

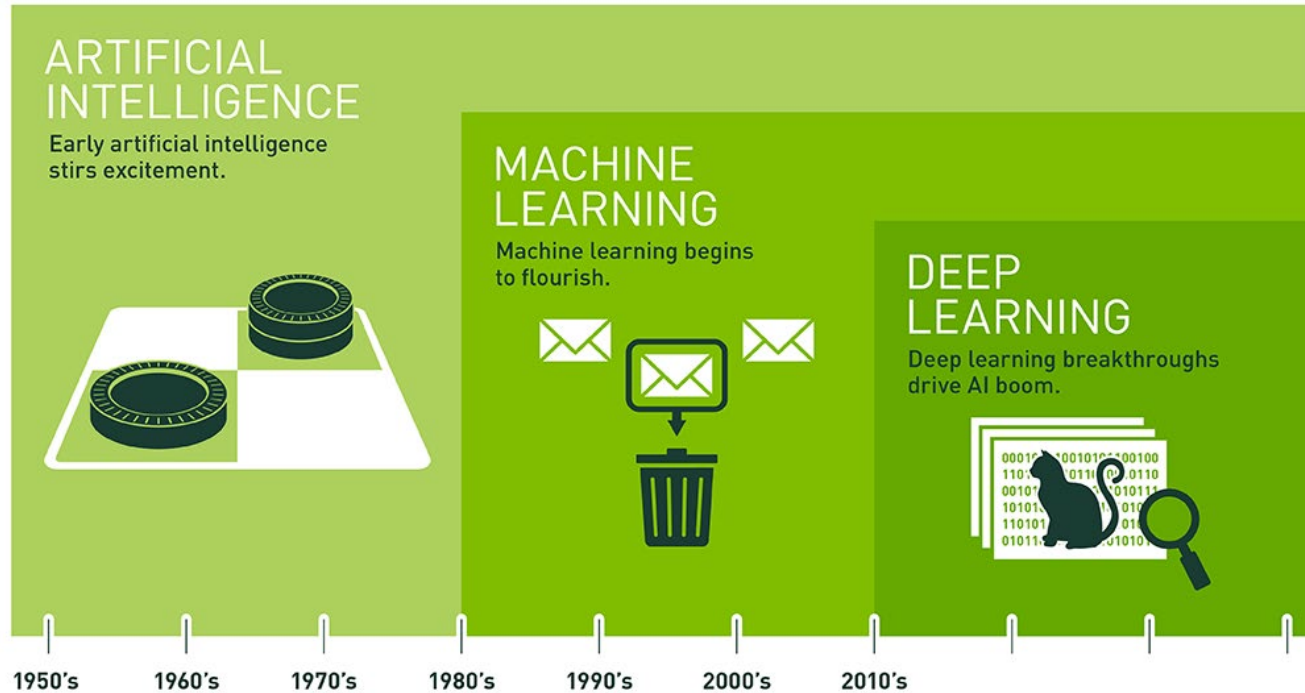
Big Data und Machine Learning

17. Mai 2022

Begriffe

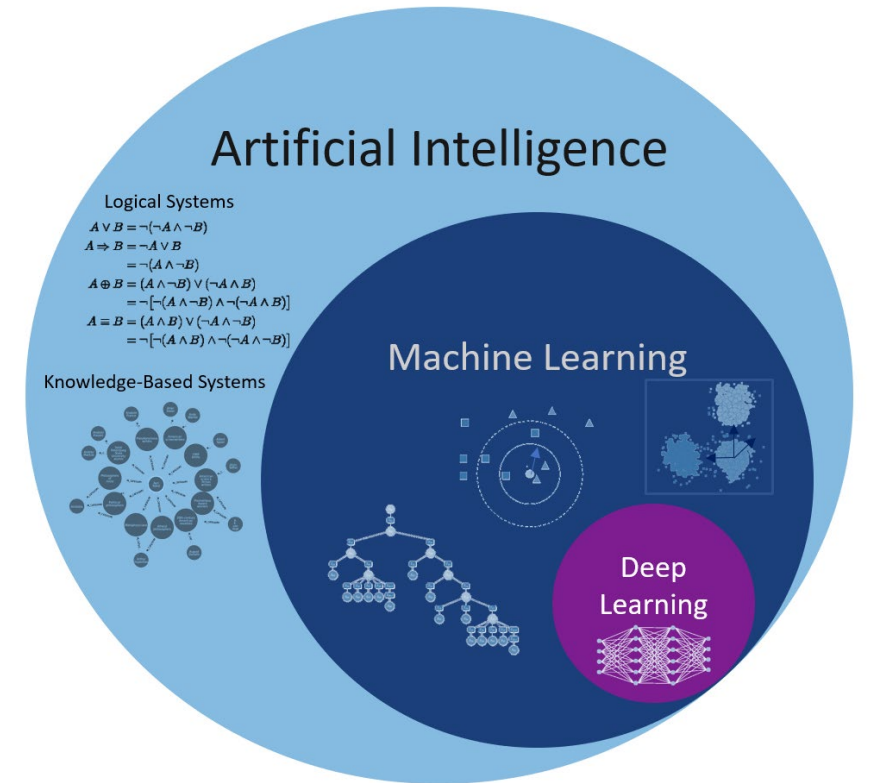
- Artificial Intelligence
- Machine Learning
- Big Data
-

