

Institut für Computational Engineering

Klaus Frick



Können Maschinen denken?

Ursprung und Grundbegriffe der künstlichen Intelligenz

summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

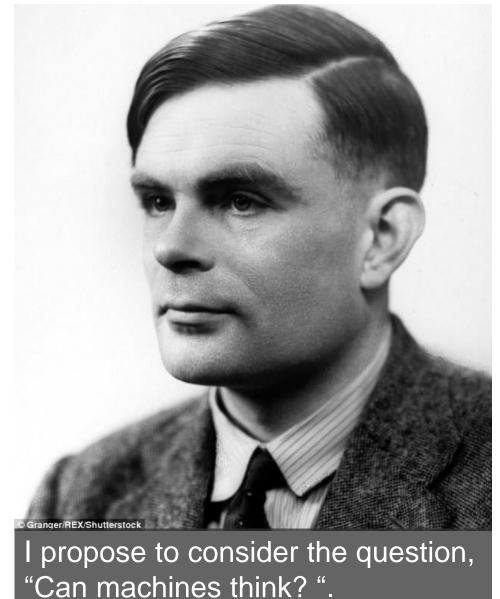


701 C 11 ·

Alan Turing

- Britischer Mathematiker/Logiker.
- Entschlüsselt im 2. WK die Codiermaschine «Enigma» der deutschen Wehrmacht.
- Artikel: «Computing Machinery and Intelligence» im Jahr 1950.
- Geburtstunde der künstlichen Intelligenz («Erste Welle»)
- Turing-Test: Test, um künstliche und natürliche Intelligenz zu vergleichen (imitiation game).
- CAPTCHA





"Can machines think?".



Dartmouth 1956

- Acht-wöchiger Workshop über Künstliche Intelligenz.
- Hauptantragsteller
 - J. McCarthy
 - M.L. Minsky
 - N. Rochester
 - C.E. Shannon
- Ziel:

An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves



Quelle: Margaret McCarthy



Erste KI Welle

Minsky (1970):

"In from three to eight years we will have a machine with the general intelligence of an average human being"

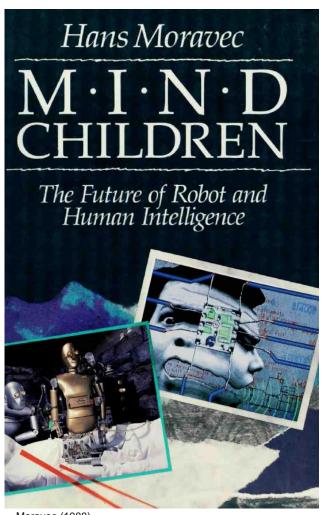
- Erster Al Winter: 1974 1980
 - Erkenntnis über die Komplexität der meisten realen Probleme führt zu NP-schweren Problemen
 - Limitierte Rechenpower bzw. Speicherkapazität
 - Funding wird eingestellt.
- Boom: 1980 1987

 - Expertensysteme werden erfolgreich (LISP)
 Hinton und Rumelhart: Backpropagation Algorithmus für Neuronal Netze
 - Fifth Generation Computer Projekt in Japan
- Zweiter Al Winter: 1987 1993
 - Desktop-Rechner von Apple und IBM lösen AI-Hardware (LISP Machines) ab.
 - Eifth Generation Computer hält die Erwartungen nicht.
 - Über 300 Al-Firmen sind bis 1993 bankrott





Moravec Paradox



- Hans Moravec (austro-kanadischer Robotikforscher)
- KI kann für Menschen «schwierige» Probleme leicht lernen, aber versagt bei «alltäglichen» Aufgaben

"it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility"

- Dies führt zu einer steigenden Skepsis gegenüber Strong Al.
- Neuausrichtung der Forschung in AI in den 1990er Jahren (2. Welle): Narrow AI.



