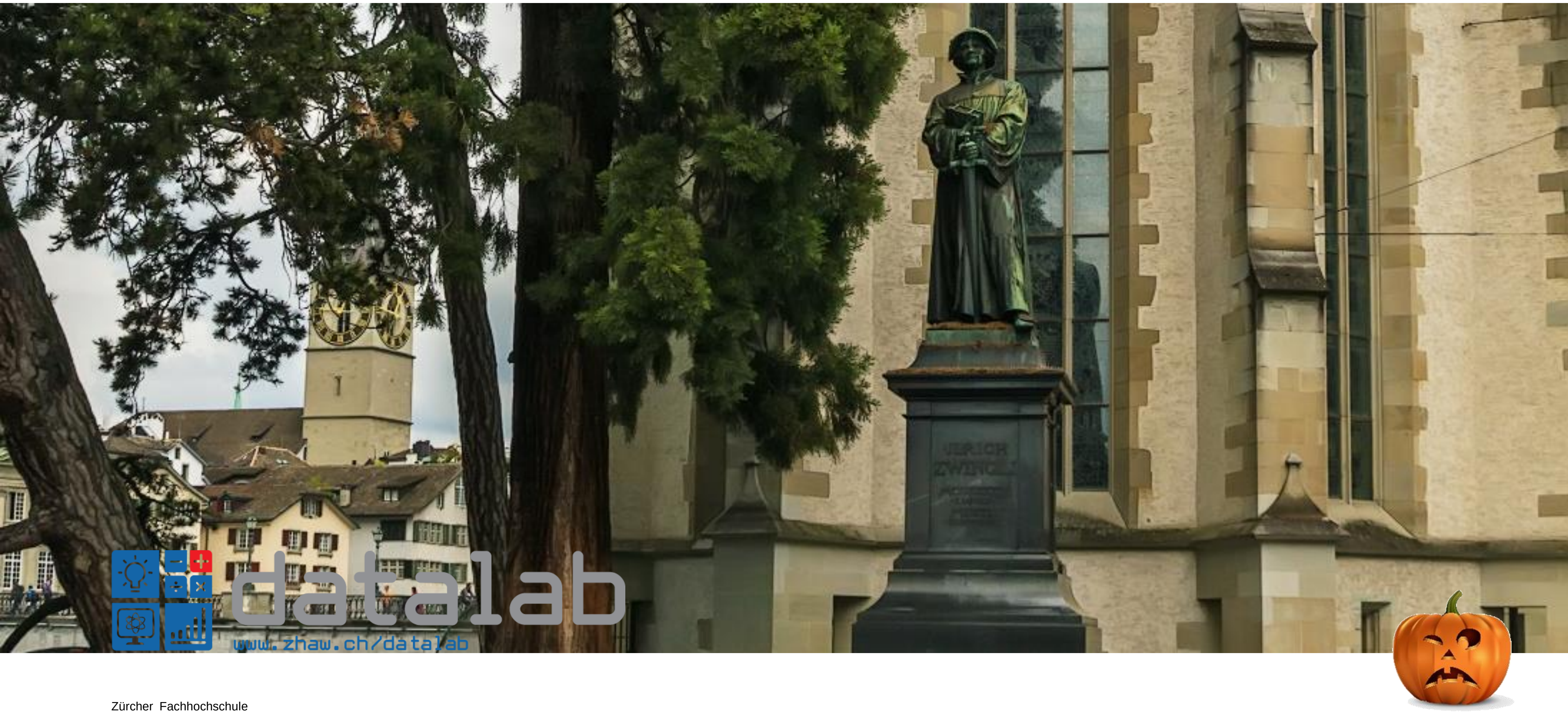
Künstliche Intelligenz – was, wo & wohin?

Volkshochschule Winterthur, 31. Oktober 2019

Thilo Stadelmann

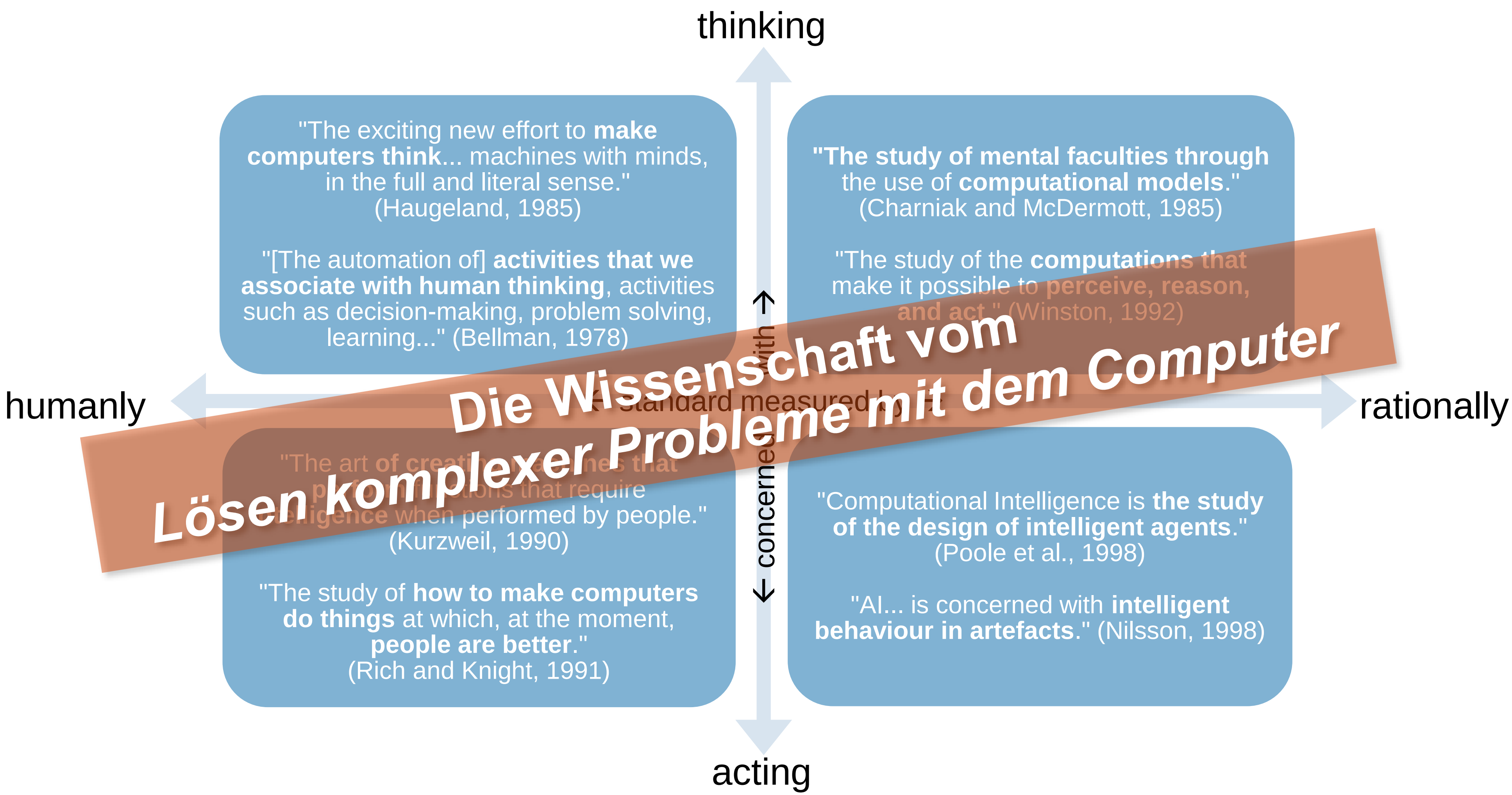


Zürcher Hochschule
für Angewandte Wissenschaften

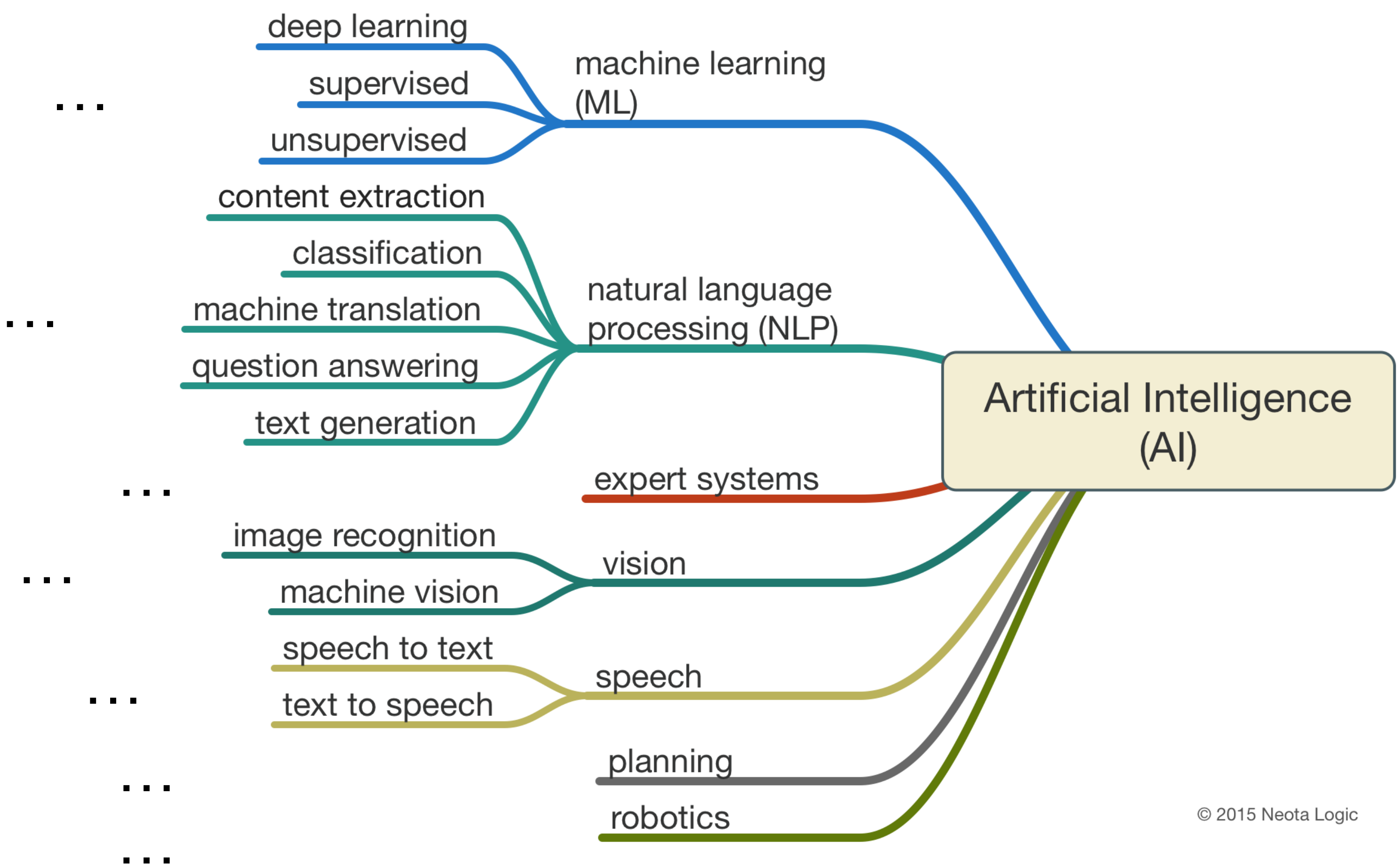
1

Was ist Künstliche Intelligenz?

Was ist künstliche Intelligenz?



Was gehört zu künstlicher Intelligenz?