Entwicklung



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.


Quelle: <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>



Quelle: <https://data-science-blog.com/blog/2018/05/14/machine-learning-vs-deep-learning-wo-liegt-der-unterschied/>

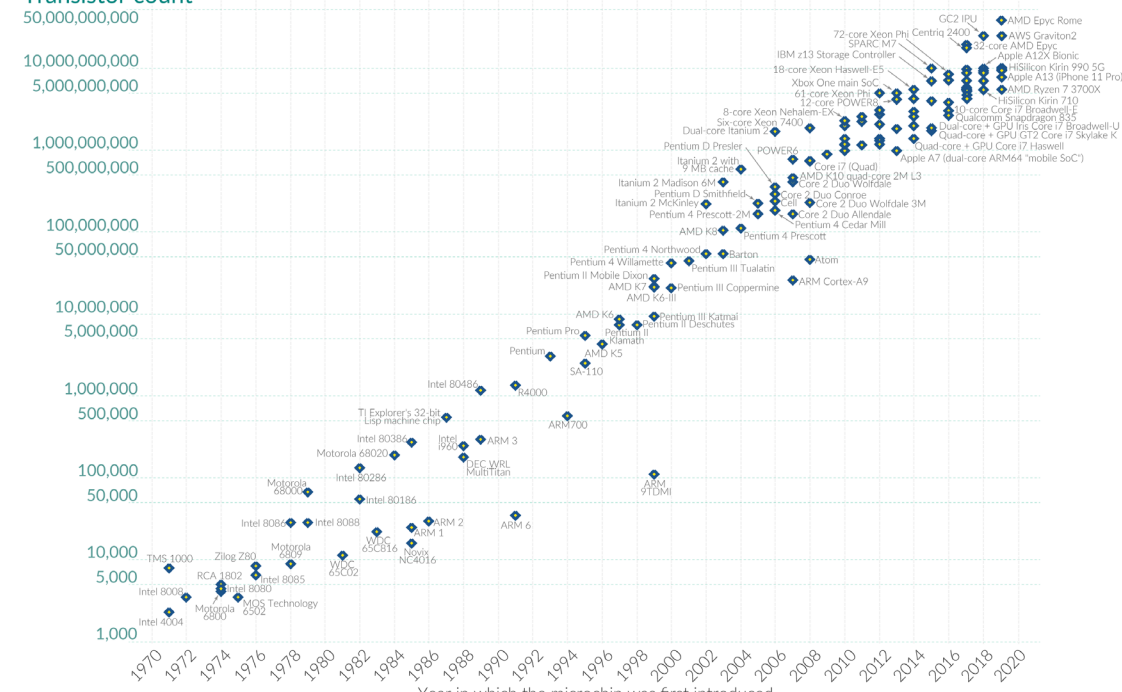
Entwicklung: Moores Law

- Transistorendichte in einem Schaltkreis verdoppelt sich in gleichbleibenden und regelmässigen Abständen.

Moore's Law: The number of transistors on microchips doubles every two years 

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.