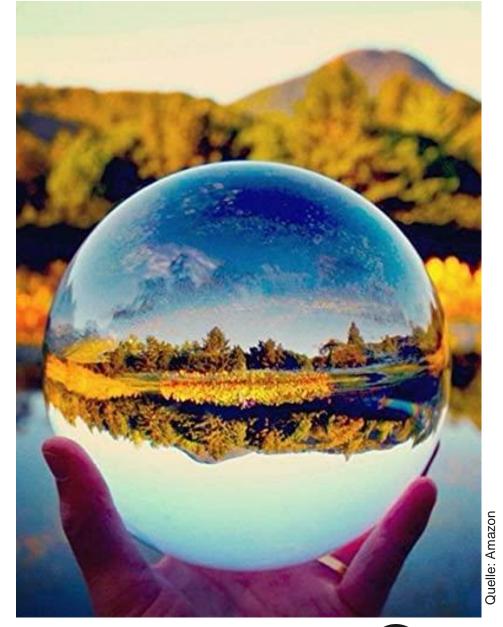


Predictions

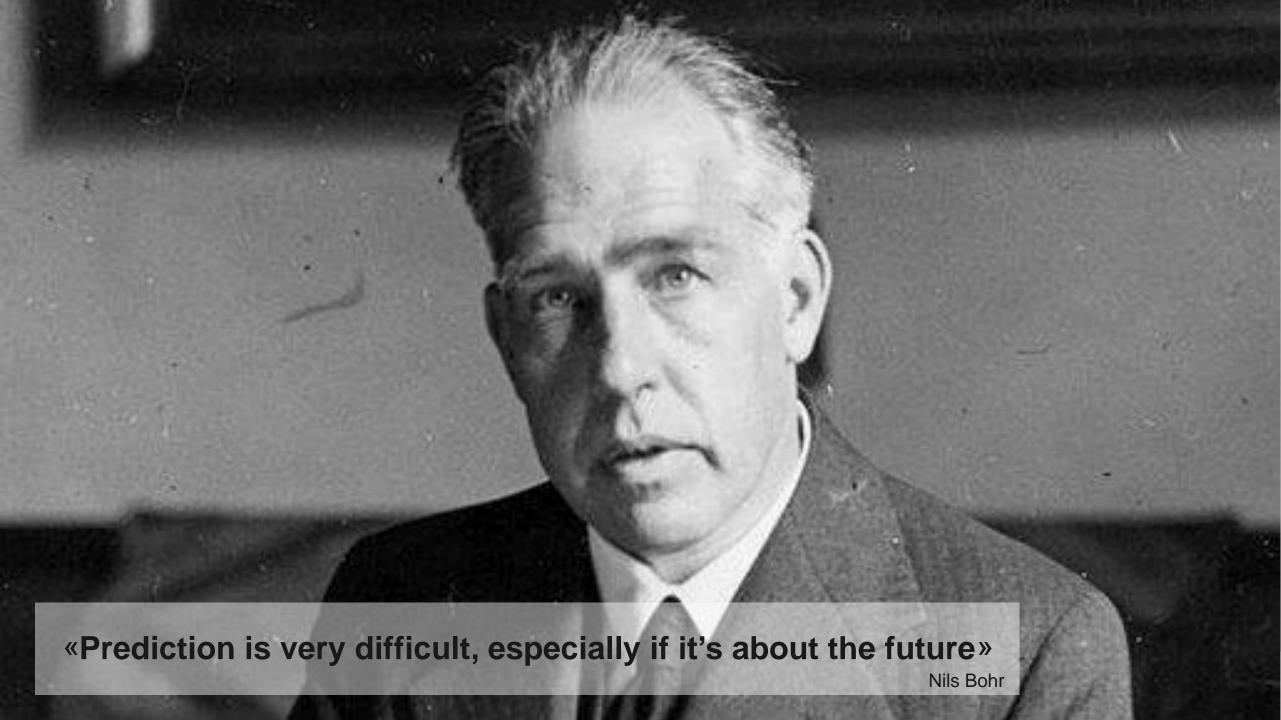
Kernstück aller moderner KI Systeme sind **Vorhersage- Modelle**

Vorhersagen ist der Prozess, fehlende Information bereitzustellen. Vorhersage-Modelle verwenden Information die man hat (z.B. Daten), um Information zu generieren, die man nicht hat.

- Dies ist die Definition von Predictive Modeling (A. Agrawal, J. Gans und A. Goldfarb 2018).
- Vorhersage-Modelle basierend auf Daten werden mit supervised Machine Learning realisiert.
- Einige KI Systeme beinhalten datenbasierte Vorhersage-Modelle, aber nicht alle.







Wo ist Ceres?

- Am 1. Januar 1801 entdeckt G. Piazzi den Zwergplaneten "Ceres" und notiert seine Positionen, verliert den Planeten am 11. Februar aber wieder.
- C. F. Gauss erhält im Oktober 1801 die Daten von Piazzi und berechnet eine **Vorhersage** der momentanen Position der Ceres. Der Planet wird am 7. Dezember 1801 wiederentdeckt.
- Nach eigenen Angaben verwendet Gauss zur Schätzung der Position die Methode der kleinsten Quadrate.

