



UNIKLINIKUM
SALZBURG
LANDESKRANKENHAUS

KI – was ist das eigentlich?

Marco Meinschad

Definitionen

- **Künstliche Intelligenz (KI)**

aus Wikipedia:

Künstliche Intelligenz ist ein Teilgebiet der Informatik, es umfasst alle Anstrengungen, deren Ziel es ist, Maschinen intelligent zu machen. Dabei wird Intelligenz verstanden als die Eigenschaft, die ein Wesen befähigt, angemessen und vorausschauend in seiner Umgebung zu agieren; dazu gehört die Fähigkeit, Sinneseindrücke wahrzunehmen und darauf zu reagieren, Informationen aufzunehmen, zu verarbeiten und als Wissen zu speichern, Sprache zu verstehen und zu erzeugen, Probleme zu lösen und Ziele zu erreichen.

Jede Methode/Algorithmus, die menschliche Intelligenz nachahmt (oder ihr entspricht)

- **Machine Learning (ML)**

Algorithmen für spezielle Aufgaben, die direkt aus Daten („Erfahrung“) lernen, ohne explizite Programmierung

- **Starke KI – schwache KI**

Das „Leben“ einer KI

- Erstellen einer Architektur
- Training
- Anwenden des Gelernten auf unbekannte Daten

→ Ziel ist die Generalisierung

Das „Leben“ einer KI

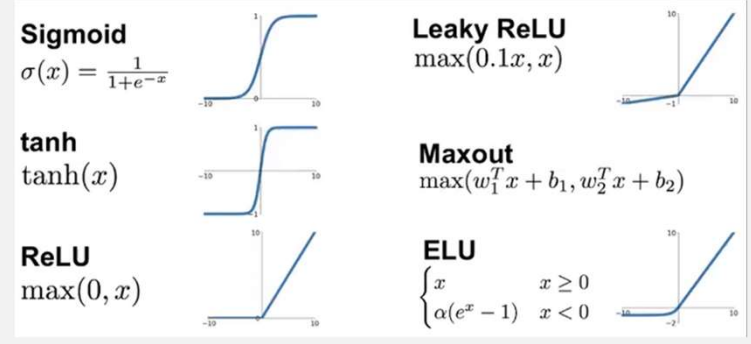
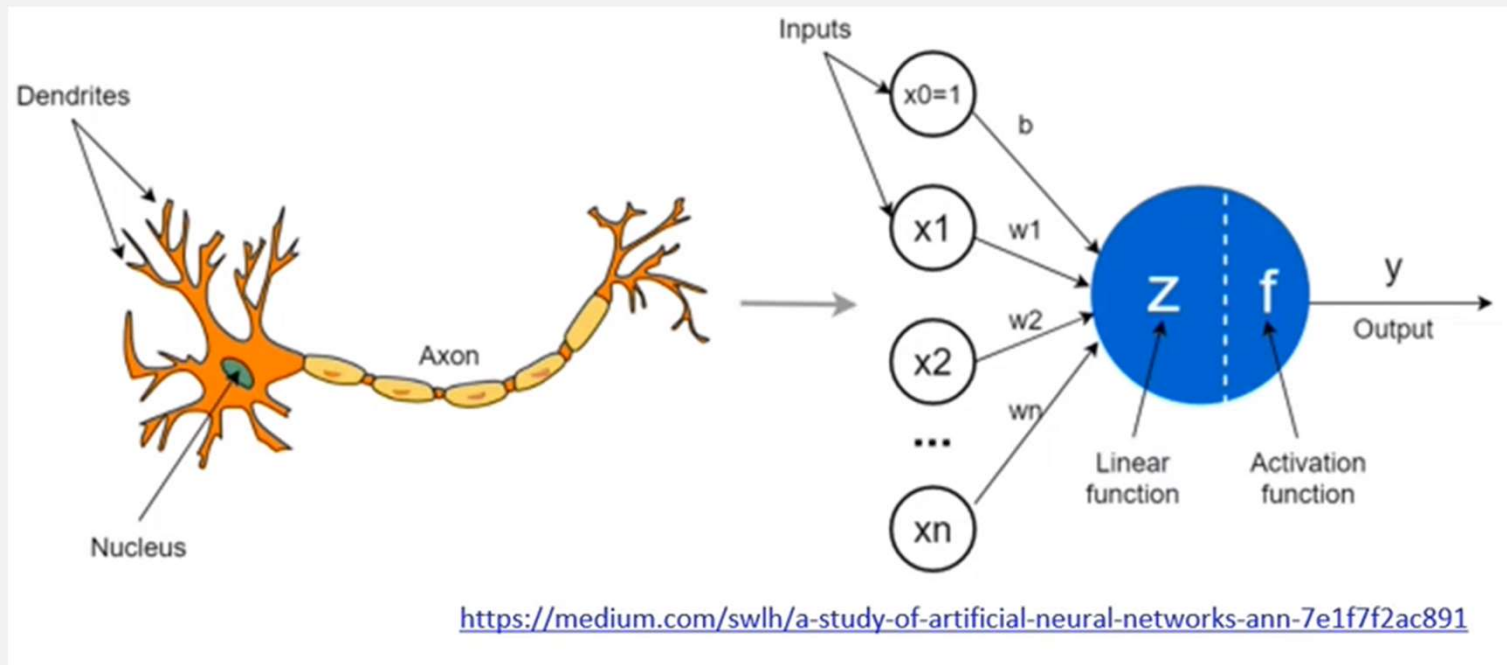
- Erstellen einer Architektur
- Training
- Anwenden des Gelernten auf unbekannte Daten