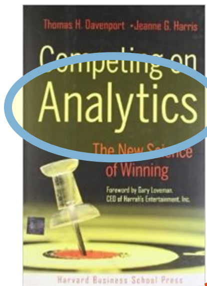


KI im Kontext

Turing's "Computing Machinery and Intelligence"
Dartmouth meeting: "Artificial Intelligence" adopted
Neural networks almost disappear: 1st "**AI Winter**"
Expert systems industry **booms**
Neural networks return to popularity
Expert systems industry **busts**: 2nd "AI Winter"
Human-Level AI back on agenda

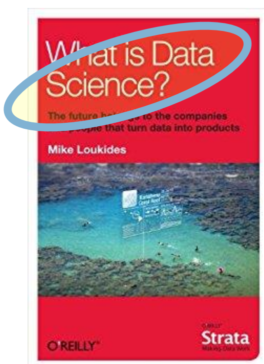
1950 1960 1970 1980 1990 2000

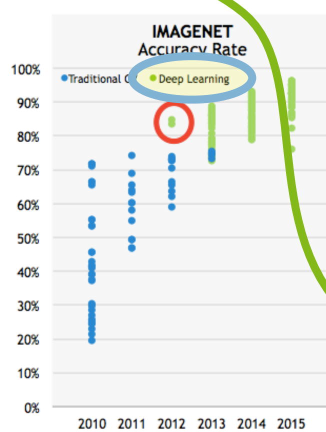
Zürcher Hochschule für Angewandte Wissenschaften

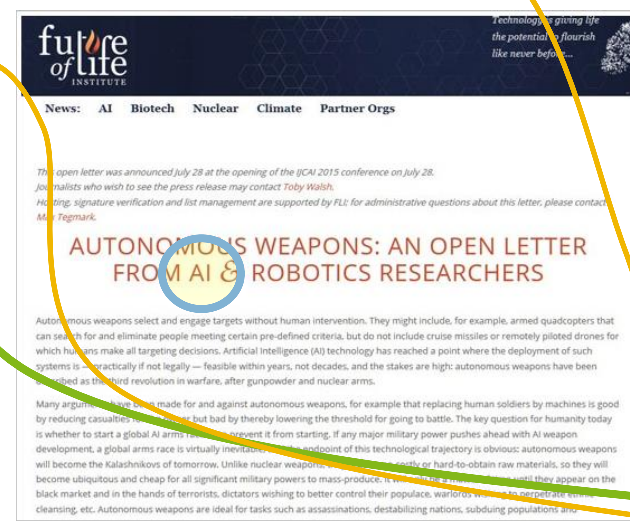


...

2007 2012 2016

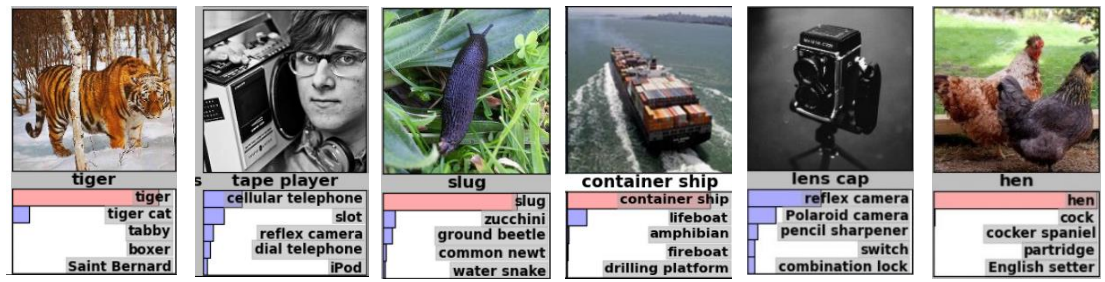




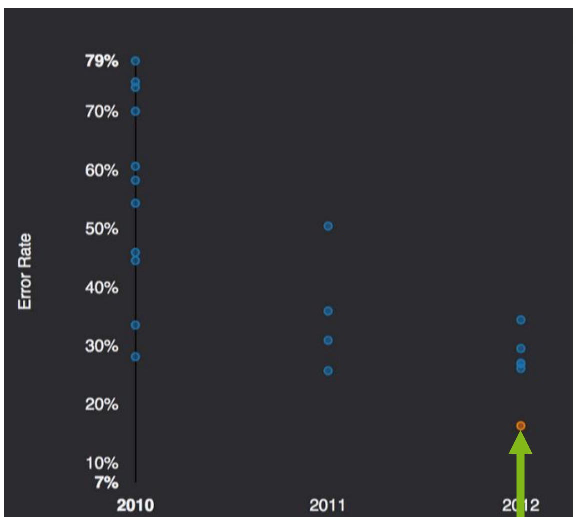


Was ist passiert?
Der ImageNet Wettbewerb

1000 Kategorien
1 Mio. Beispiele



…



A. Krizhevsky verwendet als erster ein sog. «Deep Neural Network» (CNN)

2015: Computer *haben "Sehen" gelernt*

4.95% Microsoft (06. Februar)
→ Besser als Menschen (5.10%)

4.80% Google (11. Februar)

4.58% Baidu (11. Mai)

3.57% Microsoft (10. Dezember)

Idee: Mehr «Tiefe» um Merkmale automatisch zu lernen

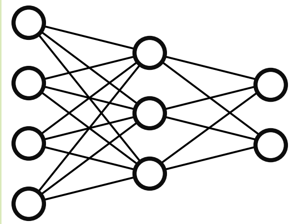
Classical image processing

Feature extraction (SIFT, SURF, LBP, HOG, etc.)

(0.2, 0.4, …)

(0.4, 0.3, …)

Classification (SVM, neural network, etc.)

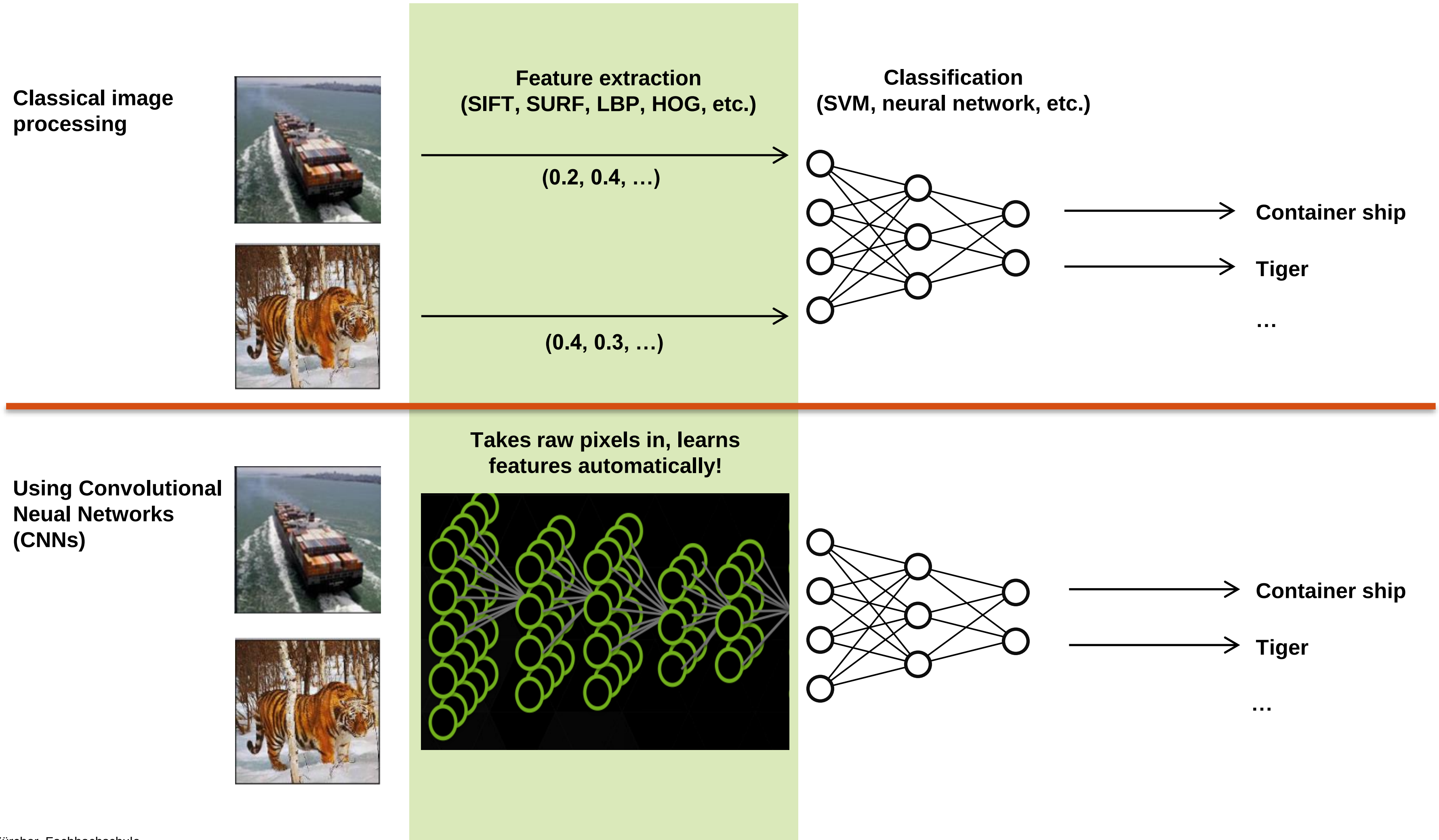


Container ship

Tiger

…

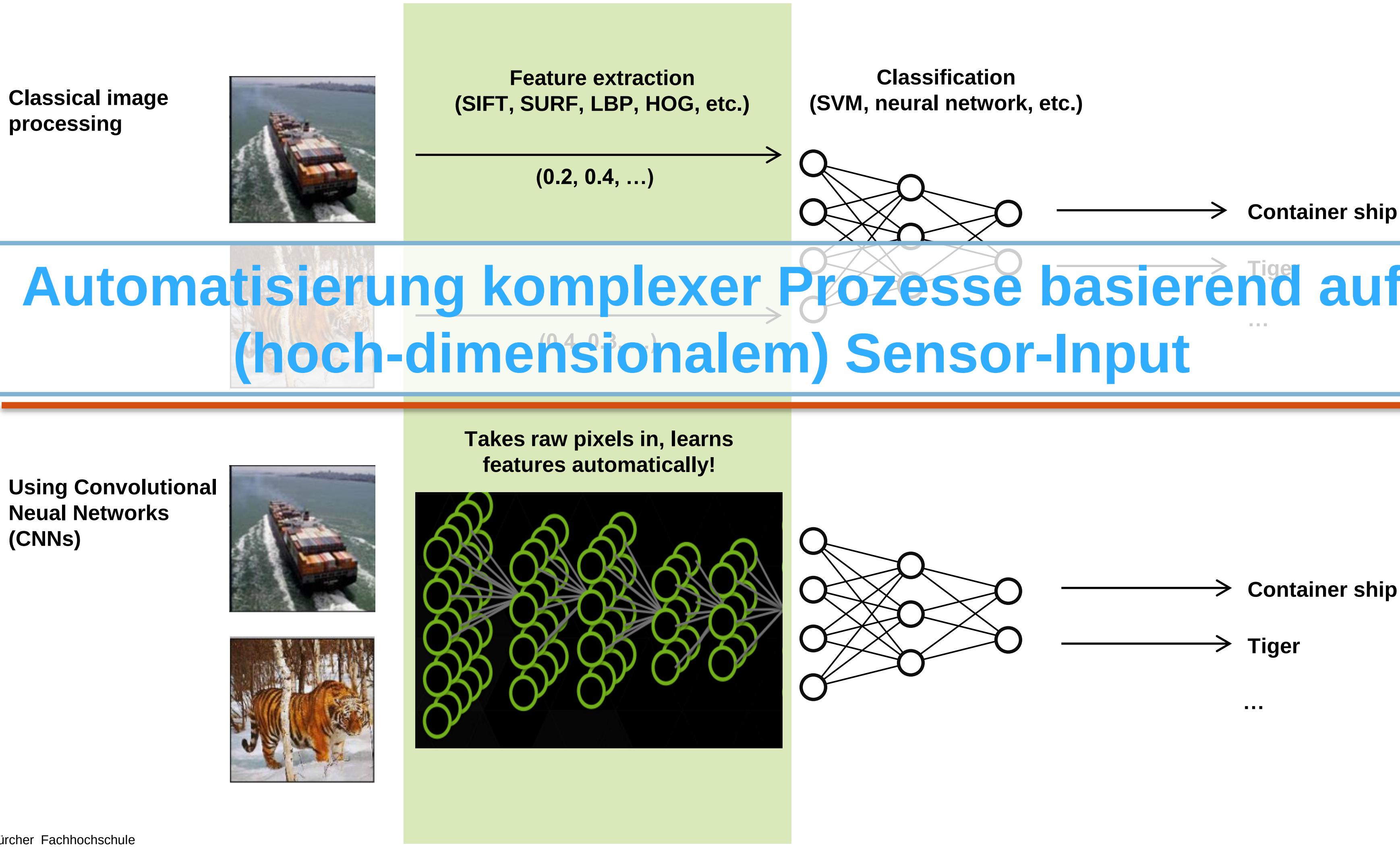
Idee: Mehr «Tiefe» um Merkmale automatisch zu lernen