Beobschtungen des zu Palermo d. C'Jan. 1821 von Prof. Pinzu neu entdeckten Gaftirns.

180	1	bot	nen- eit	A	afit	ade alg	in	A	eig	ung don	A	bw	eicl	١.	1	che	Li	inge	G	Bu	estr.		+	20		d.		ftans
Jan.	2	8 39	17.8 4.6 53.3	13	26	63.	85.	51	43	27,-8	15	41	435	5.	1	23 23	19	44-	3/3	2	24.9	.9	11	2	28, 6	slo.	993	6317
	10 11	8 30	42.1 15.8 17.5 26,1	3	25 25 25	23 32, 29.	15 1:: 73	51 51	35 23 23	47.3 1.5 26.6	16	47	32,	0	1	23	7	59.	5 2	53 29	55,6	9	20	10	17.	9,	992	6418 7641
	14 17 18	7 50	31.7 11.3 148.5	3.	25 25	55.	72	.51 .51	22	45, 0	16	40	13.	7	1	23	12	1,	2 2	12	56,7	9	24	14	13.5	9,	992	8809
	21 22 23	7 24 7 20 7 10	2.7 21,7 43.5 51.3	3	26 26 27	49,	42	51 51 51	38 42 46	34, 1	17	58 3	18,	5	1 7 1	23 23 23	34 39 44	21, 1,	3 1	42 38	6,0 28,1 52,1	IO IO	2	20 21 22	40, 32,	9.	993 993 993	1434 1886 2348
Febr	30 31	6 4	52,9 8 26,4 4 59,9 1 35,8	3	30 30	48, 17, 47.	25 20	52 52 52	27 34 41	18. 8 48. 0	17	48 53	11, 21, 36,	5	1 1	24 24 24	30 38 46	7.	3 1	10	16,0 54.6 30.0	10	11	27 28	28.	9,	993 993 993	6332 7007 7703
	5 8	6 2	31,5 1 39,2 1 58,2	3	33	58.	70 50	53 53	15	37.5	18	31	23.	2	1	25	53	43,	4 0	54	23.9	10	16	31	45,	9.	994	3276



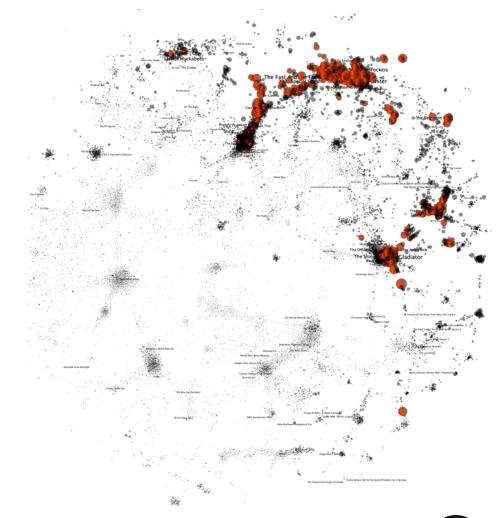


Netflix Price 2006

2006 stellt Netflix 100.480.507
 Filmbewertungen online:

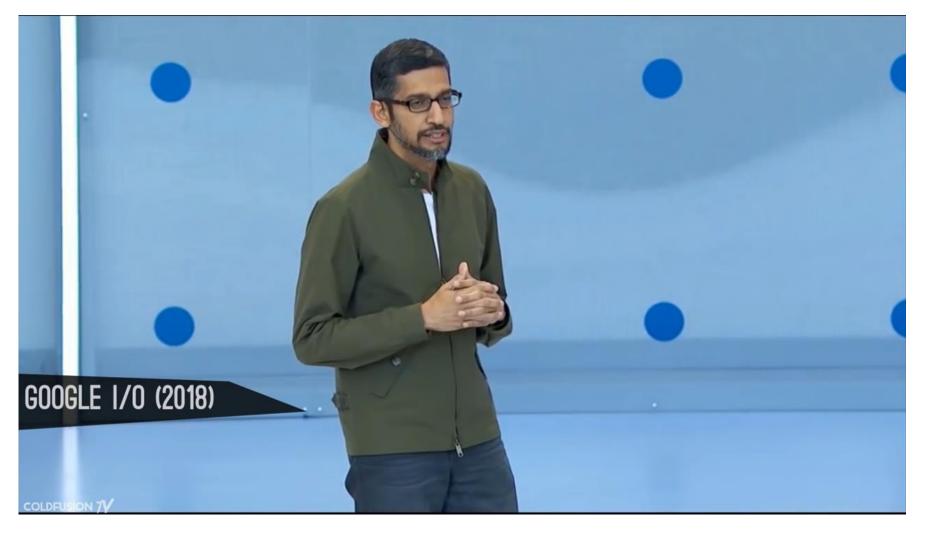
```
[User-ID; Movie-ID; Date; Score]
```