Transistor count
50,000,000,000

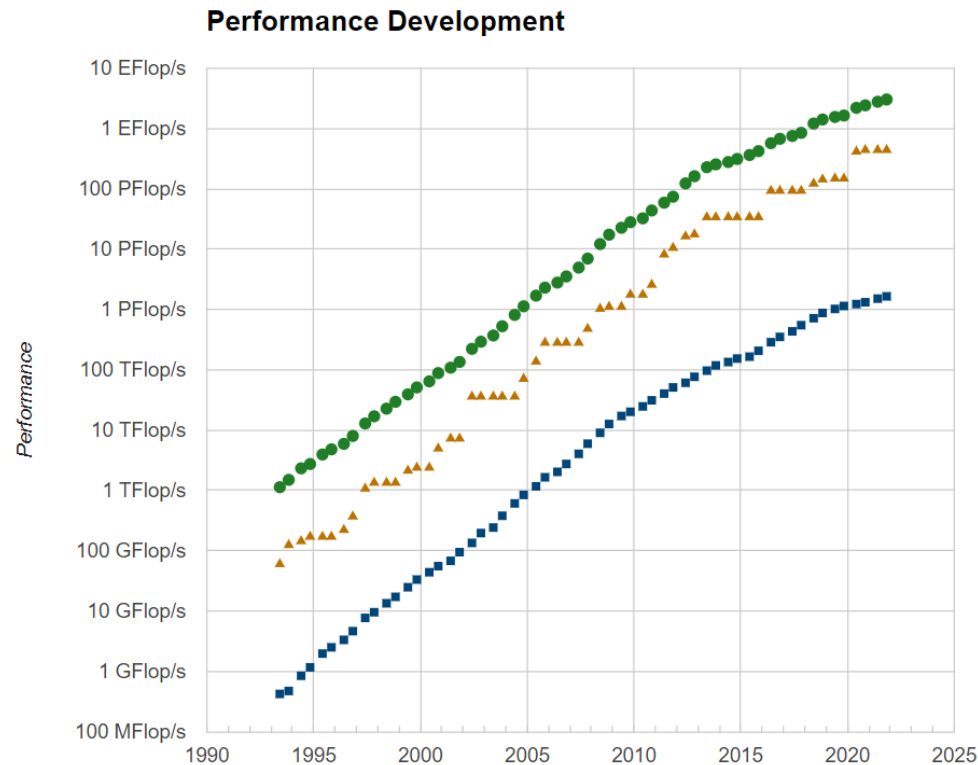


Data source: Wikipedia (wikipedia.org/wiki/Transistor_count) OurWorldinData.org – Research and data to make progress against the world's largest problems. Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

Quelle: <https://ourworldindata.org/uploads/2020/11/Transistor-Count-over-time.png>

Entwicklung: Rechengeschwindigkeit

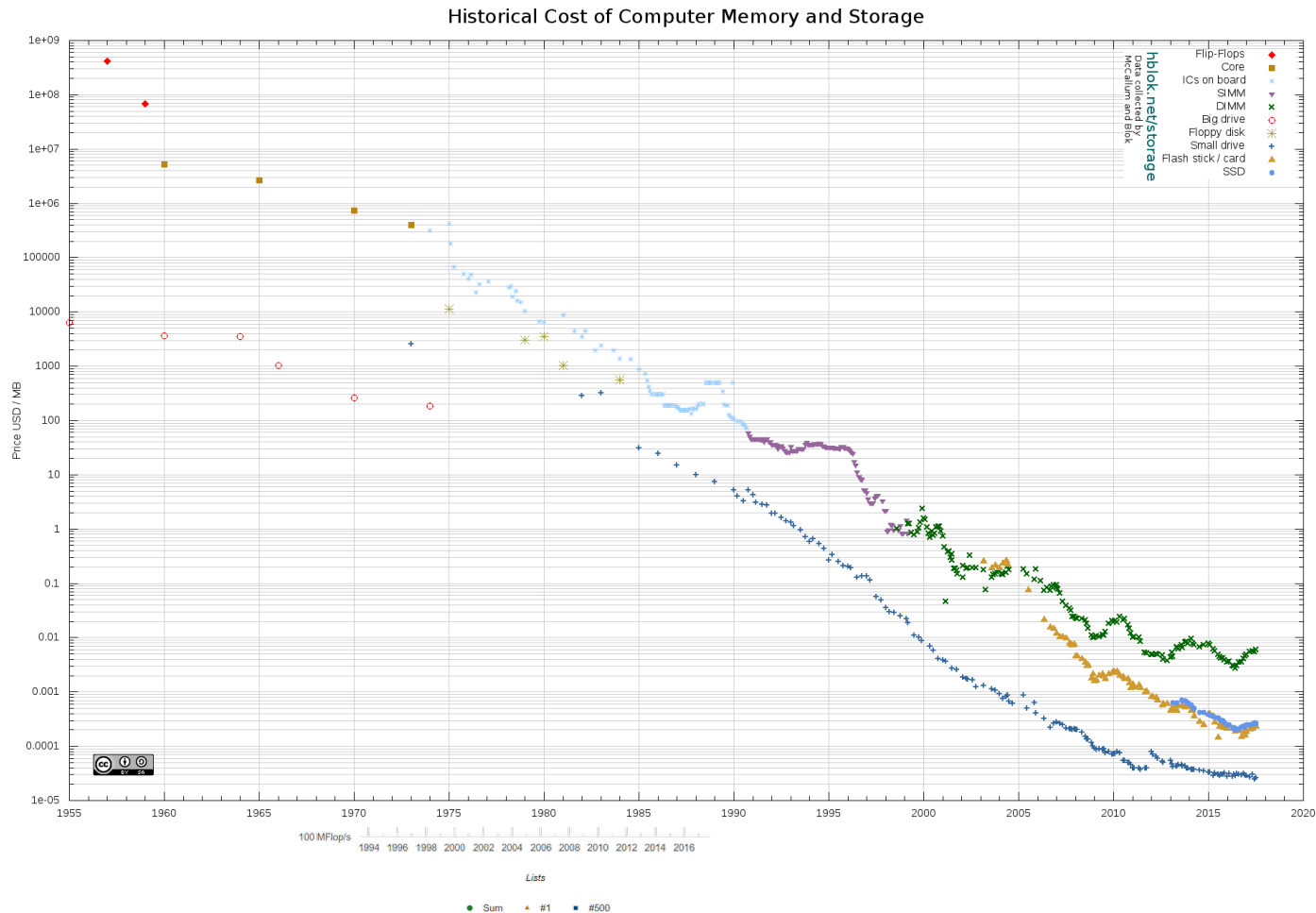
- FLOPS = «Floating Point Operations per Second», Giga: 10^9 , Tera: 10^{12} , Peta: 10^{15} , Exa: 10^{18}



