

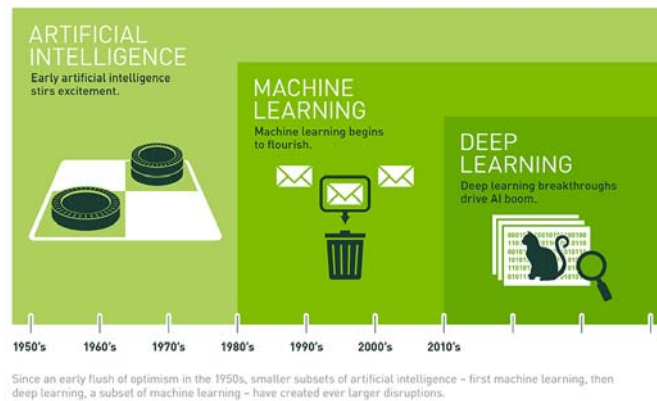
Big Data und Machine Learning

13. Februar 2018

Begriffe

- Artificial Intelligence
- Machine Learning
- Big Data
-

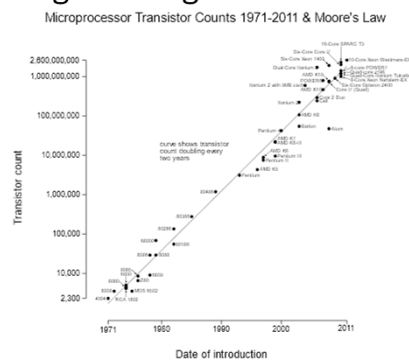
Entwicklung



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

Entwicklung: Moores Law

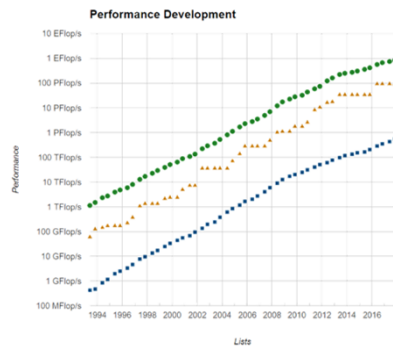
- Transistorendichte in einem Schaltkreis verdoppelt sich in gleichbleibenden und regelmäßigen Abständen.



Quelle: https://de.wikipedia.org/wiki/Mooresches_Gesetz

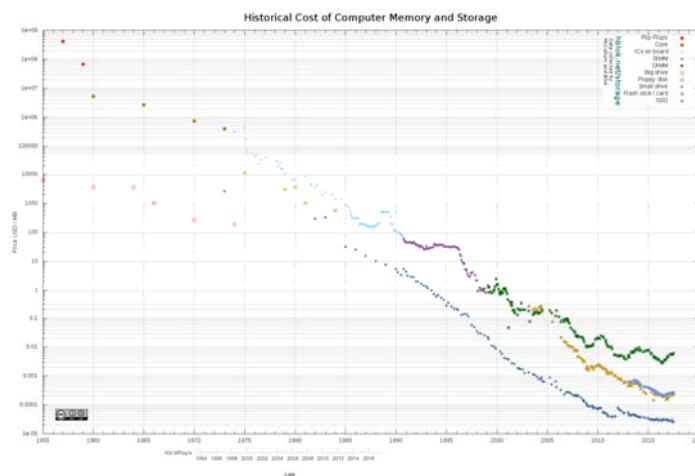
Entwicklung: Rechengeschwindigkeit

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



Quelle: <https://www.top500.org/statistics/perfdevel/>

Entwicklung: Speicherpreise



Quelle: https://hblol.net/storage_data/storage_memory_prices_2017-12_large.png

Exkurs: Logarithmische Skalen und Wachstum

Slaughterbots

- NZZ am Sonntag, 3. Dezember 2017



Mag ich, mag ich nicht: OCEAN

- Spektrum, 12. Januar 2015, PNAS 27. Januar 2015

COMPUTER ERRECHNET UNSERE PERSÖNLICHKEIT

Was Facebook-Likes über uns verraten

Zeig mir deine Facebook-Likes ... und ich sag dir, wie du bist: Schon wenige "Likes" auf Facebook genügen, um die Persönl engsten Freunde es können.

<http://www.spektrum.de/news/was-facebook-likes-ueber-uns-verraten/1326951>

von Jan Dönges



© Stock / iStockphoto

Die Persönlichkeit eines Menschen lässt sich erstaunlich präzise mit Hilfe eines Computers bestimmen – zumindest wenn derjenige es Youyou von der University of Cambridge und Kollegen haben eine Software darauf programmiert, aus den "Gefällt-mir-Angaben" eine überraschend gut: Ihr Programm erreichte höhere Trefferquoten als Arbeitskollegen, Freunde oder sogar Familienangehörige des Jew

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.

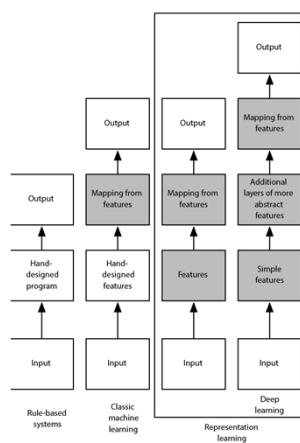
X_i ist dabei ein Vektor in \mathbb{R}^n , Y_i nimmt Werte in \mathbb{N}_0 an (z.B. 0 und 1 oder, 1,2,3,4,5, etc.)

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: Zeichenerkennung

- Bekannt: Einzelne, handgeschriebene Zahlen von US-Postleitzahlen als Bild-Datei
- Gesucht: Zahlwert der Bild-Datei zwecks automatischer Sortierung

0	1	2	3	4	7	8	9	0	1	2	3	4	5	6	7	8	6	
2	6	4	7	5	5	4	7	8	9	2	9	3	9	3	8	2	0	5
1	0	4	2	6	5	3	5	3	8	0	0	3	4	1	5	3	0	8
0	6	2	7	1	1	8	1	7	1	3	8	9	7	6	7	4	1	6
5	1	7	1	9	8	0	6	9	4	9	9	3	7	1	9	2	2	5
8	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	0	
3	4	5	6	7	8	9	8	1	0	5	5	1	9	0	4	1	9	
4	7	7	8	5	0	6	5	5	3	3	3	9	8	1	4	0	6	
0	6	2	1	1	3	2	8	8	7	8	4	6	0	2	0	3	6	
5	9	9	3	2	4	9	4	6	5	3	2	8	5	9	4	1		
1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7		
1	2	3	4	5	6	7	8	9	6	4	2	6	4	7	5	5		
2	9	3	9	3	8	2	0	9	8	0	5	6	0	1	0			
5	5	4	3	4	1	5	3	0	8	3	0	6	2	7	1			
3	8	5	4	2	0	9	7	6	7	4	1	6	8	4				
1	9	8	0	6	9	4	9	9	6	2	3	7	1					
0	1	2	3	4	5	6	7	8	0	1	2	3						
2	3	4	5	6	7	8	9	2	1	2	1	3						
5	7	9	9	4	7	0	3	4	1	4								
6	6	4	6	3	5	7	2	5	9									

Lernalgorithmen: Klassifizierung von Punkten

- Bekannt: Ordinate und Abszisse und Farbe von 1000 Punkten
- Gesucht: Farbe von neu dazukommenden Punkten, bei denen nur Ordinate und Abszisse gegeben sind.