- Insgesamt 480.189 Benutzer und 17.770 Filme.
- Ziel: Sage für jeden User die Bewertung für jeden Film voraus. (Recommender system).
- Das Team, das als erstes ein Vorhersage-Modell baut, dass einen 10% besseren match erzielt, erhält 1 Mio \$.
- Ursprung der data science competitions



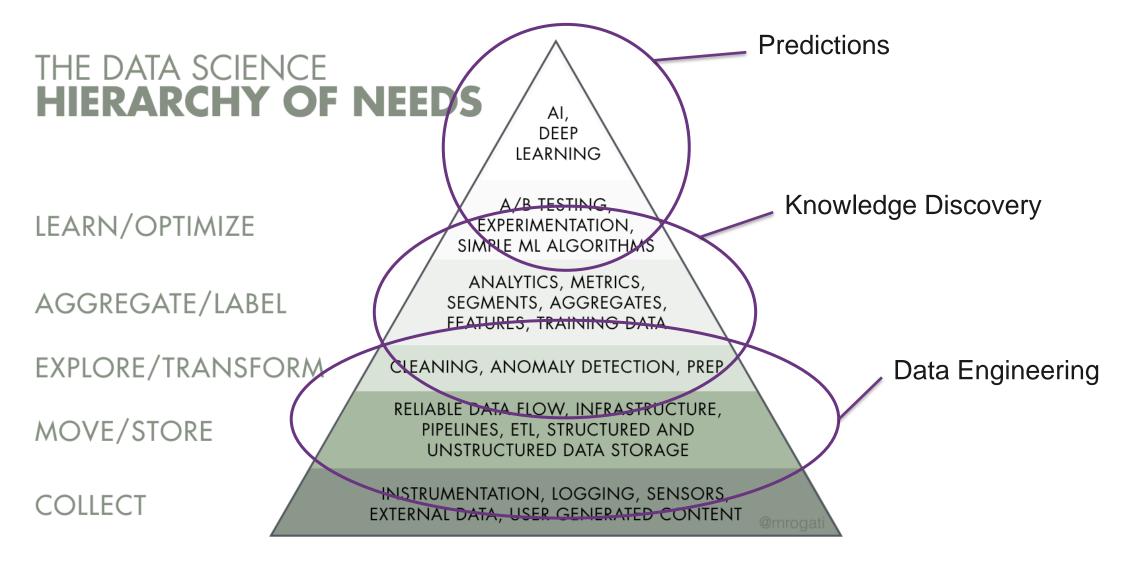


Google Duplex 2018





Al 1/1 der Fraxis 30. September 2020



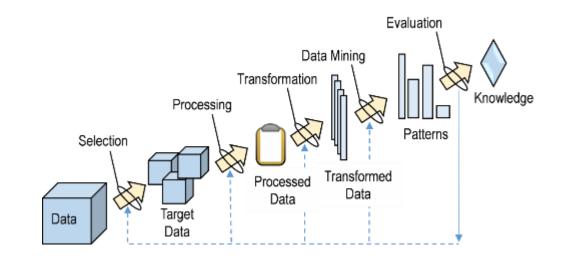


Knowledge Discovery

Ziel in komplexen industriellen Prozessen ist es

stichhaltige, neue, potentiell brauchbare und schlussendlich verständliche Muster in Daten zu identifizieren.

- Dies ist die Definition von knowledge discovery in databases (KDD) nach Gregory Piatetsky-Shapiro (1989).
- Synonym: Data Mining
- Schliesst klassische statistische Verfahren mit ein.
- Ist ein Prozess und keine Methode.