Exkurs: Präfixe

Präfix	Beschreibung	Faktor	
E	Exa	$10^{18} =$	1000000000000000000
P	Peta	$10^{15} =$	1000000000000000
T	Tera	$10^{12} =$	1000000000000
G	Giga	$10^9 =$	1000000000
M	Mega	$10^6 =$	1000000
k	Kilo	$10^3 =$	1000
h	Hecto	$10^2 =$	100
da	Deko	$10^1 =$	10
da	Deci	$10^{-1} =$	0.1
c	Centi	$10^{-2} =$	0.01
m	Milli	$10^{-3} =$	0.001
μ	Micro	$10^{-6} =$	0.000001
n	Nano	$10^{-9} =$	0.000000001
p	Pico	$10^{-12} =$	0.000000000001
f	Femto	$10^{-15} =$	0.000000000000001
a	Atto	$10^{-18} =$	0.000000000000000001

Entwicklung: Speicherpreise



Quelle: https://hblok.net/storage_data/storage_memory_prices-2017-12_large.png

Exkurs: Logarithmische Skalen und Wachstum

Lernalgorithmen: Charakteristika

Ein Lernalgorithmus hat üblicherweise mindestens drei Charakteristiken

- Eine **Aufgabe** (task, T)
 - Entscheide auf Grund Bild ob Mann oder Frau
 - Schätze auf Grund Kaufverhalten Cumulus Haushaltseinkommen
 - Entscheide ob Brustkrebs auf MRI Bild vorliegt oder nicht.
 - Spracherkennung, Übersetzung, etc.
- Eine **Gütemessung** (performance measure, P)
 - Missklassifikationsrate
 - Absolute oder quadratische Abweichung
- Eine **Erfahrungdatensatz** (experience, E)
 - In Kenntnis der wahren Begebenheit («supervised learning»)
 - In Unkenntnis der wahren Begebenheit («unsupervised learning»)

Lernalgorithmen: Toy-Example

- Auf Grund der Körpergrösse die Schuhgrösse vorhersagen

Lernalgorithmen: Einschränkung EF

Wir beschränken uns als Beispiel auf sogenannte «supervised classification tasks», d.h.,

- wir kennen die **Attribute** (features) einer Beobachtung: X_i
- wir kennen die **Klasse** (outcome) einer Beobachtung: Y_i
- wir möchten neuen Beobachtungen mit **bekanntem** Attributen und **unbekannter** Klasse eine Klasse zuordnen.