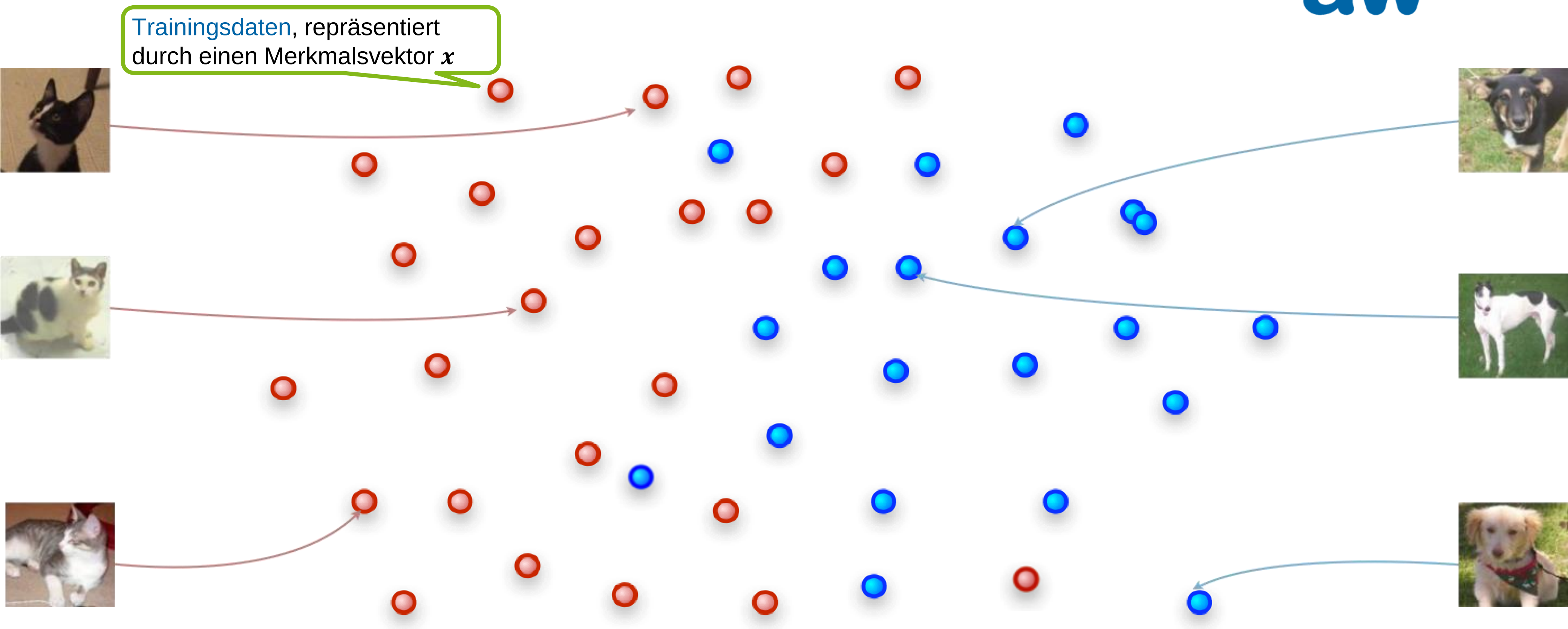
Zürcher Hochschule
für Angewandte Wissenschaften

Idee: Mehr «Tiefe» um Merkmale automatisch zu lernen

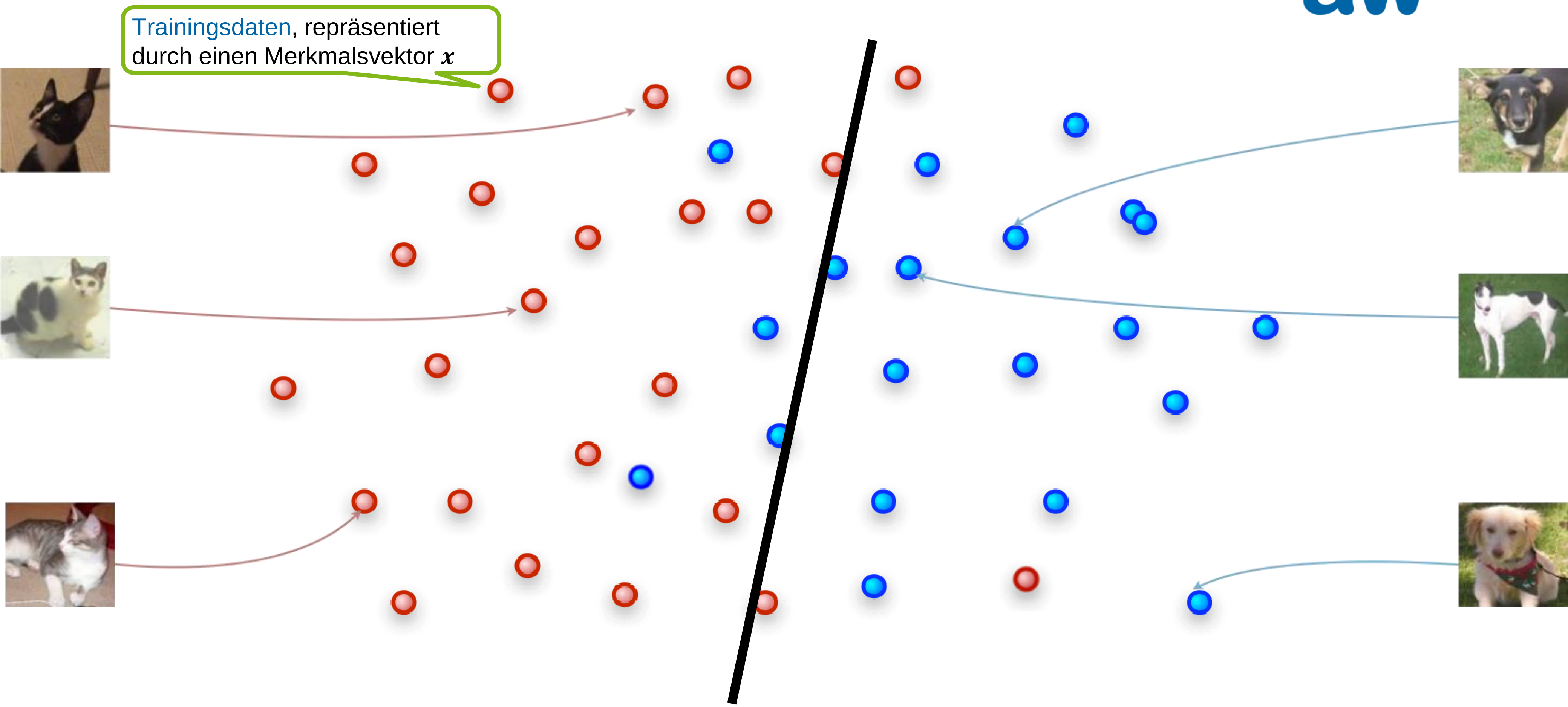
Automatisierung komplexer Prozesse basierend auf (hoch-dimensionalem) Sensor-Input



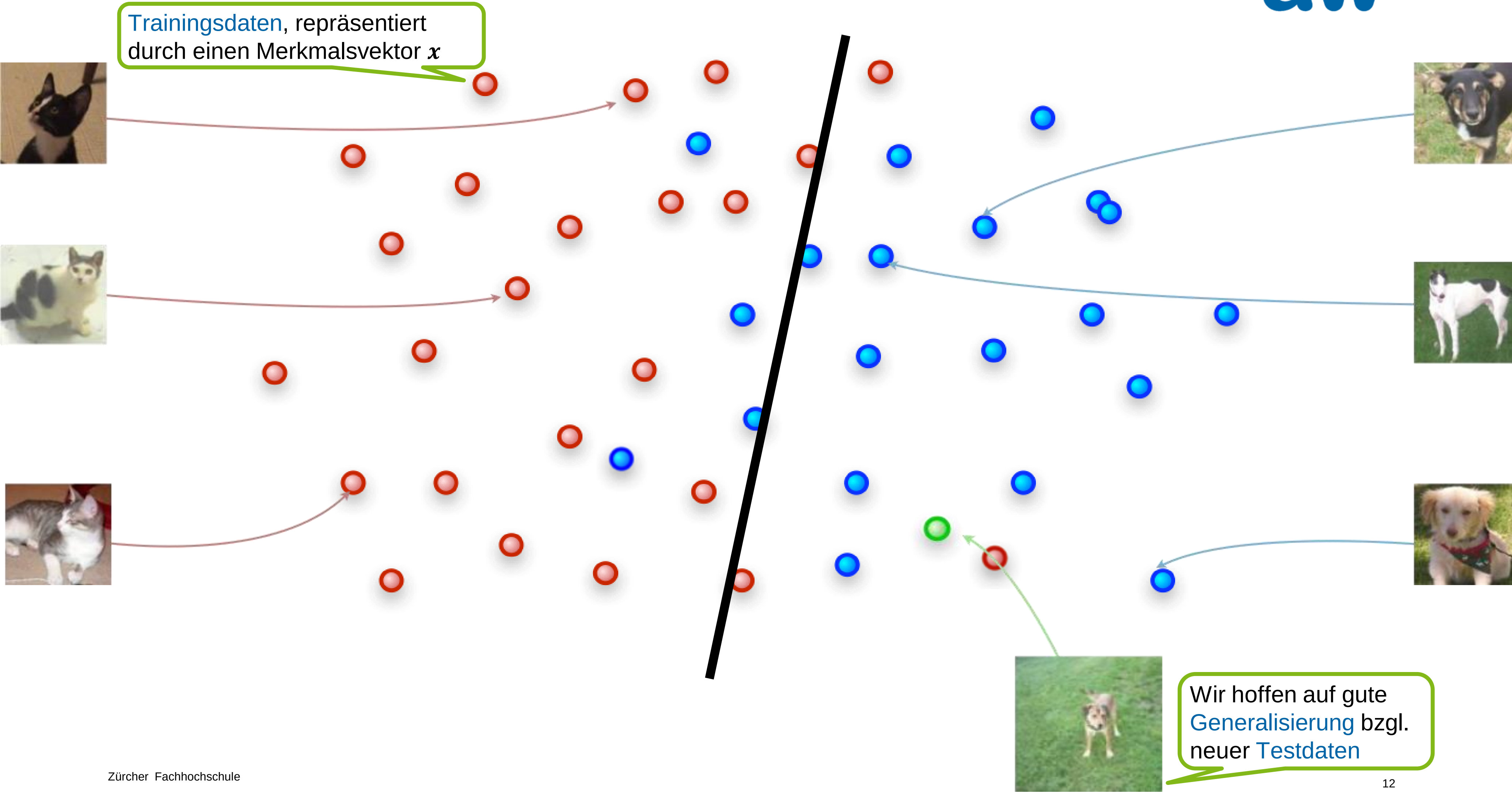
Supervised Machine Learning im Überblick