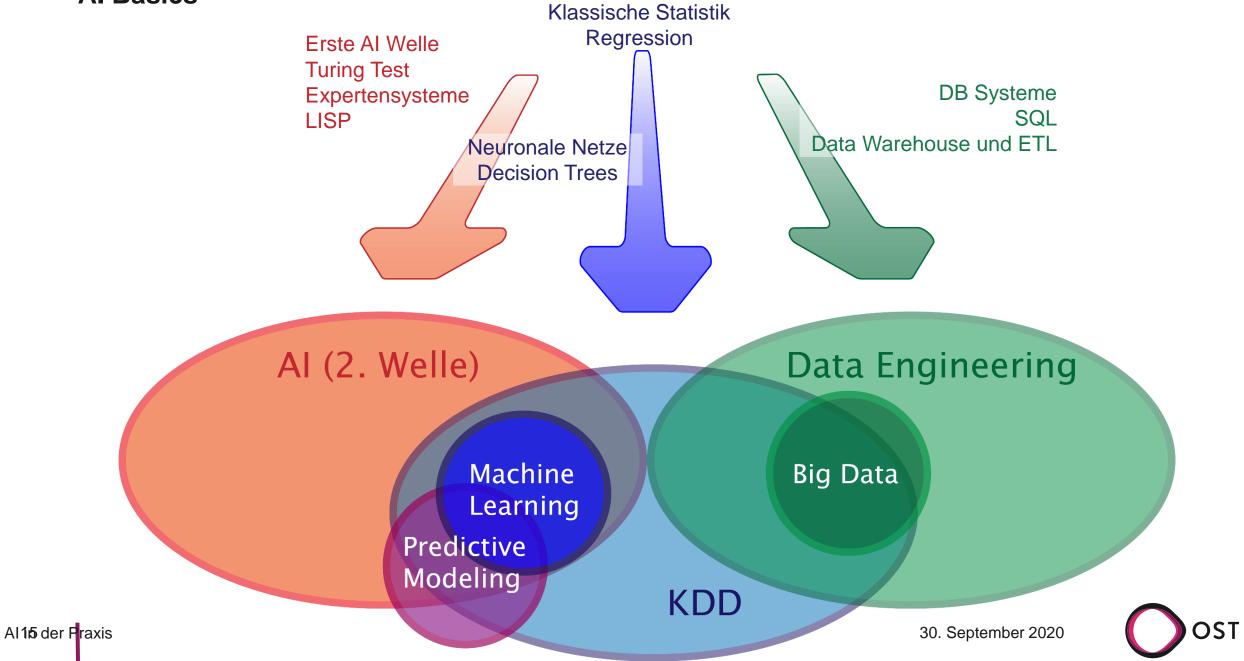
Data Engineering

Aufgabe des Data Engineering ist es

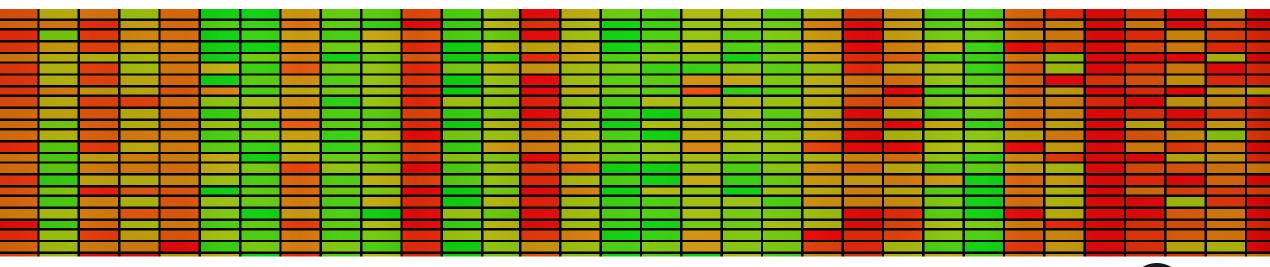
Daten-Pipelines zu erstellen und zu pflegen sowie die Daten zu bereinigen, zu verknüpfen und in einer geeigneten Form zu speichern.

- Klassische Aufgaben: Datenbanken (SQL) und Data-Warehouses (ETL Prozess)
- Die Grösse (volume) und Geschwindigkeit (velocity) und Vielseitigkeit (variety) sind Teil des Data Mining Problems. Man spricht dann von Big Data
- Cloud Computing (AWS, Microsoft Azure, etc.) bietet neue Möglichkeiten.





Einsatz von Machine Learning im Rheintal





Machine Learning

Herbert Alexander Simon



«Learning is any process by which a system improves performance from experience. Machine Learning is concerned with computer programs that automatically improve their performance through experience»

Tom Mitchell



«A computer program ist said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.»

OST

Al 1 rd der Praxis 30. September 2020

Arten von Machine Learning

ARTIFICIAL INTELLIGENCE (AI)

Jede Technik (in Form von Software), die seine Umgebung wahrnimmt und anhand dieser Inputs menschliche Intelligenz imitiert, um die Wahrscheinlichkeit ein gewisses Ziel zu erreichen zu maximieren.