→ Ziel ist die Generalisierung



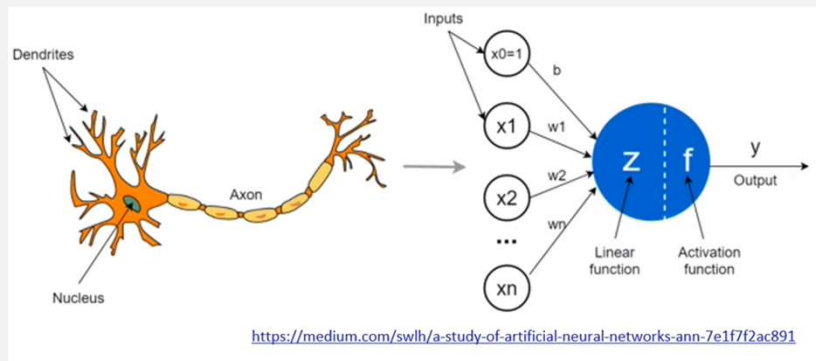
Aufbau eines Neuronales Netzes (NN)

Künstliches Neuron

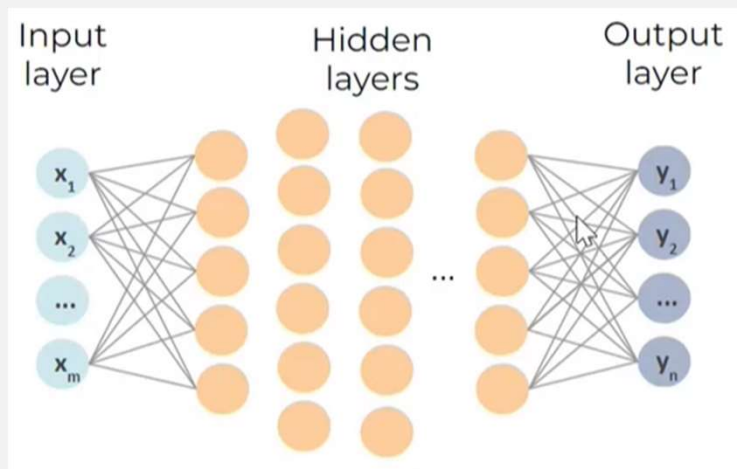


Aufbau eines Neuronales Netzes (NN)