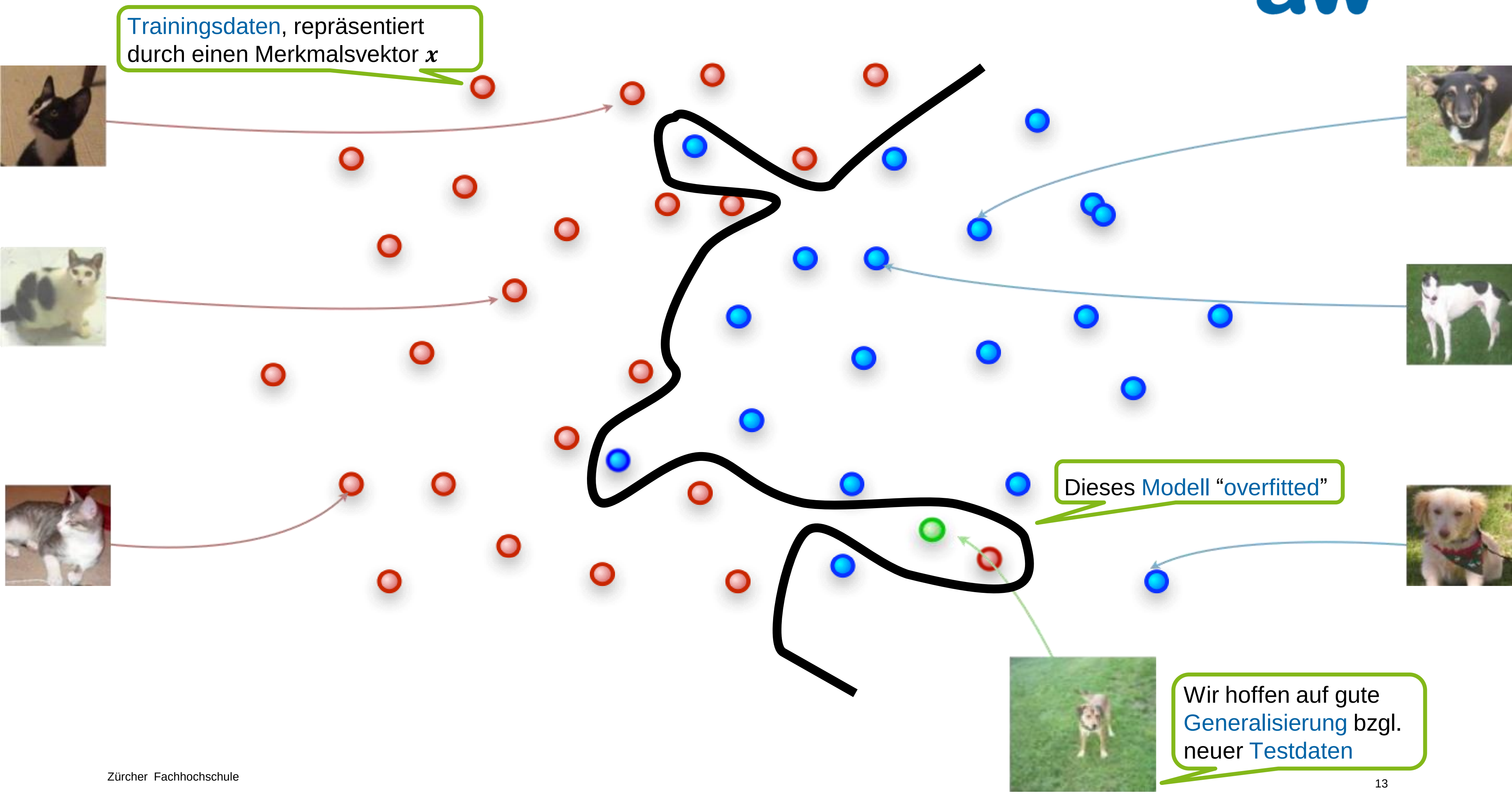
Supervised Machine Learning im Überblick



Supervised Machine Learning im Überblick



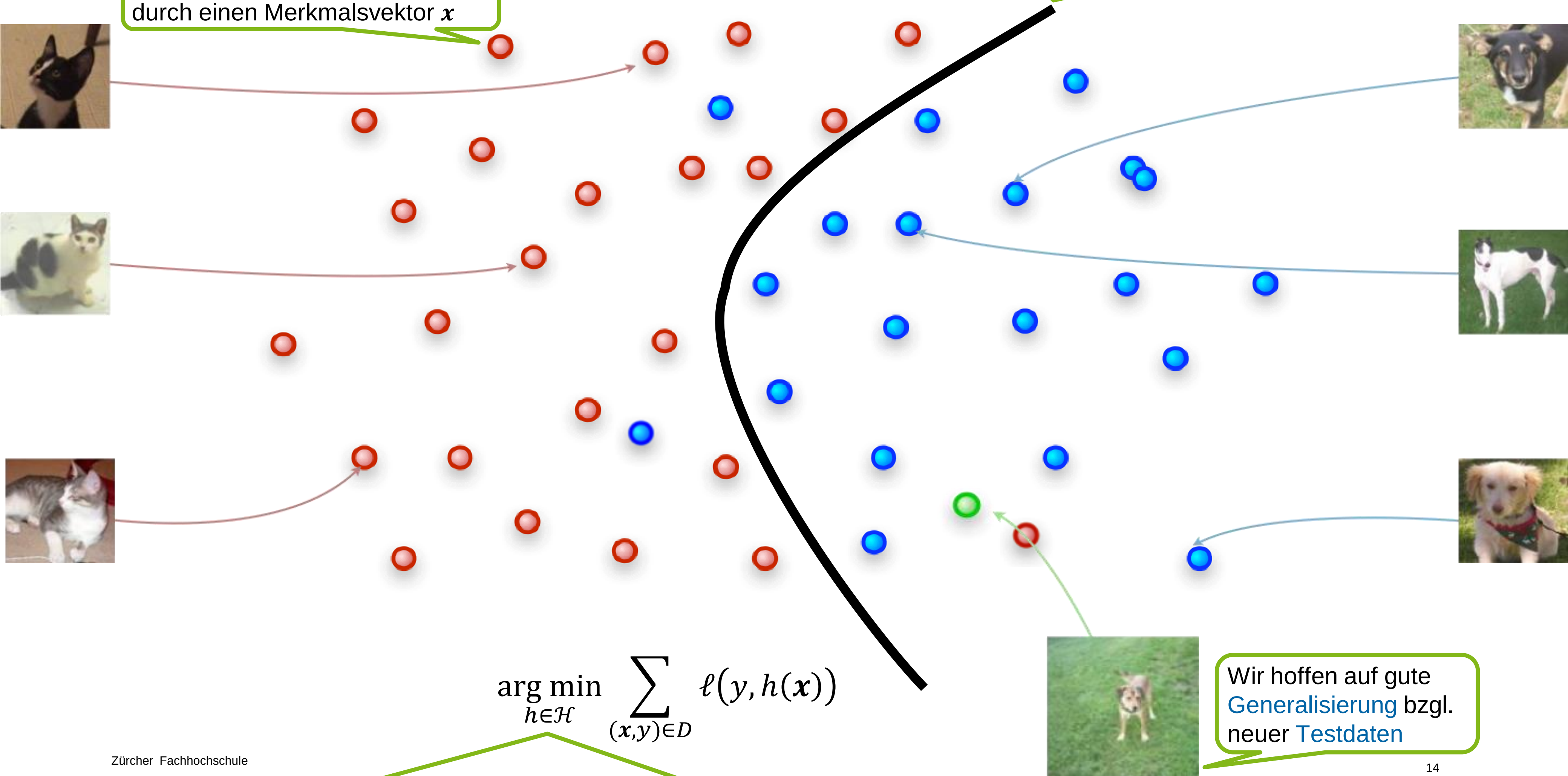
Supervised Machine Learning im Überblick



Supervised Machine Learning im Überblick

Trainingsdaten, repräsentiert durch einen Merkmalsvektor $\boldsymbol{x}$

Dieses Modell scheint gut zu passen



Wir hoffen auf gute Generalisierung bzgl. neuer Testdaten

$$\underset{h \in \mathcal{H}}{\arg\min} \sum_{(\boldsymbol{x},y)\in D} \ell\big(y, h(\boldsymbol{x})\big)$$

Wir suchen gute Modelle in einem Hypothesenraum $\mathcal{H}$ durch Minimierung des Loss ℓ zwischen Label y und Resultat $h(\boldsymbol{x})$

Was→ Wo?→ Wohin?

2

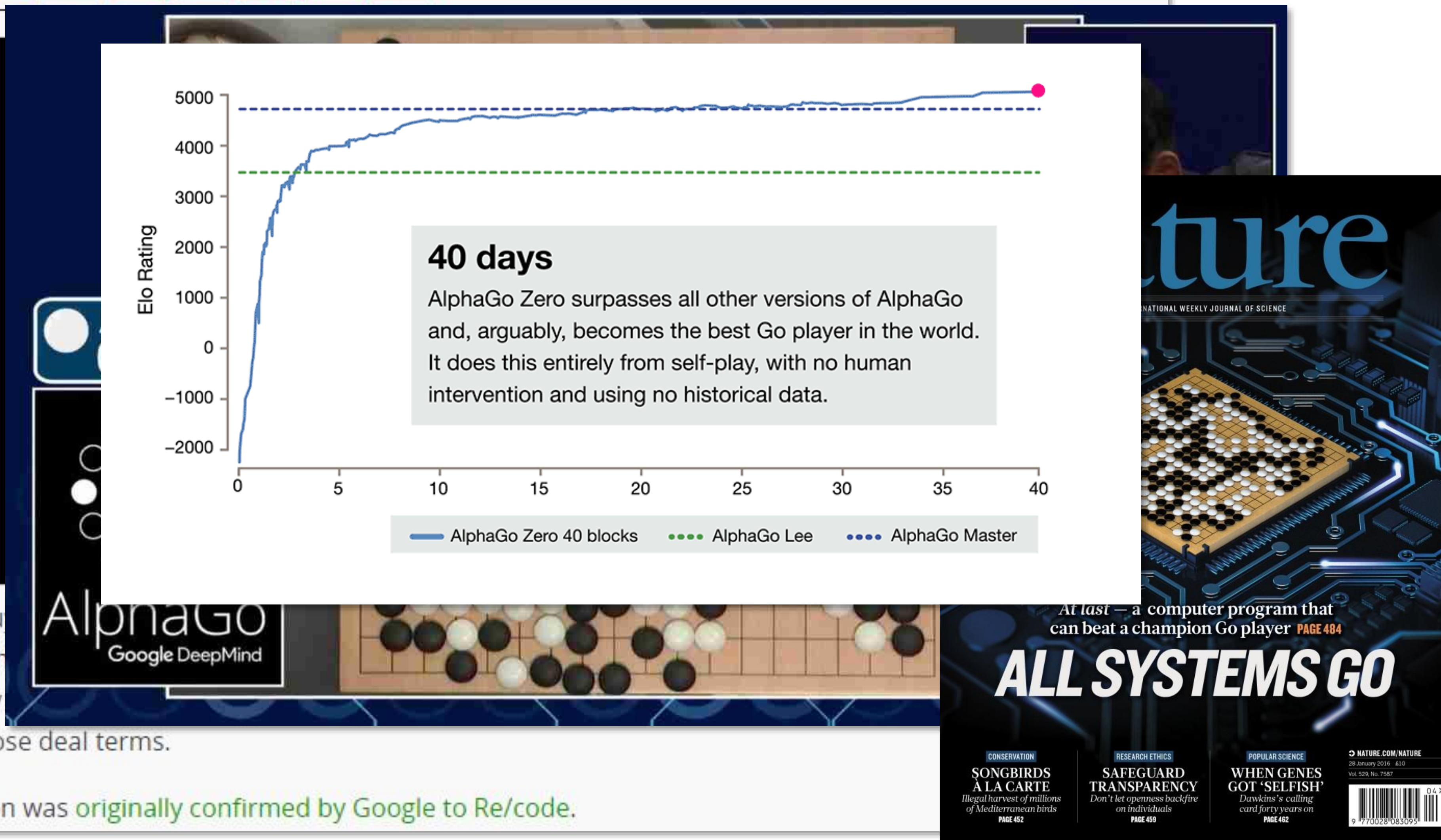
Wo wird das bereits eingesetzt?