Teilgebiete: Computer Vision, Robotik, Natural Language Processing, Route Planing, Machine Learning, ...

MACHINE LEARNING

Eine spezielle Sammlung an Algorithmen, realisiert in Software, die basierend auf mathematischen und/oder statistischen Verfahren es Computern erlaubt, die Wahrscheinlichkeit ein Ziel zu erreichen anhand von Erfahrungen (in Form von Daten) zu erhöhen.

Supervised Learning (Lernen durch Nachahmung)

Unsupervised Learning (Lernen durch Suchen)

Reinforcement Learning (Lernen durch Belohnung)



Supervised Learning

Ausgangslage

- Trainingsdaten sind in ausreichendem Umfang vorhanden.
- Die Daten liegen als panel data vor.
- Jede Beobachtung besteht aus messbaren Merkmalen (features) und einer Response (label).
- Die Labels wurden von einem Experten (supervisor) vergeben

Ziel

Finde einen Algorithmus bzw. eine Funktion, die für zukünftige Beobachtungen der Merkmale die Response vorhersagt (ohne supervisor).

	Merkmal 1	Merkmal 2	Merkmal 3	Merkmal 4	Merkmal 5	 Merkmal p	Response
1	2.31	1.0	-13	True	0	2.2	OK
2	2.33	1.2	-13	False	0	2.1	OK
3	2.18	1.5	-12	False	1	1.1	NOK
4	2.18	1.1	-13	True	0	1.0	OK
5	2.17	1.1	-12	True		1001	NOK
n	2.01	1.0	-5	True	0	0.8	OK

2.08	1.2	15	False	1	1.8	???





Soft Sensor

Zweck

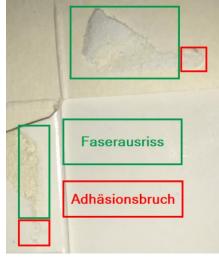
- Monitoring von Verpackungsanlagen von pharmazeutischen Produkten
- Klebestellen von Schachteln sind qualitäts- und sicherheitsrelevant

Konzept

- Klebelaschen werden automatisch mit Klebstoff benetzt und mittels Motoren in einem Aufrichtmodul angepresst.
- Überwachung der Motordrehmomente und Schleppfehler ermöglicht Bewertung der Klebestellen









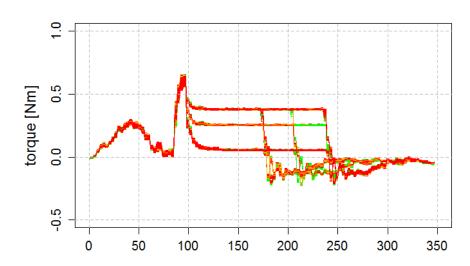
Soft Sensor

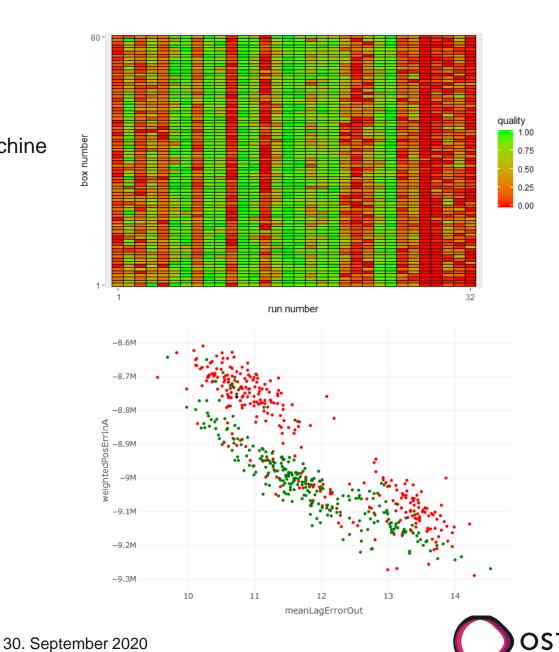
Ansatz

- Datenbasierte Kalibration der Klebemodule auf jeder Maschine
- Bestimmung optimaler Betriebsparameter mittels DoE
- Feature-Erzeugung aus Motorenkurven
- Regressionsanalyse der Klebstellenqualität

Herausforderungen

- Robust Feature Engineering und Einbettung auf Maschine
- Übertragbarkeit bzw. Generalisierung der Modelle