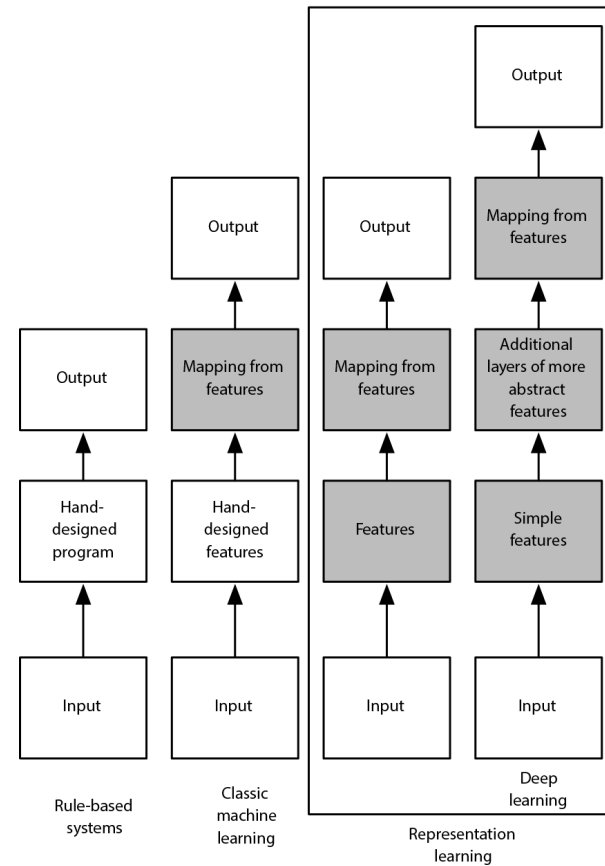
X_i ist dabei ein Vektor in \mathbb{R}^n , Y_i nimmt Werte in \mathbb{N}_0 an (z.B. 0 und 1 oder, 1,2,3,4,5, etc.)

Lernalgorithmen: Neuronale Netze

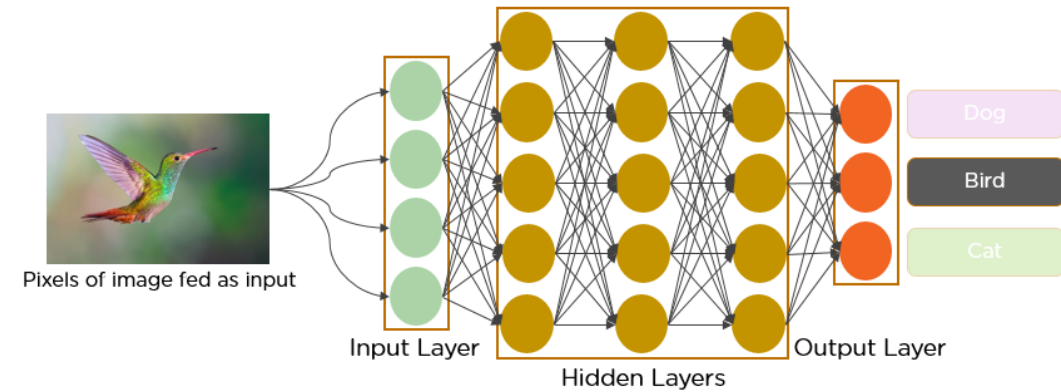
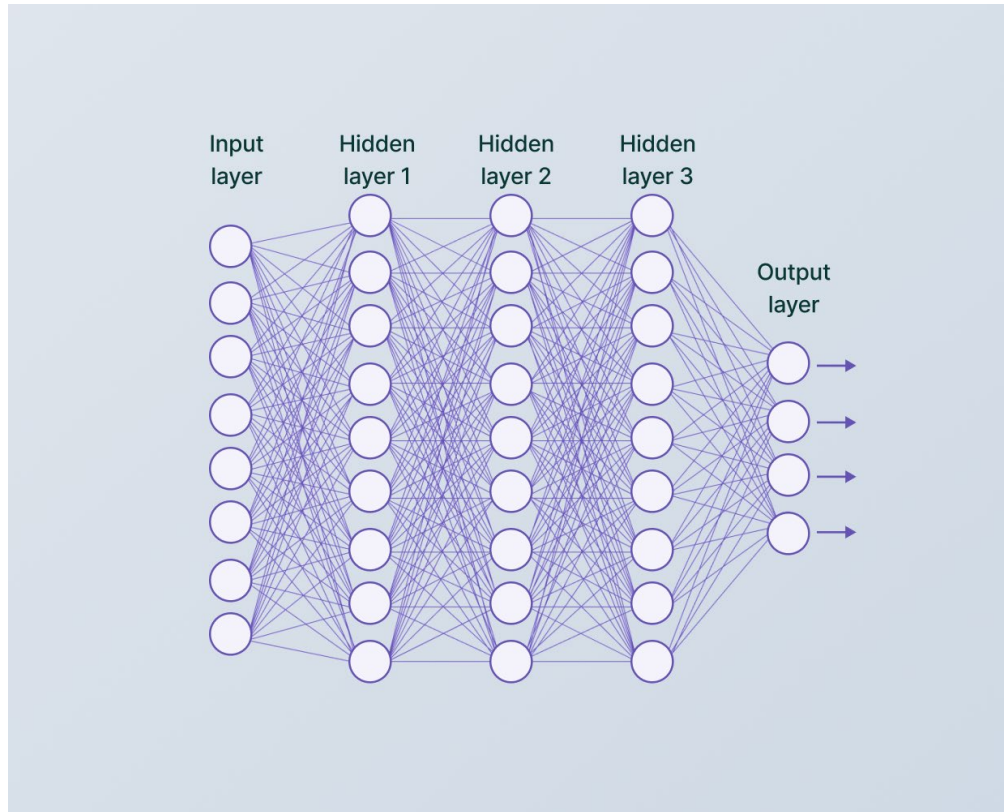
Neuronale Netze haben über die Jahre stark an Popularität gewonnen sind aber letzte Endes schon lange bekannt:

- Cybernetics in den 50er und 60er Jahren
- Connectionism in the 80er und 90er Jahren
- Artificial Neural Networks (ANNs)
- «Deep Learning» ist dabei der aktuell häufig verwendete Begriff


Lernalgorithmen: Neuronale Netze



Lernalgorithmen: Neuronale Netze



Misslearning

Download PDF 

Article | Published: 31 May 2021

AI for radiographic COVID-19 detection selects shortcuts over signal

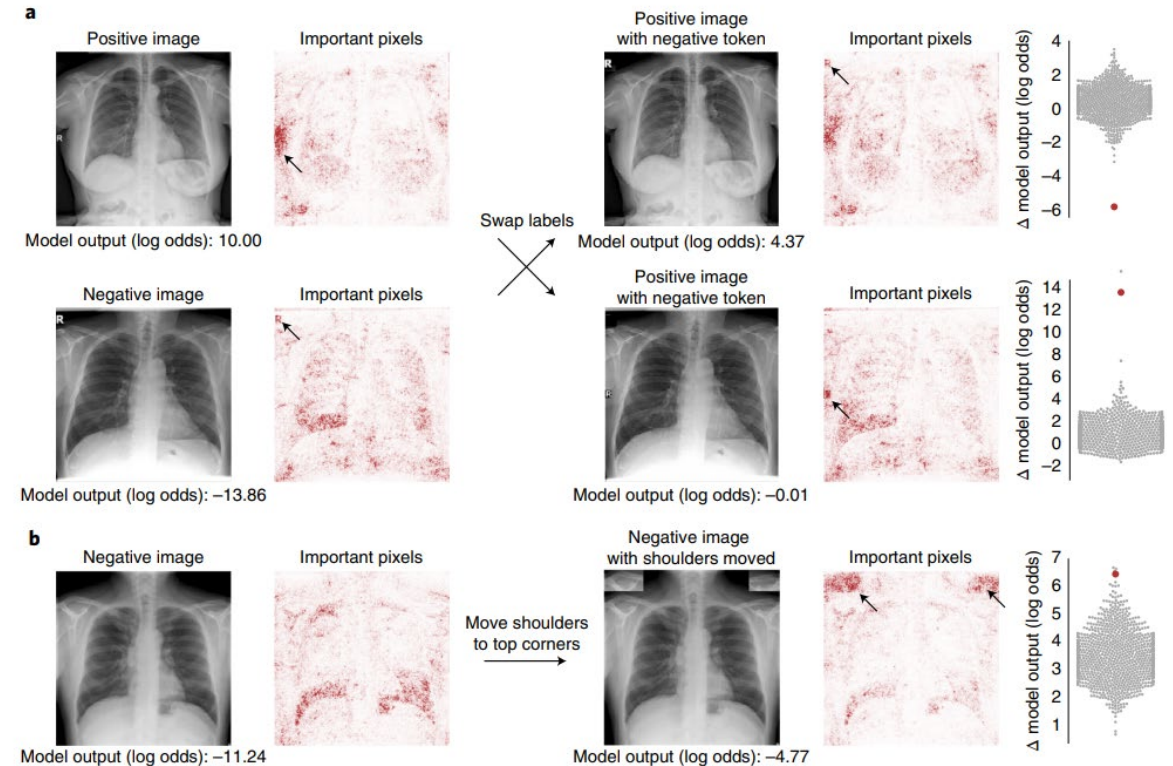
Alex J. DeGrave, Joseph D. Janizek & Su-In Lee 

Nature Machine Intelligence 3, 610–619 (2021) | [Cite this article](#)

20k Accesses | 40 Citations | 379 Altmetric | [Metrics](#)

Abstract

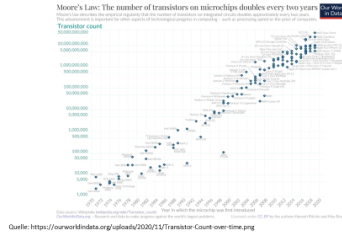
Artificial intelligence (AI) researchers and radiologists have recently reported AI systems that accurately detect COVID-19 in chest radiographs. However, the robustness of these systems remains unclear. Using state-of-the-art techniques in explainable AI, we demonstrate that recent deep learning systems to detect COVID-19 from chest radiographs rely on confounding factors rather than medical pathology, creating an alarming situation in which the systems appear accurate, but fail when tested in new hospitals. We observe that the approach to obtain training data for these AI systems