Künstliches Neuron



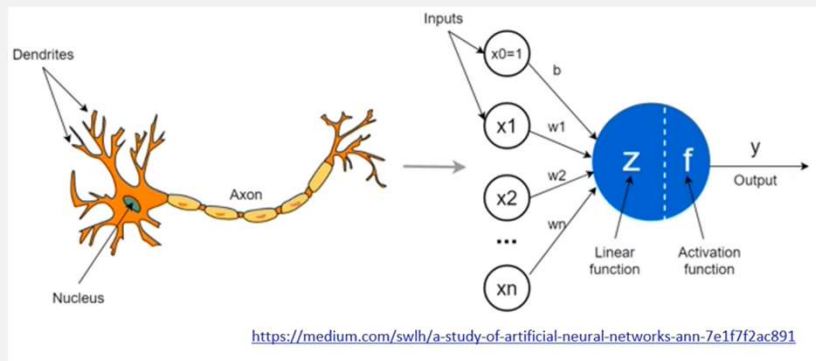
(tiefes) neuronales Netz



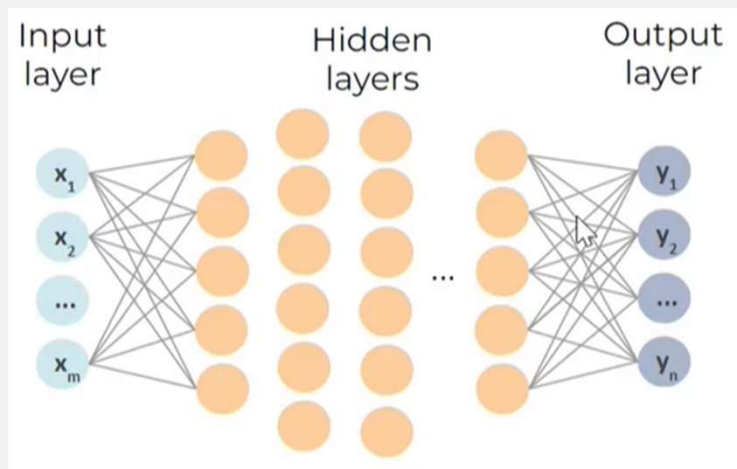
Expressivität bzw. Repräsentationskraft eines NN
 → kann eine Funktion beliebig genau approximieren

Aufbau eines Neuronales Netzes (NN)

Künstliches Neuron



(tiefes) neuronales Netz



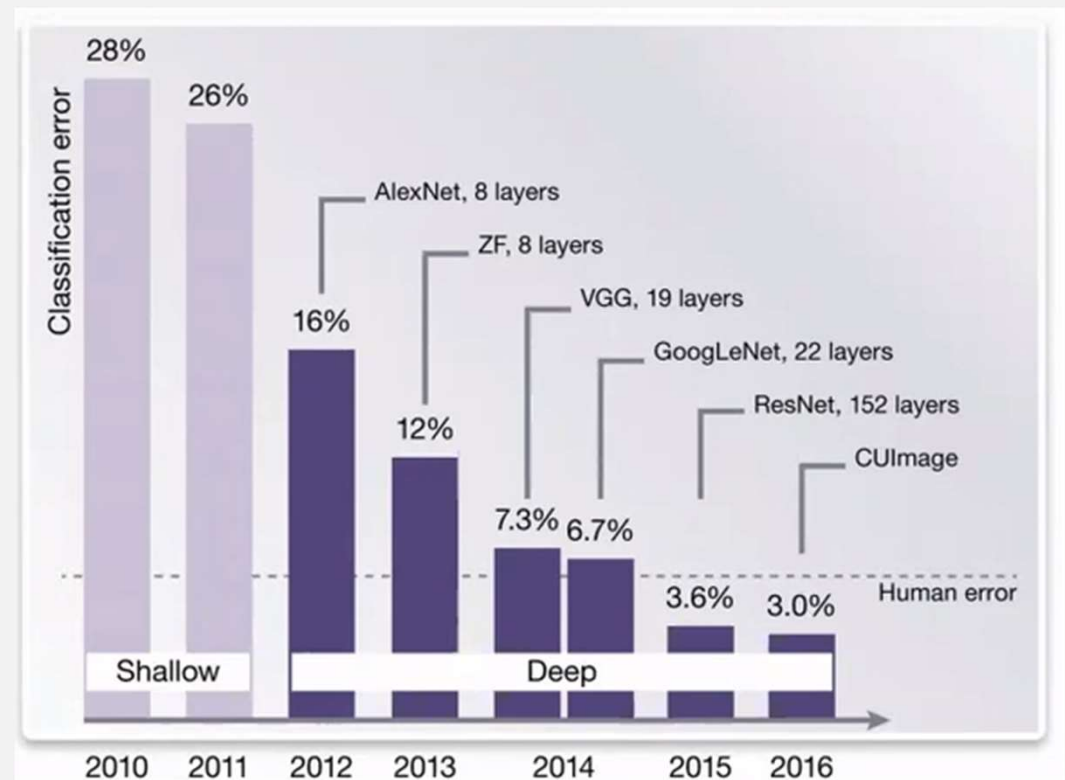
Expressivität bzw. Repräsentationskraft eines NN
 → kann eine Funktion beliebig genau approximieren

- Enges + tiefes NN
 → bessere Generalisierung als flach und breit

ILSVRC – ImageNet Large Scale Visual Recognition Challenge (ab 2010)