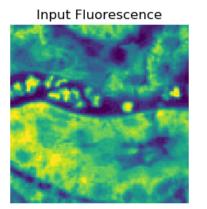
Digital Pathology

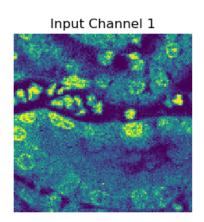
Zweck

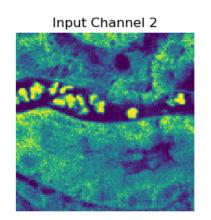
- Beschleunigung des Diagnoseprozesses in der Pathologie
- Langwierige F\u00e4rbung von Gewebeschnitten soll beschleunigt werden

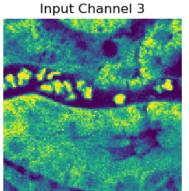
Konzept

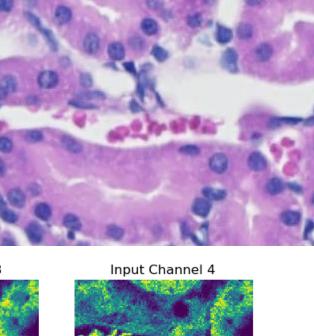
- Neuartige Multimodal Mikroskoptechnik entwickelt von Prospective Instruments in Dornbirn (Campus-V)
- Aus der reichhaltigen Information von multimodalen Bildern soll ein virtuelle Färbung der Probe erstellt werden.

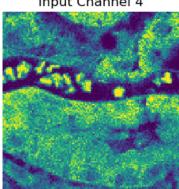














PROSPECTIVE

PROSPECTIVE

Digital Pathology

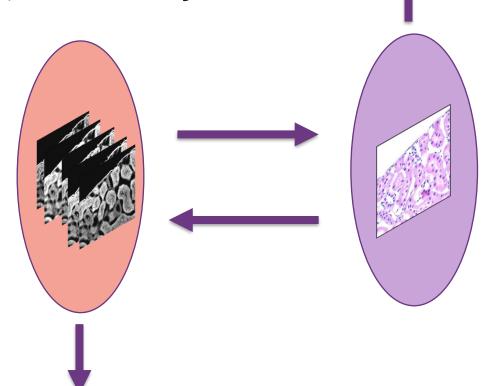
Ansatz

- Supervised Machine Learning zum Erlernen des Stainings
- Deep Learning Architektur basierend auf Faltungsnetzen (CNN)
- Generator Diskriminator Setup

Herausforderungen

- Multimodale Input Bilder und gefärbte Schnitte sind nicht identisch.
- Keine eins-zu-eins Korrespondenz der Pixel





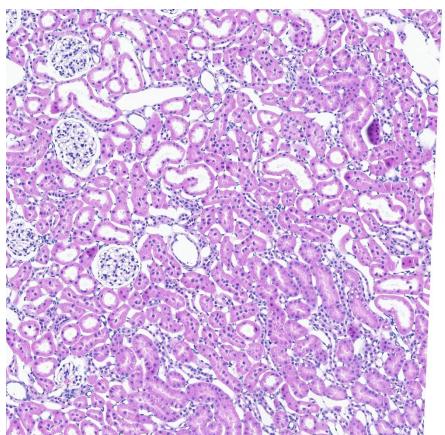


Pix2Pix vergleich

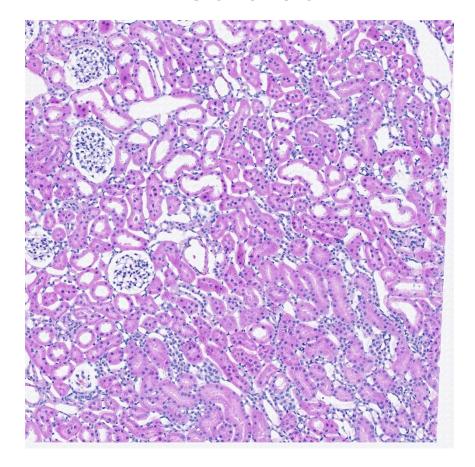
PROSPECTIVE

Digital Pathology

Target



Predicted





Unsupervised Learning

Ausgangslage

- Trainingsdaten sind in ausreichendem Umfang vorhanden.
- Die Daten liegen als panel data vor (3. Normalform)
- Jede Beobachtung besteht aus messbaren Merkmalen (features)
- Es gibt aber keine Response

Ziel

- Entdecke Muster und Gruppen in den Daten (Clustering).
- 2. Ersetze den Datensatz durch eine neue Tabelle mit viel weniger Spalten ohne viel Information zu verlieren (Dimensionsreduktion)