Google Acquires Artificial Intelligence Startup DeepMind For More Than $500M

Posted Jan 26, 2014 by Catherine Shu (@catherineshu)

Zürcher Hochschule für Angewandte Wissenschaften

zh aw



Google will bu
reports that th
in talks to buy
couldn't disclose deal terms.

The acquisition was originally confirmed by Google to Re/code.

MIT Techn
Revie

Computing

Algorith
Artistic
Other In

A deep neural n
other images.

by Emerging Tech

The nature of arti
of Vincent Van G
Edvard Munch's
humans recogni

Deep neural networks can now transfer the style of one photo onto another

And the results are impressive

by James Vincent | @jjvincent | Mar 30, 2017, 1:53pm EDT

SHARE TWEET LINKEDIN

Original photo Reference photo Result

You've probably heard of an AI technique known as "style transfer" — or, if you haven't heard of it, you've seen it. The process uses neural networks to apply the look and feel of one image to another, and appears in apps like Prisma and Facebook. These style transfers, however, are stylistic, not photorealistic. They look good because they look like they've been painted. Now a group of researchers from Cornell University and Adobe have augmented

NOW TRENDING

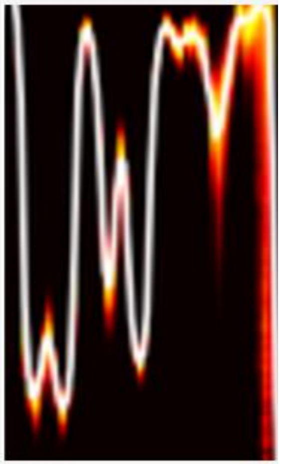
Zürcher Hochschule für Angewandte Wissenschaften

Künstliche Intelligenz

WaveNet läs[...]

von Henning Steier / 12.9.201[...]

Die Google-Tochter DeepM[...] macht auch Musik.

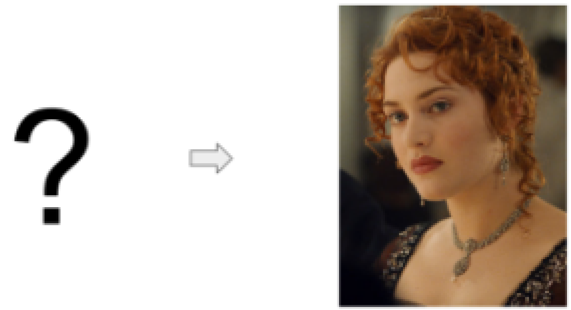


DeepMind lässt WaveNet Spra[...]

Die Google-Tochter Deep[...] Spiel «Go» Schlagzeilen: [...] einen der besten mensch[...] Londoner Unternehmen [...] erzeugt Sprache, die sehr [...] im Blogeintrag des Unter[...] Massstab nimmt. Man ha[...]

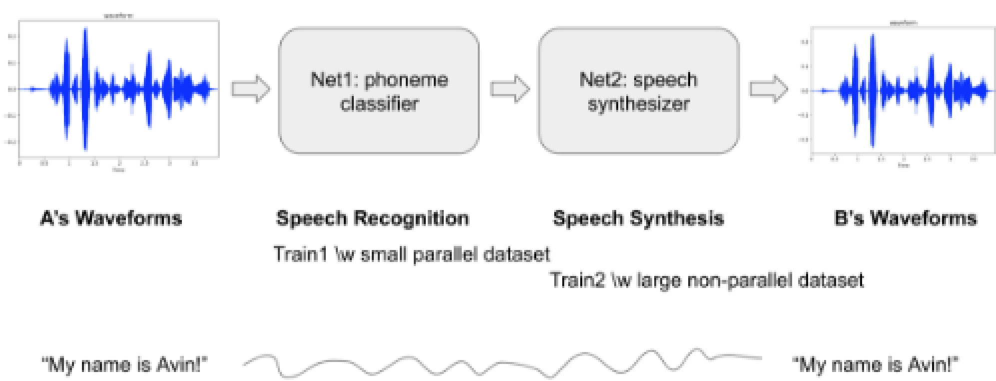
Intro

What if you could imitate a famous celebrity's voice or sing like a famous singer? This project started with a goal to convert someone's voice to a specific target voice. So called, it's voice style transfer. We worked on this project that aims to convert someone's voice to a famous English actress Kate Winslet's voice. We implemented a deep neural networks to achieve that and more than 2 hours of audio book sentences read by Kate Winslet are used as a dataset.



Model Architecture

This is a many-to-one voice conversion system. The main significance of this work is that we could generate a target speaker's utterances without parallel data like <source's wav, target's wav>, <wav, text> or <wav, phone>, but only waveforms of the target speaker. (To make these parallel datasets needs a lot of effort.) All we need in this project is a number of waveforms of the target speaker's utterances and only a small set of <wav, phone> pairs from a number of anonymous speakers.



[...]nerierte Sprache
[...]s Texteingabe»

[...]nerierte Musik
[...]ne Inhaltsvorgabe»

Zürcher Hochschule
für Angewandte Wissenschaften

zh
aw

…und die Liste liesse sich fortsetzen!

Brandon Amos   About   Blog

Image Completion with Deep Learning in TensorFlow

August 9, 2016

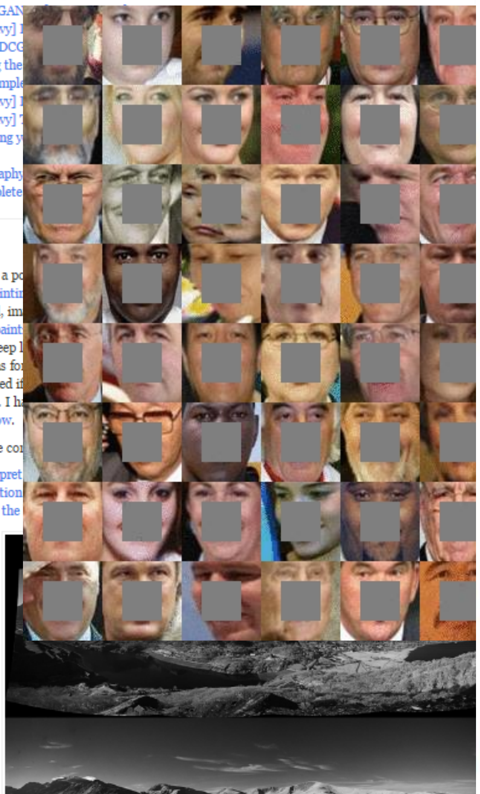
- Introduction
- Step 1: Interpreting images as samples from a probability distribution
  - How would you fill in the missing information?
  - But where does statistics fit in? These are images.
  - So how can we complete images?
- Step 2: Quickly generating fake images
  - Learning to generate new samples from an unknown probability distribution
  - [ML-Heavy] Generative Adversarial Net (GAN) building blocks
  - Using $G(z)$ to produce fake images
  - [ML-Heavy] Training DCGANs
  - Existing GAN
  - [ML-Heavy]
  - Running DCG
- Step 3: Finding the
  - Image comple
  - [ML-Heavy]
  - [ML-Heavy] T
  - Completing y
- Conclusion
- Partial bibliography
- Bonus: Incomplete

Introduction

Content-aware fill is a p… completion and inpaintin… do content-aware fill, im… "Semantic Image Inpaint… shows how to use deep l… some deeper portions fo… section can be skipped i… from images of faces. I h… completion.tensorflow.

We'll approach image co…

1. We'll first interpret
2. This interpretation
3. Then we'll find the



Andrej Karpathy blog   About   Hacker's guide to Neural Networks

The Unreasonable Effectiveness of Recurrent Neural Networks

GEEK.COM

TECH

Nvidia AI Generates Fake Faces Based On Real Celebs

BY STEPHANIE MLOT 10.31.2017 :: 10:00AM EST

32 SHARES

I'm getting a distinctly mid-90s "The Rachel" vibe from the woman in the top left corner (via Nvidia)

STAY ON TARGET

AI Shelley Pens Truly Creepy Horror Stories–And You Can Help

Neural Network Serves Up Truly Frightening Halloween Costume Ideas

Celebrity scandals are about to get a lot more complicated.

Nvidia has **developed** a way of producing photo-quality, AI-generated human profiles—by using famous faces.

the morning paper

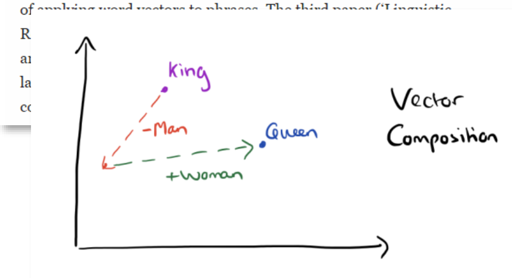
The amazing power of word vectors

APRIL 21, 2016

For today's post, I've drawn material not just from one paper, but from five! The subject matter is 'word2vec' – the work of Mikolov et al. at Google on efficient vector representations of words (and what you can do with them). The papers are:

- **Efficient Estimation of Word Representations in Vector Space** – Mikolov et al. 2013
- **Distributed Representations of Words and Phrases and their Compositionality** – Mikolov et al. 2013
- **Linguistic Regularities in Continuous Space Word Representations** – Mikolov et al. 2013
- **word2vec Parameter Learning Explained** – Rong 2014
- **word2vec Explained: Deriving Mikolov et al's Negative Sampling Word-Embedding Method** – Goldberg and Levy 2014

From the first of these papers ('Efficient estimation…') we get a description of the *Continuous Bag-of-Words* and *Continuous Skip-gram* models for learning word vectors (we'll talk about what a word vector is in a moment…). From the second paper we get more illustrations of the power of word vectors, some additional information on optimisations for the skip-gram model (hierarchical softmax and negative sampling), and a discussion



Beispiele aus der angewandten Forschung

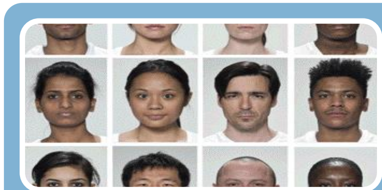
…mit lokalen Industriepartnern (KMUs)

Zürcher Hochschule für Angewandte Wissenschaften

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency



Gesichtserkennung für Stadionzutritt

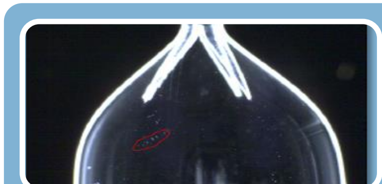
[!] DEEPIMPACT

- Nutzen: *Robustes* Personenidentifikationssystem
- Wirkung: Unterstützung bei Entwicklung; Datenqualität schränkte ein

Automatische Artikelsegmentierung

ARGUS DATA INSIGHTS

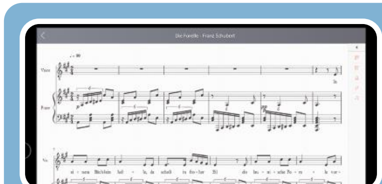
- Nutzen: vollautomatisches Produkt in niedrigem Preissegment
- Wirkung: Einführung dank *Teamausbau* geglückt



Visuelle Qualitätskontrolle in Produktion

BW-TEC

- Nutzen: vollautomatischer Triage & Bearbeitung normaler Fälle
- Wirkung: macht *Familienunternehmen* zu Technologieanbieter



Digitalisierung von Musikalien

SCOREPAD

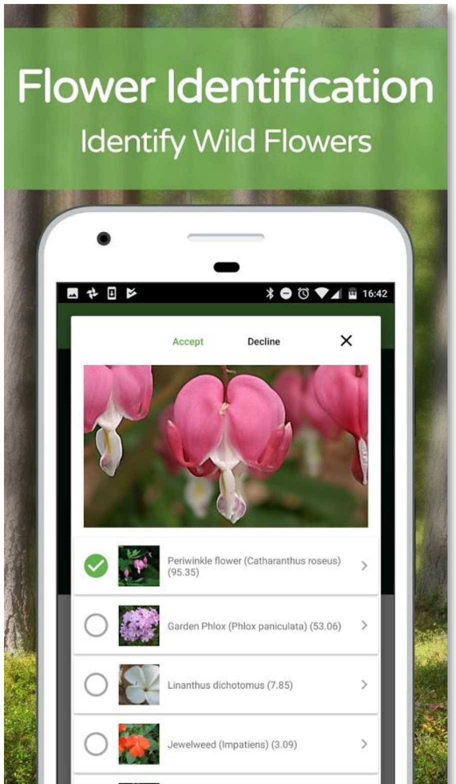
- Nutzen: Enabler für digitales Geschäftsmodell
- Wirkung: 5 Jahre nach Start ist entwickelte Technologie *grösstes Asset*