- Wettbewerb von KIs, bzw. Convolutional Neural Networks (CNNs)
 - 1,2 Mio. Bilder
(1 Mio. zum Training, 200k zum Testen)
 - 1000 Kategorien
- LeNet-5 (1995)
~ 62000 Parameter
- Alexnet (2012)
~ 60 Mio Parameter
- Inception (= GoogLeNet) (2014)
~ 24 Mio Parameter

Expressivität und
Generalisierung eines NN

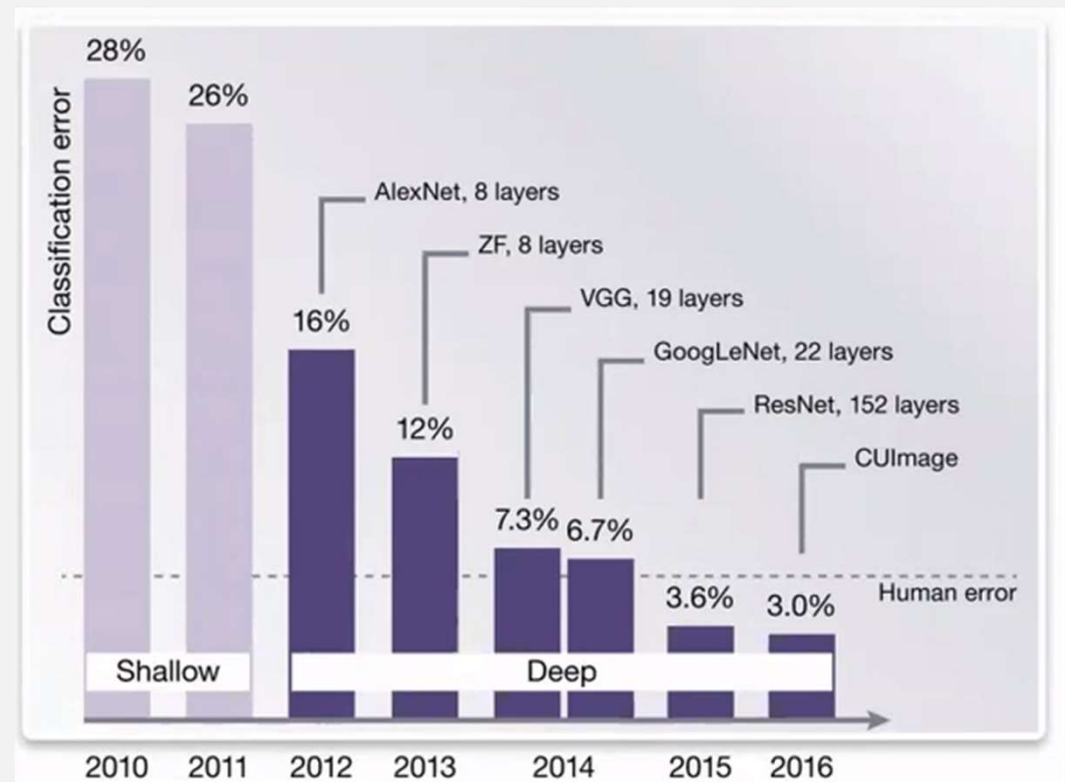


ILSVRC – ImageNet Large Scale Visual Recognition Challenge (ab 2010)

- Wettbewerb von KIs, bzw. Convolutional Neural Networks (CNNs)
 - 1,2 Mio. Bilder
(1 Mio. zum Training, 200k zum Testen)
 - 1000 Kategorien

- LeNet-5 (1995)
~ 62000 Parameter
- Alexnet (2012)
~ 60 Mio Parameter
- Inception (= GoogLeNet) (2014)
~ 24 Mio Parameter

- ChatGPT 3 (2020)
~ 175 Mrd Parameter



Trainings-Paradigmen

Supervised learning



(e.g. like having a teacher that shows you examples of problems for which s.he already tells you the answer)

Unsupervised learning



(e.g. like giving you a bunch of data and let you explore it and extract useful information by yourself)