Curse of Dimensionality

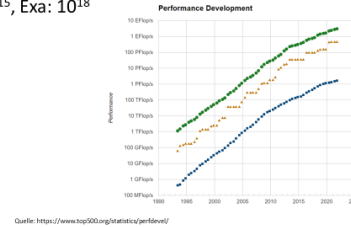
Entwicklung: Moores Law

- Transistorendichte in einem Schaltkreis verdoppelt sich in gleichbleibenden und regelmässigen Abständen.

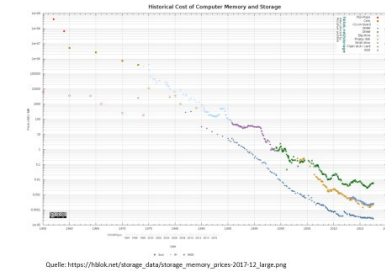


Entwicklung: Rechengeschwindigkeit

- FLOPS = «Floating Point Operations per Second», Giga: 10^9 , Tera: 10^{12} , Peta: 10^{15} , Exa: 10^{18}



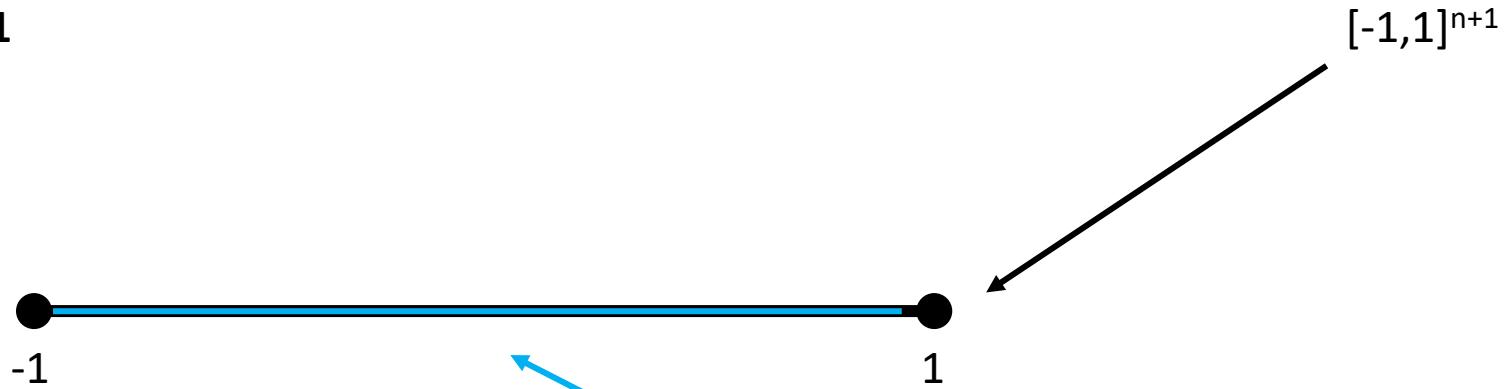
Entwicklung: Speicherpreise



Wieso sind neuronale Netze nicht noch viel besser?!