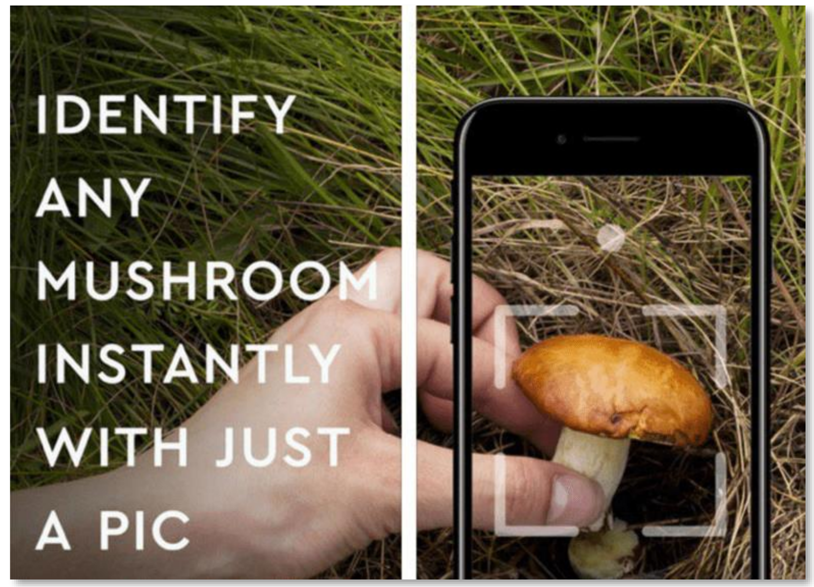
Was→ Wo?→ Wohin?

3

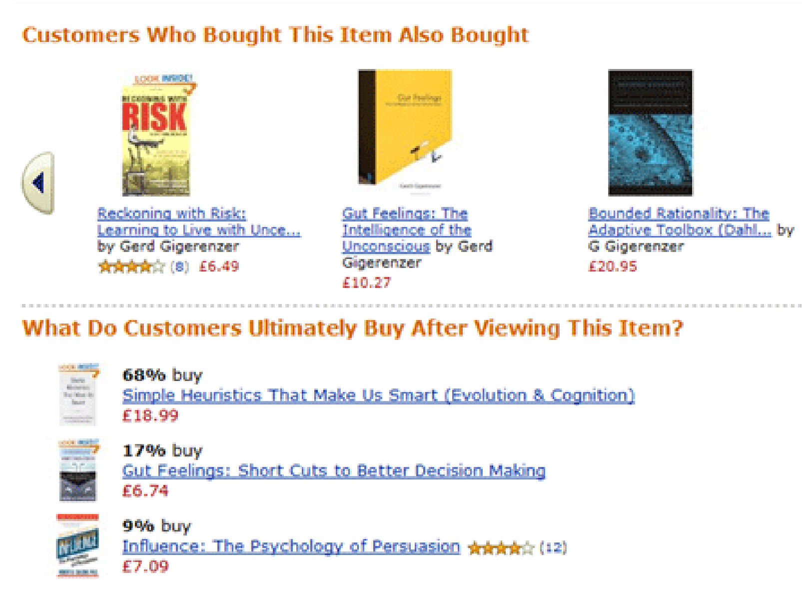
Wohin könnte und das einmal führen?

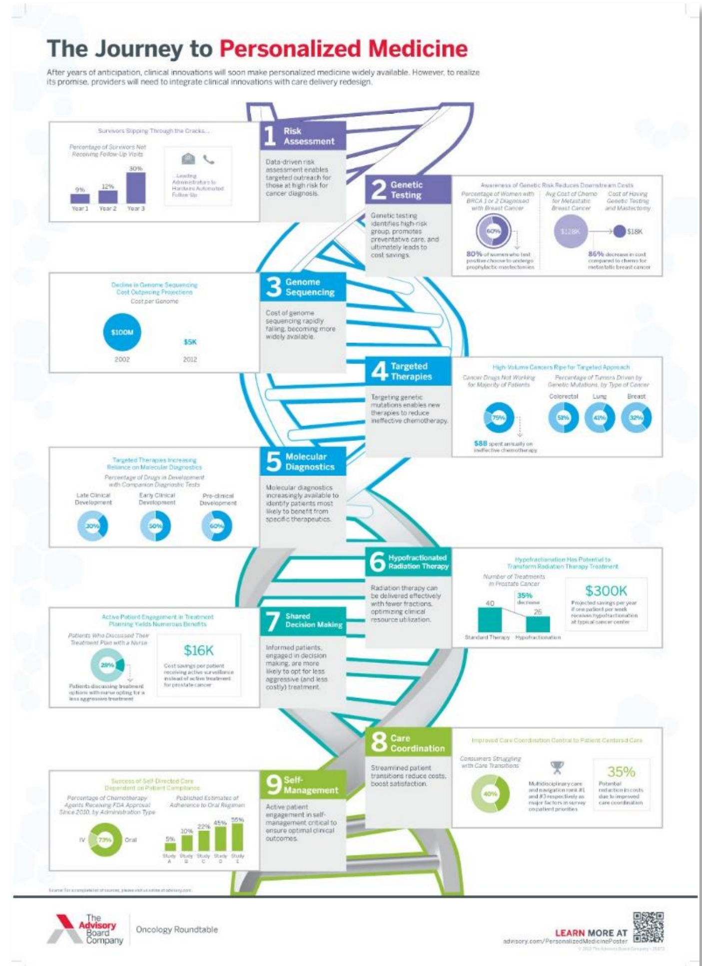
Beispiel: Machbar vs. gefährlich
Technologie: Computer Vision mit Deep Learning





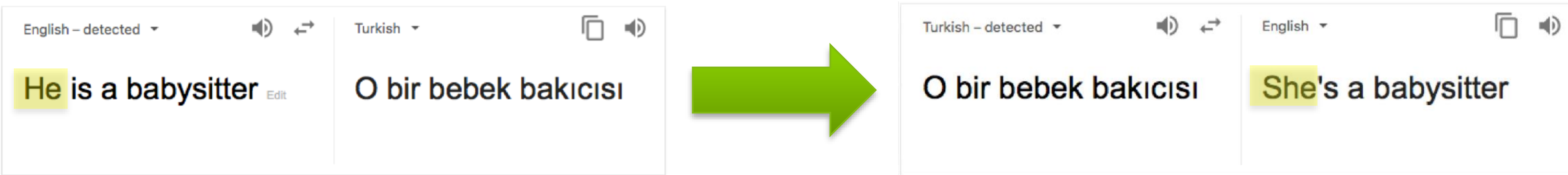
Beispiel: Markterfolg vs. regulatorische Hürden
Technologie: Recommender Systems

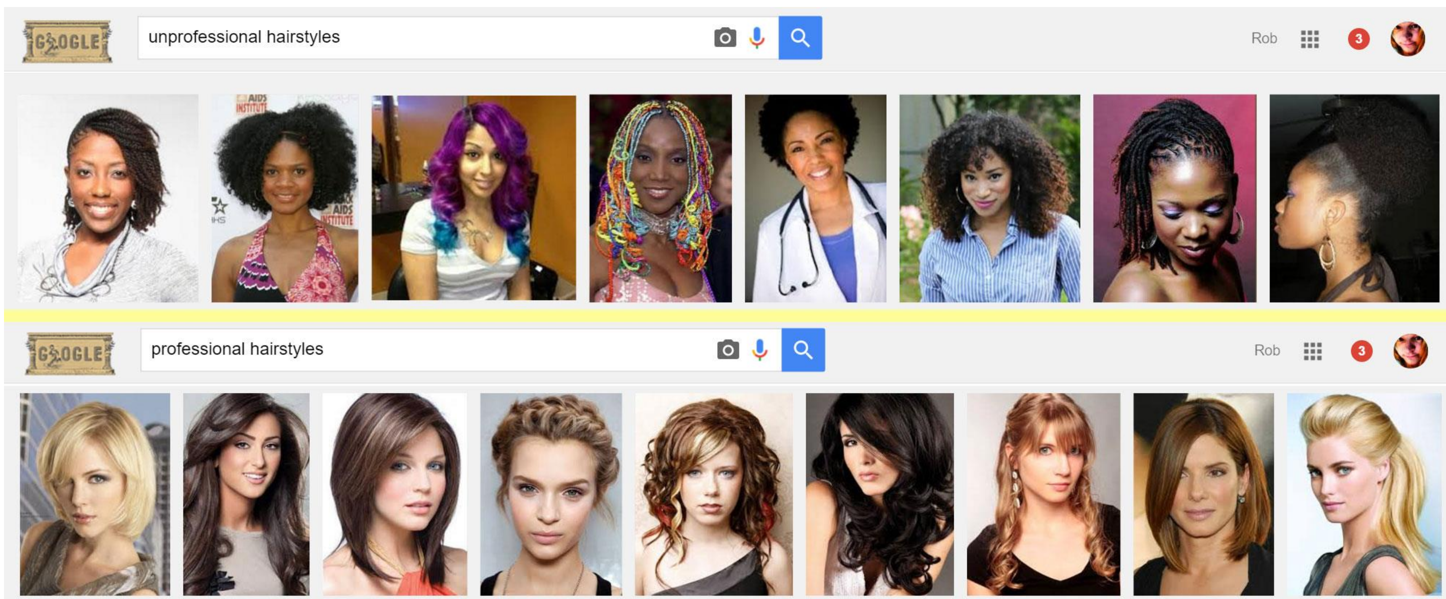




Zürcher Hochschule
für Angewandte Wissenschaften

Beispiel: Statistik vs. Bias
Technologie: Machine Learning

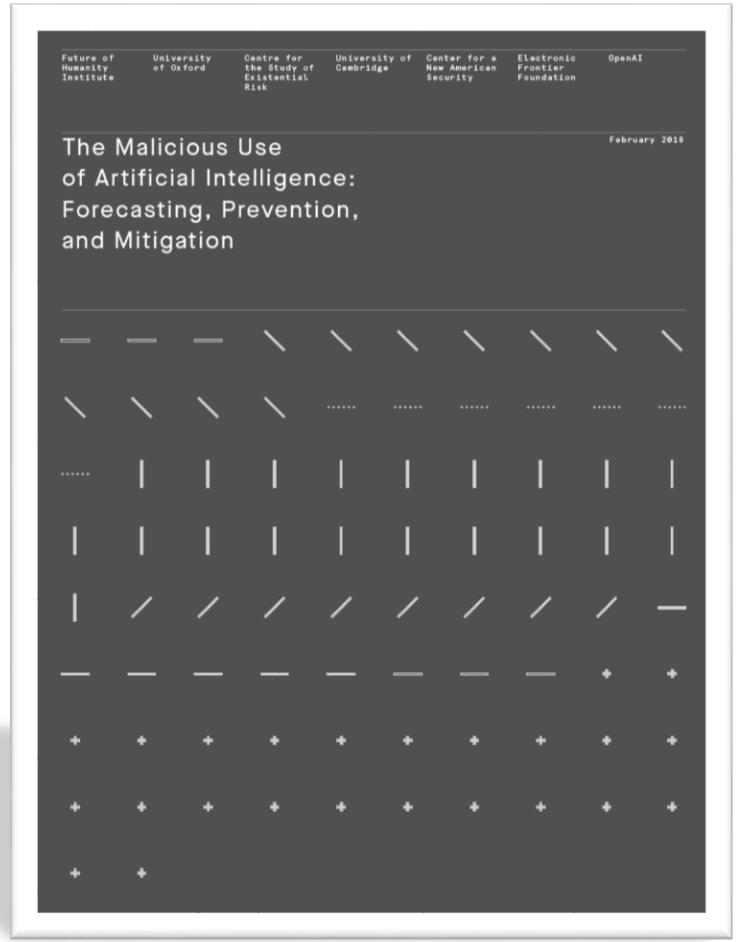




See also: Nassim Nicholas Talib, «*The Black Swan: The Impact of the Highly Improbable*», 2007

Gefahren durch KI?

