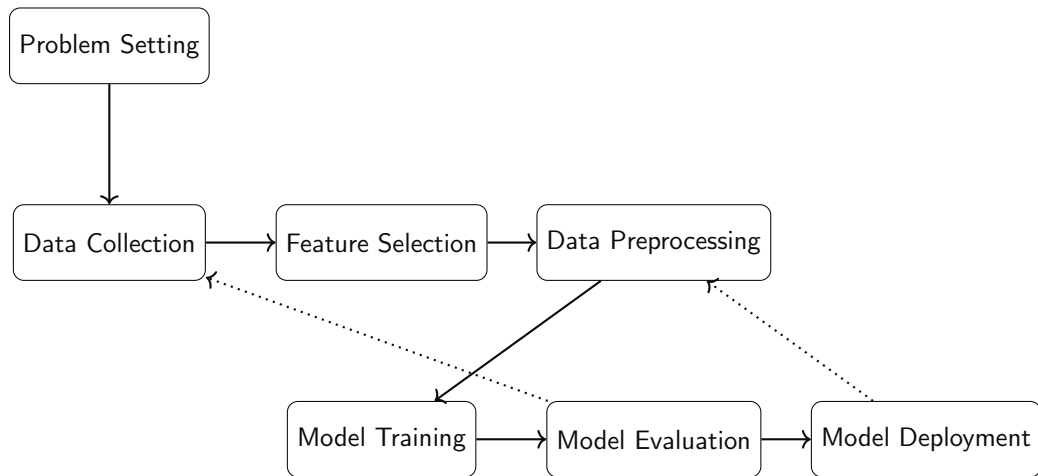


Minds constructs a (falsifiable) theory or hypothesis about reality to test against.



Algorithms (directly) fit a parametric model to the data. **Minds** are usually unable to conceptualize the trained model.

Development Cycle



```
100 class fast_qlinear(torch.autograd.Function):
101     def forward(ctx, a, b, scales, zeros):
102
103         m, k = a.shape
104         _, n = b.shape
105
106         quant_groupsize = 128
107         block_size_m = 16
108         block_size_n = 32 # [N = 4096 // 32] = 128 blocks
109         block_size_k = 256
110         group_size_m = 8
111         num_warps = 4
112         num_stages = 8
113         total_blocks_m = triton.cdiv(m, block_size_m)
114         total_blocks_n = triton.cdiv(n, block_size_n)
```

Python and ML libraries (PyTorch, tensorflow, JAX etc.)

Train a Model with Python

1. How to Model Intelligence?

2. When is Learning Possible?

3. How Do Machines Learn?

4. How Do Humans Train Machines?

5. How Do Humans Teach Machines?

Wekinator is a free, open source software that allows anyone to use machine learning to build new musical instruments, gestural game controllers, computer vision or computer listening systems, and more (Fiebrink, 2019).

www.wekinator.org

ml5js is an open source library with a goal of making machine learning approachable for a broad audience of artists, creative coders, and students.

Examples

InteractML brings machine learning visual scripting to Unity, empowering game creators to develop novel gameplay mechanics and control schemes without writing code. You can create ML models by joining nodes together and visualize in real-time the data from a Unity scene in the graph (Bernardo, Zbyszyński, Grierson, & Fiebrink, 2020).

<https://interactml.com/>

Teachable Machine is Google's online tool which you can use to recognize sound, images, and poses.

<https://teachablemachine.com/>

Marcelle is a software for composing interactive machine learning workflows and interfaces (Françoise, Caramiaux, & Sanchez, 2021).

<https://marcelle.dev/>

Interactive Machine Learning: The Marcelle Toolkit

Marcelle Example - Dashboard

Data Management

Training

Batch Prediction

Real-time Prediction

webcam

☐ activate video

mobileNet

Using Mobilenet v1 with alpha = 1.

Instance label

Capture instances to the training set

dataset browser

This dataset contains 65 instances.



[View More](#)

Bene



[View More](#)

Nicht Bene



[View More](#)

“Machine learning allows for the creation of new creative outcomes and new creative relationships between people and machines.” – Rebecca Fiebrink

Any questions?

References I

- Bernardo, F., Zbyszyński, M., Grierson, M., & Fiebrink, R. (2020, apr). Designing and evaluating the usability of a machine learning API for rapid prototyping music technology. *Frontiers in Artificial Intelligence*, 3. doi: 10.3389/frai.2020.00013
- Cardon, D., Cointet, J.-P., & Mazieres, A. (2018). Neurons spike back: The invention of inductive machines and the artificial intelligence controversy. *Rezeaux*, 36, 173–220.
- Dreyfus, H. L. (1965). *Alchemy and artificial intelligence*. Santa Monica, CA: RAND Corporation.
- Dreyfus, H. L., & Dreyfus, S. E. (1986). From socrates to expert systems: The limits of calculative rationality. In C. Mitcham & A. Huning (Eds.), *Philosophy and technology ii: Information technology and computers in theory and practice* (pp. 111–130). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-009-4512-8_9
- Esposito, E. (2022). *Artificial communication*. The MIT Press. doi: 10.7551/mitpress/14189.001.0001

References II

- Farrell, H., Gopnik, A., Shalizi, C., & Evans, J. (2025). Large AI models are cultural and social technologies. *Science*, 387(6739), 1153-1156. doi: 10.1126/science.adt9819
- Fiebrink, R. (2019, dec). Machine learning education for artists, musicians, and other creative practitioners. *ACM Transactions on Computing Education*, 19(4), 1–32. doi: 10.1145/3294008
- Françoise, J., Caramiaux, B., & Sanchez, T. (2021). Marcelle: Composing interactive machine learning workflows and interfaces. In *The 34th annual acm symposium on user interface software and technology* (pp. 39–53). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3472749.3474734
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1), 70–87. doi: 10.1016/j.jphysparis.2006.10.001

References III

- Friston, K. J. (2011). Embodied inference: or “I think therefore I am, if I am what I think”. In W. Tschacher & C. Bergomi (Eds.), *The implications of embodiment (cognition and communication)* (pp. 89–125). Exeter, UK: Imprint Academic.
- Hui, Y. (2024). *Machine and sovereignty: For a planetary thinking*. Minneapolis, MN: University of Minnesota Press.
- Legg, S., & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. *Minds and Machines*, 17(4), 391–444. doi: 10.1007/s11023-007-9079-x
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. Reidel Publishing Company.
- Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. New York: Viking.
- Schrödinger, E. (1944). *What is life? The physical aspect of the living cell*. Cambridge, UK: Cambridge University Press.

- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(3), 379-423.
- van Rooij, I., Guest, O., Adolphi, F., de Haan, R., Kolokolova, A., & Rich, P. (2024, dec). Reclaiming AI as a theoretical tool for cognitive science. *Computational Brain & Behavior*, 7(4), 616–636. doi: 10.1007/s42113-024-00217-5
- Zönnchen, B., Dzhimova, M., & Socher, G. (2025). From intelligence to autopoiesis: rethinking artificial intelligence through systems theory. *Frontiers in Communication*, Volume 10 - 2025. doi: 10.3389/fcomm.2025.1585321