Curse of Dimensionality

$n = 1$

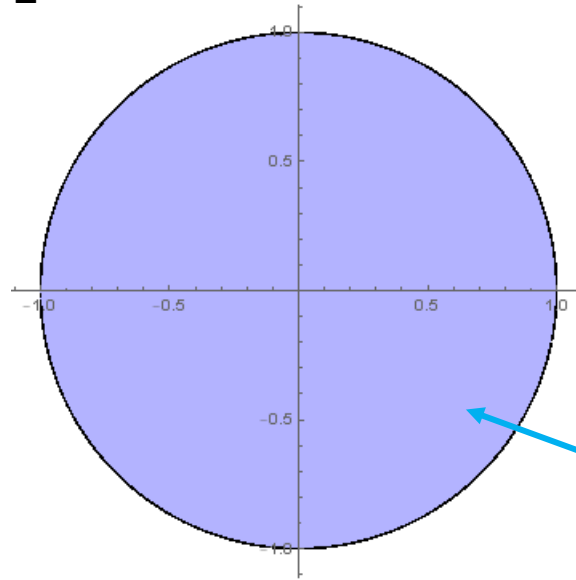


$$B^n = \{x \in \mathbb{R}^n \mid |x| \leq 1\}$$

Anteil Sphäre: 100%

Curse of Dimensionality

$n = 2$

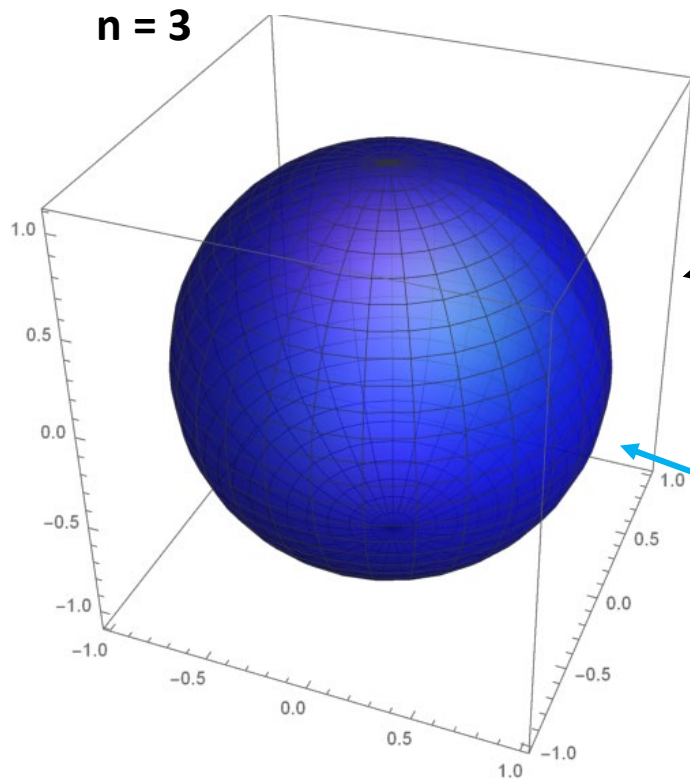


$[-1,1]^{n+1}$

$$B^n = \{x \in \mathbb{R}^n \mid |x| \leq 1\}$$

Anteil Sphäre: 79%

Curse of Dimensionality



$[-1,1]^n$

$$B^n = \{x \in \mathbb{R}^n \mid |x| \leq 1\}$$

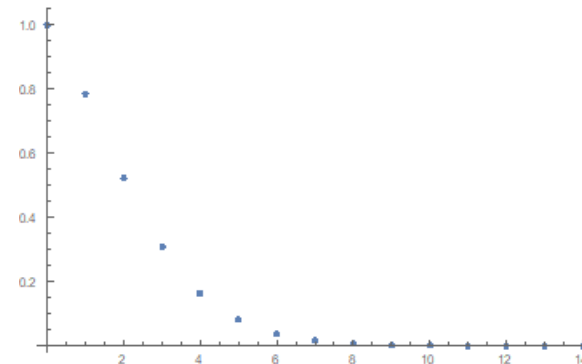
Anteil Sphäre: 52%

Curse of Dimensionality

Volumen der Einheitssphäre

$$\frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}$$

Anteil des Einheitssphärenvolumens am Einheitshyperkubus



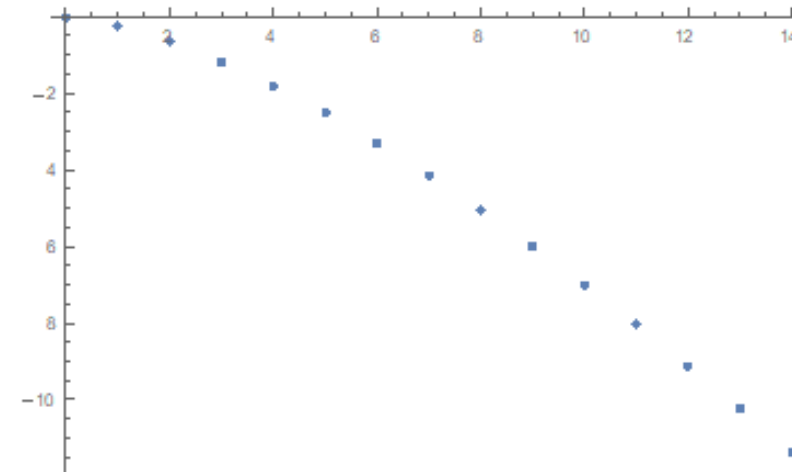
Mit zunehmender *Anzahl features* braucht es *superexponentiell* mehr Daten für die gleiche Datendichte in einer kleinen Umgebung

Curse of Dimensionality

Volumen der Einheitssphäre

$$\frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)}$$

Anteil des Einheitssphärenvolumens am Einheitshyperkubus
(log-Achse)



Mit zunehmender Anzahl *features* braucht es *superexponentiell* mehr Daten für die gleiche Datendichte in einer kleinen Umgebung

Fragen?