- KI ist per Definition eine “**dual use Technology**”
  → siehe Report von Brundage et al., 2018
- Aber: “**natürliche Dummheit**” ist die grössere Bedrohung
- **Algorithmische Ethik** und **erklärbare KI** sind in den letzten Jahren zu einem top Forschungsfeld geworden – nicht wegen der unkalkulierbaren Risiken per se, sondern:



Aussicht: Disruption
…selbst bei völliger Stagnation des technischen Fortschritts

1. Hypothese: Einsatz (aktueller) KI wird sich massiv ausbreiten (Zeitrahmen: 5 Jahre)
- Indikator: **KI-Fortschritt** momentan hauptsächlich **Industriegetrieben** (**Gewinnaussicht**); Konsumenten kaufen “bequem”; diese Incentivierung “hält den Motor am Laufen”

2. Hypothese: Dies wird unsere Gesellschaften umwälzen
- Kernfragen: Wie **verteilt** sich der algorithmisch (hauptsächlich bei Grosskonzernen) erwirtschaftete **Gewinn**? Wie verteilt sich neue **Freizeit** und **Alltagserleichterung**?

3. Hypothese: Grösste Frage wird der Umgang miteinander sein (nicht der Umgang mit KI)
- Argument: KI (etc.) “for the common good” ist ein wichtiges Thema; entscheidend wird jedoch sein, wie wir **als Gesellschaften die Regeln** für das digitalisierte Zusammenleben (s.o.) **gestalten**



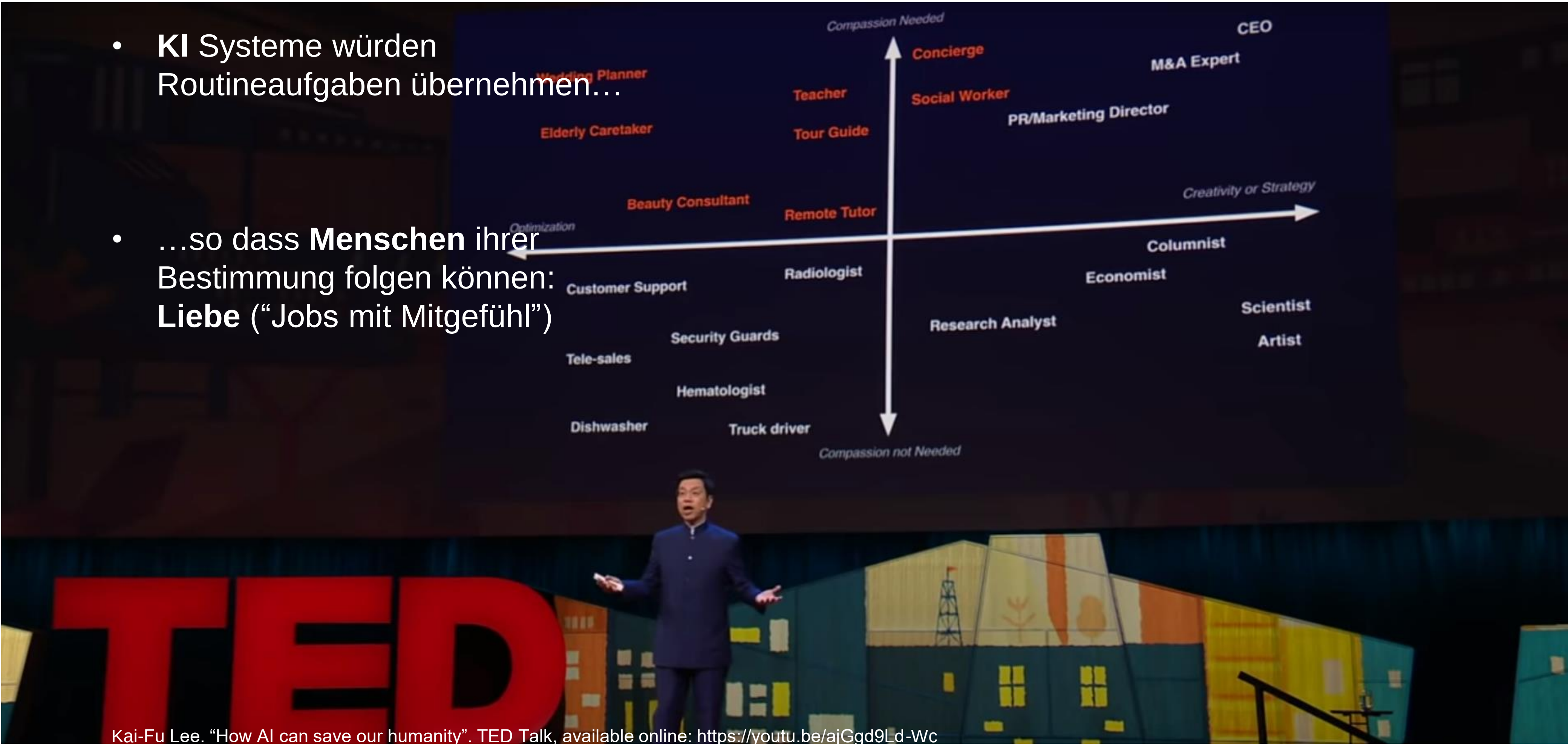
Siehe auch: Stockinger, Braschler & Stadelmann. “Lessons Learned from Challenging Data Science Case Studies”. In: Braschler et al. (Eds), *“Applied Data Science - Lessons Learned for the Data-Driven Business”*, Springer, 2019.

Eine lebenswertere Gesellschaft durch KI?

Die Vision von Kai-Fu Lee, Unternehmer & Forscher

- **KI** Systeme würden Routineaufgaben übernehmen…

- …so dass **Menschen** ihrer Bestimmung folgen können: **Liebe** ("Jobs mit Mitgefühl")



Kai-Fu Lee. "How AI can save our humanity". TED Talk, available online: https://youtu.be/ajGgd9Ld-Wc

Schlussfolgerungen

- KI automatisiert *einzelne*, komplexe, aber *redundante* Prozesse
  (meist mittels maschinellem Lernen auf menschengenerierten Beispielen)
- Deep Learning hat zu Paradigmenwechsel in *Mustererkennungsaufgaben* geführt
- Das Ergebnis könnte eine menschlichere Gesellschaft sein
- Das Zeitfenster zum Gestalten beträgt wenige Jahre (<5)

datalab
www.zhaw.ch/datalab

Swiss Alliance for
Data-Intensive Services

Zu mir:

- Prof. KI/ML, Scientific Director ZHAW digital
- Email: stdm@zhaw.ch
- Telefon: 058 934 72 08
- Web: https://stdm.github.io/
- Twitter: @thilo_on_data
- LinkedIn: thilo-stadelmann

Mehr zum Thema:

- Veranstaltungen in CH: www.data-service-alliance.ch
- Lesenswert: https://www.deeplearning.ai/thebatch/
- Ebenso: Stuart Russell, «Human compatibel», Penguin Books, 2019

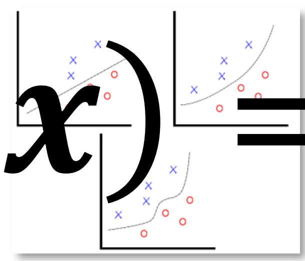
ANHANG

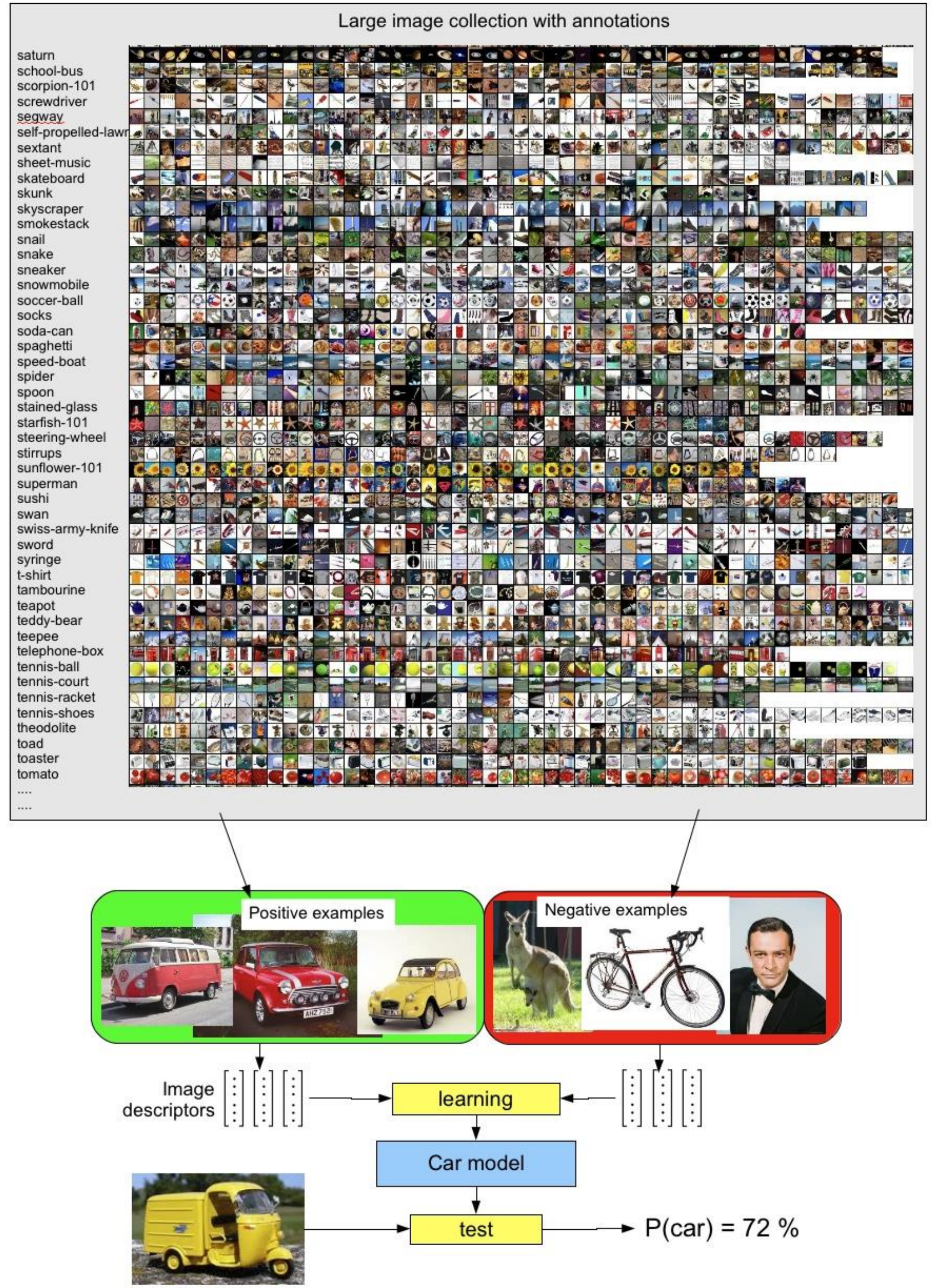
Grundlage
Induktives überwachtes Lernen

Annahme
- Ein an *genügend viele* Beispiele angepasstes Modell…
- …wird auch auf unbekannte Daten **generalisieren**