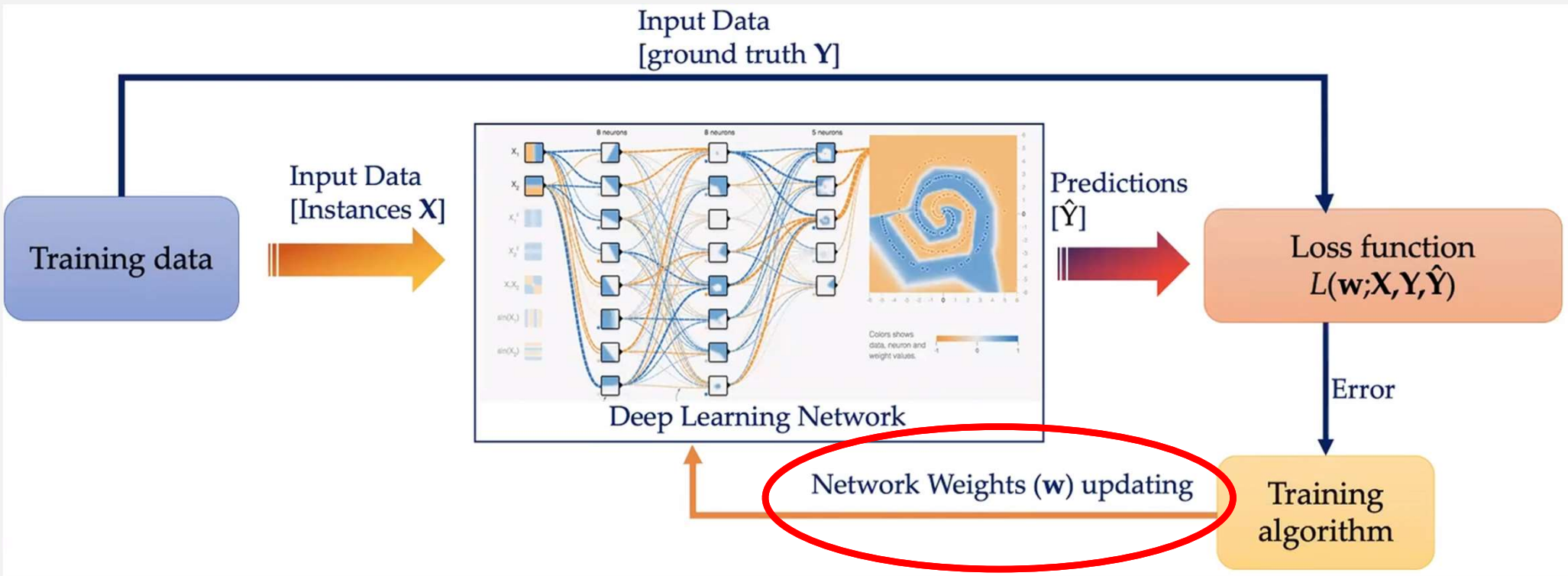
Reinforcement learning



(e.g. like a kid that interacts with the environment and learns about the consequences to his/her actions)

Quelle: Vortrag von Ana M. Barragan-Montero, UCLouvain, Brüssel

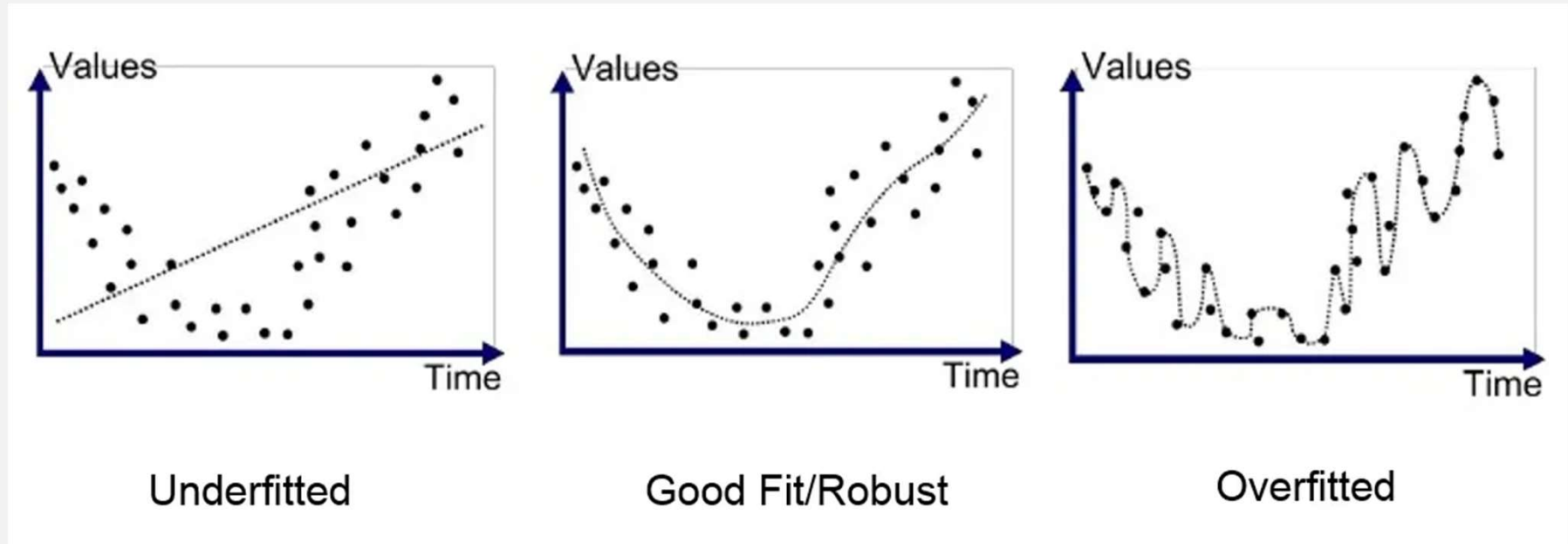
Training einer KI



Good results come from good data!

Quelle:
Vortrag von Andrea Barucci - Nello Carrara Institute of Applied
Physics, Florence, Italy

under- and overfitting



- underfitted
model too simple
large training error
- overfitted
model too complex
low training error
large generalization error

Quelle: <https://medium.com/greyatom/what-is-underfitting-and-overfitting-in-machine-learning-and-how-to-deal-with-it-6803a989c76>