Methode
- **Suchen der Parameter einer gegebenen Funktion**…
- …so dass für alle Beispiele Eingabe (Bild) auf Ausgabe («Auto») abgebildet wird

$$f(x) = y$$

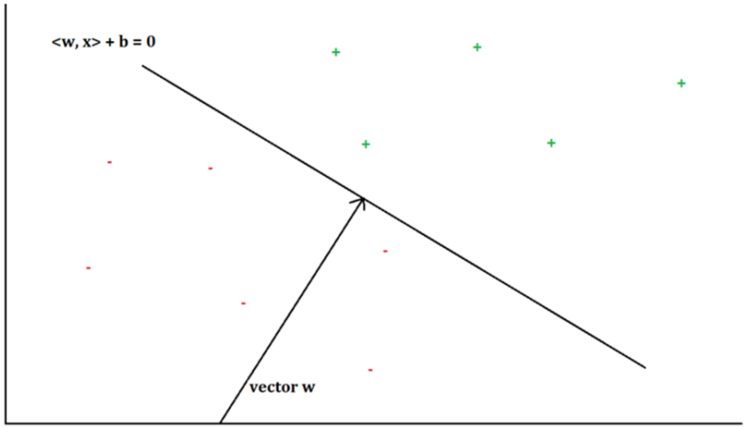


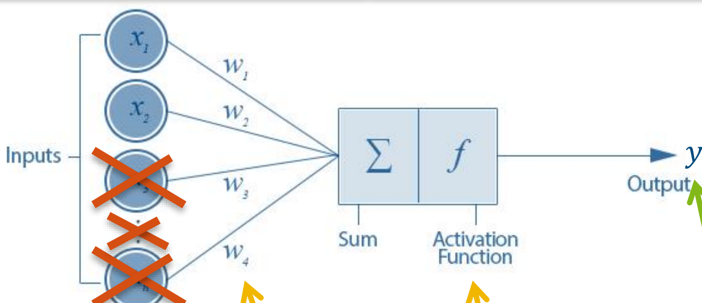


Quelle: http://lear.inrialpes.fr/job/postdoc-large-scale-classif-11-img/attribs_patchwork.jpg

Suche der Parameter *einer Funktion*?

Neuron





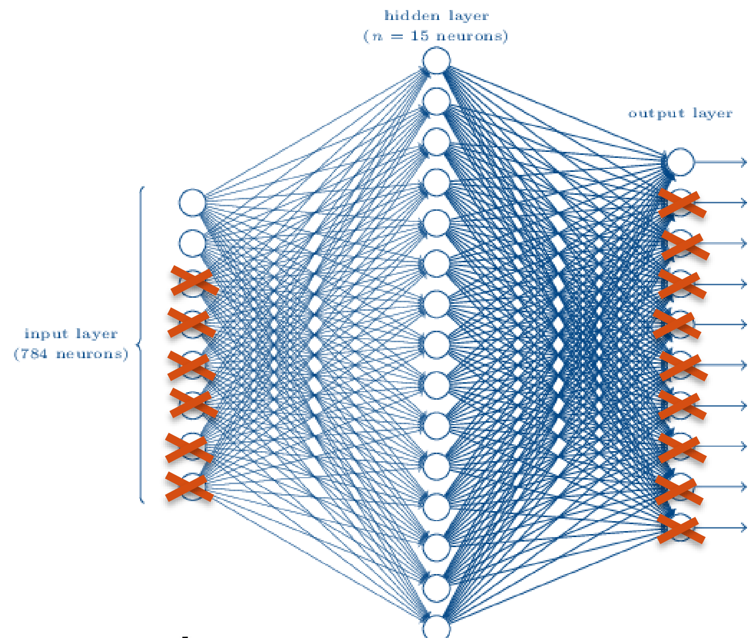
Merkmale (z.B. Pixel)

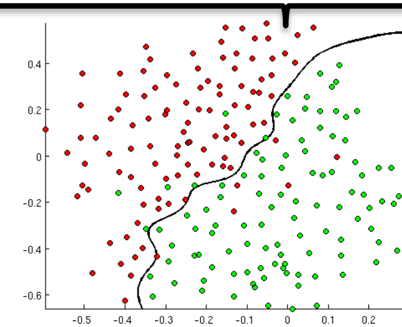
Anpassbare Parameter

Entscheidung (Schwellwert)

Ergebnis (z.B. «Auto»)

Neuronales Netz





Suche der Parameter einer Funktion?

- Unser Neuronales Netz: $f_W(x) = y$
  mit Bild x, echtem Resultat y und Parametern $\boldsymbol{W}$
  ($\boldsymbol{W} = \{w_1, w_2, \ldots\}$ anfangs zufällig gewählt)

- Fehlermass: $l(\boldsymbol{W}) = \frac{1}{N}\sum_{i=1}^{N}(f_W(x_i) - y_i)^2$
  Durchschnitt der quadratischen Abweichungen über alle Bilder (Loss)

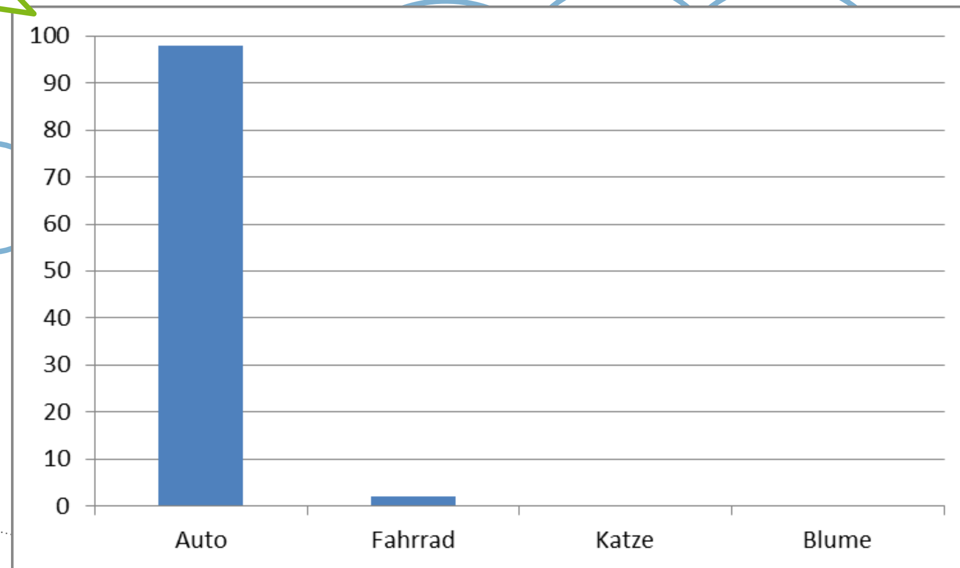
$$l(\boldsymbol{W}) = \underbrace{\frac{1}{N}\sum_{i=1}^{N}}_{\text{Durchschnitt (über alle Beispiele)}}\underbrace{(f_W(x_i) - y_i)}_{\text{Differenz IST – SOLL (Fehler)}}\underbrace{{}^2}_{\text{Bestraft grosse Fehler überproportional stärker}}$$

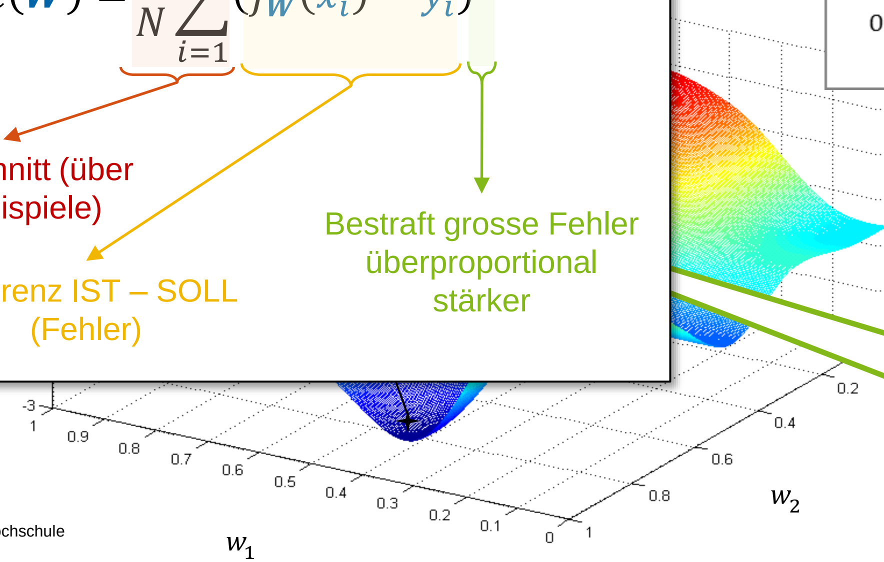
Durchschnitt (über alle Beispiele)

Differenz IST – SOLL (Fehler)

Bestraft grosse Fehler überproportional stärker

Wahrscheinlichkeit [%] für bestimmtes Ergebnis



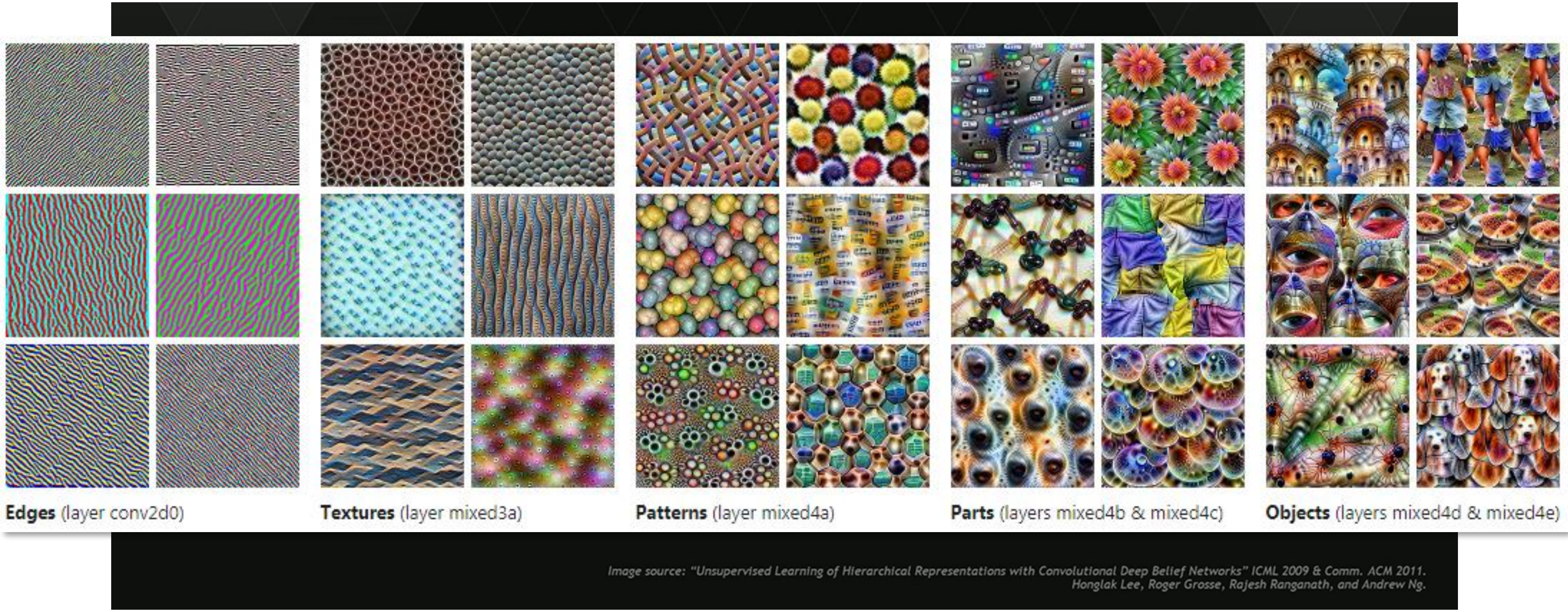


← Fehlerlandschaft

Methode: Anpassung der Gewichte von f in Richtung der steilsten Steigung (abwärts) von J

Was «sieht» das Neuronale Netz?
Hierarchien komplexer werdender Merkmale

Zürcher Hochschule
für Angewandte Wissenschaften



Quellen: https://www.pinterest.com/explore/artificial-neural-network/
Olah, et al., "Feature Visualization", Distill, 2017, https://distill.pub/2017/feature-visualization/.