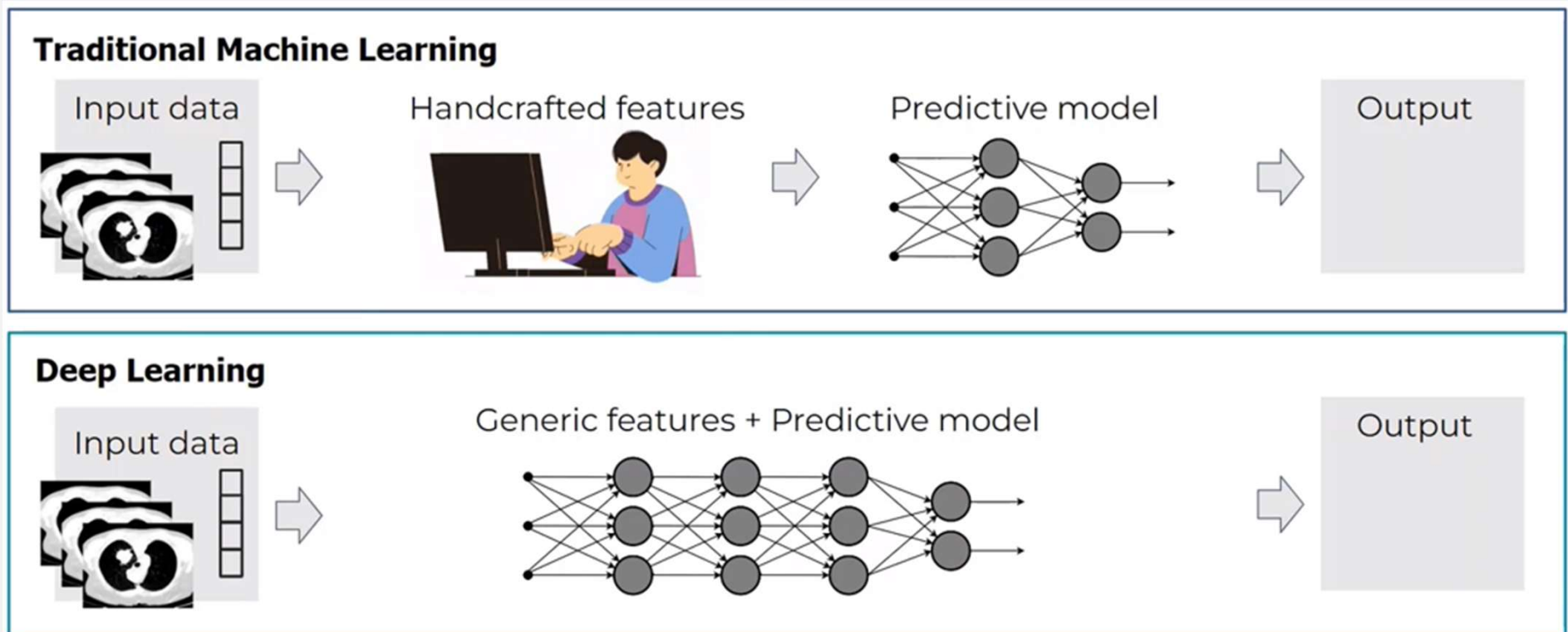
Feature Extraction

Beispiel: Kantenerkennung



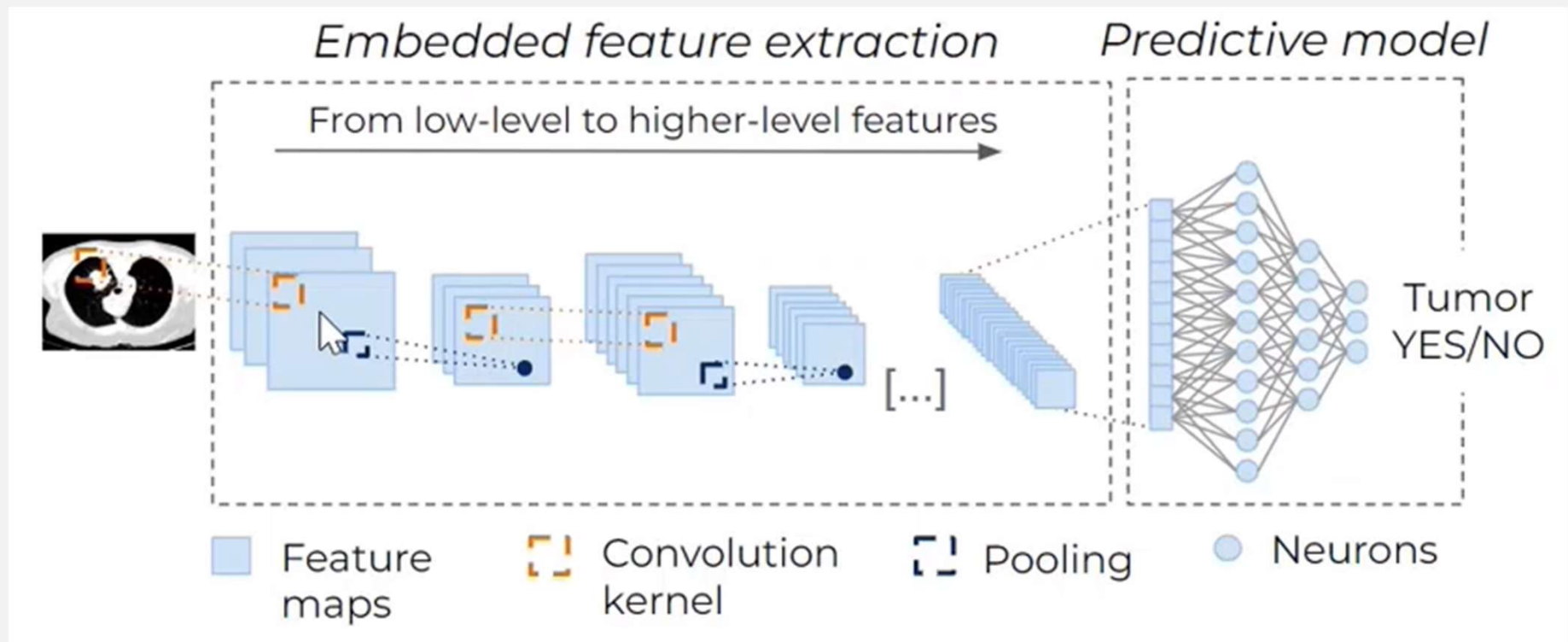
https://en.wikipedia.org/wiki/Edge_detection

Machine Learning vs. Deep Learning workflows





Quelle: Vortrag von Ana M. Barragan-Montero, UCLouvain, Brüssel

Convolutional Neural Network (CNN)



Quelle:
 Barragán-Montero, Ana, et al. "Artificial intelligence and machine learning for medical imaging: A technology review." *Physica Medica* 83 (2021): 242-256.

Confounding features

Original Image	Noised Image
Lifeboat: 89.20%, Scotch Terrier: 0.00%	Lifeboat: 0.03%, Scotch Terrier: 99.77%
	
Lifeboat (89.2%) → Scotch Terrier (99.8%)	
Danny Karmon, Daniel Zoran, Yoav Goldberg - LaVAN: Localized and Visible Adversarial Noise (2018)	
Janelle Shane - Aiweirdness.com	



Confounding features

PLOS MEDICINE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

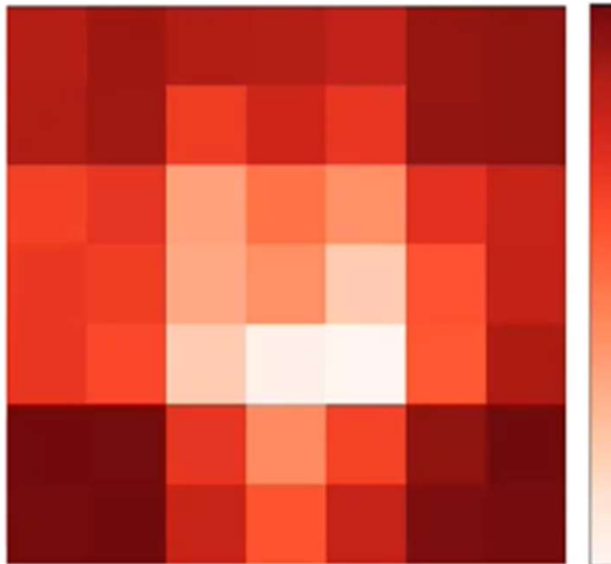
Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study

John R. Zech , Marcus A. Badgeley , Manway Liu, Anthony B. Costa, Joseph J. Titano, Eric Karl Oermann 

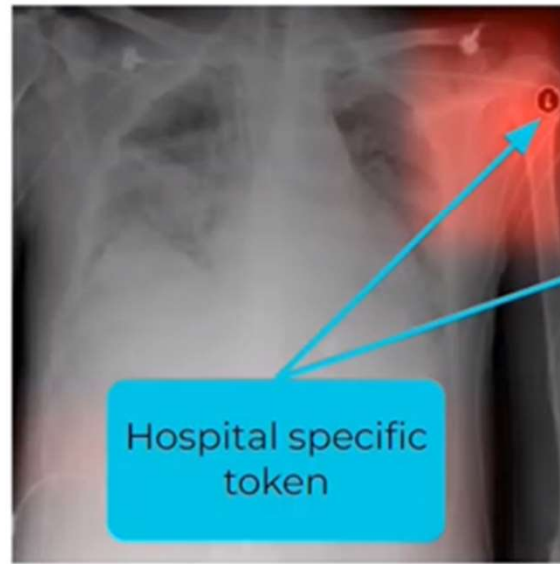
Published: November 6, 2018 • <https://doi.org/10.1371/journal.pmed.1002683>

Class Activation Maps (CAM)

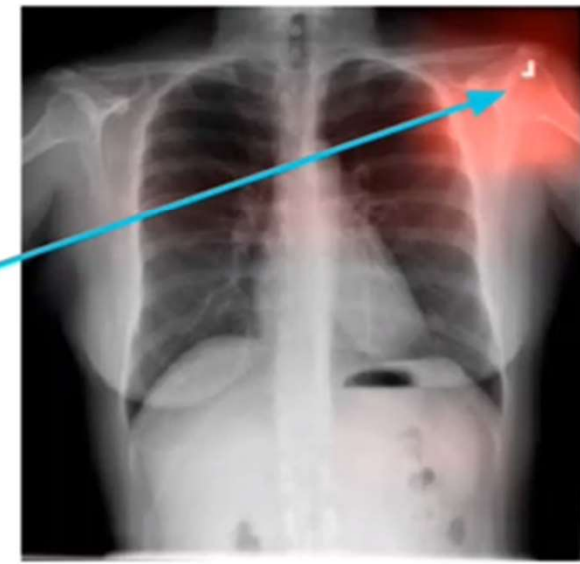
Average CAM
(over several patients)



CAM patient 1
(overlay with X-ray image)



CAM patient 2
(overlay with X-ray image)



- 3 Zentren, 158k Thoraxröntgen
Pneumonie-Prävalenz **34,2%** / 1,2% / 1,0%

Weitere wichtige KI Typen

- Generative Adversarial Networks (GANs)
z.B. Midjourney, Stable Diffusion, etc.



- Large Language Models (LLMs), bzw. Generative Pretrained Transformer (GPT)
z.B. ChatGPT (OpenAI), LaMDA (= Bard (Google)), LLaMA (Meta), etc.

Quelle:
https://en.wikipedia.org/wiki/Generative_adversarial_network

Zusammenfassung KIs derzeit

- Spezialisierte Systeme für dezidierte Aufgaben
- Keine explizite Programmierung
Generalisierung auf Grundlage von Trainingsdaten
- Qualität der Trainingsdaten ist entscheidend
- Keine „starke“ KI, keine „Intelligenz“

Vielen Dank 😊

Marco Meinschad
m.meinschad@salk.at
+43 5 7255-57937