	Merkmal 1	Merkmal 2	Merkmal 3	Merkmal 4	Merkmal 5	 Merkmal p
1	2.31	1.0	-13	True	0	2.2
2	2.33	1.2	-13	False	0	2.1
3	2.18	1.5	-12	False	1	1.1
4	2.18	1.1	-13	True	0	1.0
5	2.17	1.1	-12	True		1001
n	2.01	1.0	-5	True	0	0.8

	2.08	1.2	15	False	1		1.8
--	------	-----	----	-------	---	--	-----



Unsupervised Learning

Monitoring ölfreier Kompressoren

Zweck

- Überwachung von ölfreien Kompressoren (Monitoring)
- Insbesondere während der Entwicklung

Konzept

- Kompressoren werden im Testbetrieb mit einer Fülle von Sensoren ausgestattet.
- Ziel ist das Monitoring des Gesundheitszustandes des Kompressors.
- Welche Sensoren sind entscheidend (für die Serie)?

Ansatz

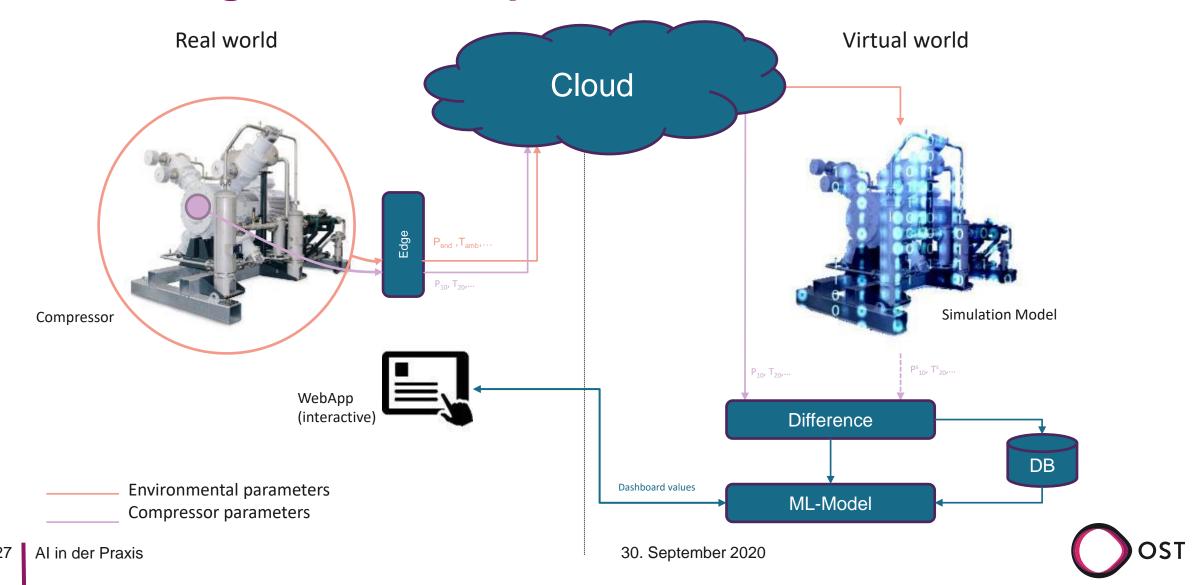
- Thermodynamisches Simulationsmodell der Drücke und Temperaturen im Kompressor gegeben die extrinsischen Parameter.
- Monitoring der Differenz der modellierten und gemessenen Werte mittels unsupervised learning





Unsupervised Learning

Monitoring ölfreier Kompressoren



Reinforcement Learning



Fine Film

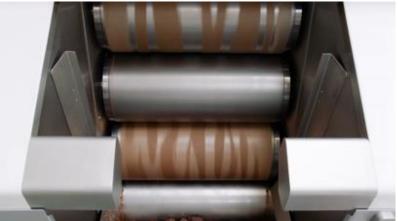
Zweck

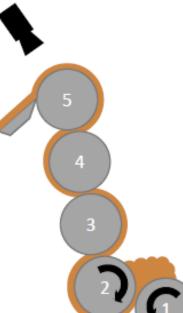
- Fünfwalzwerk verkleinert die in der Schokoladenmasse vorhandenen Zuckerkristalle
- Für das Erreichen der notwendigen Korngösse ist eine homogene Belegung der Walzen entscheidend

Konzept

- Schokoladenmasse wird durch fünf Walzen gezogen
- Der Film wird mittels verschiedener Temperaturen und Drücke gesteuert
- Überwachung der Qualität mittels Zeilenkamera (FineFilm System).









Reinforcement Learning



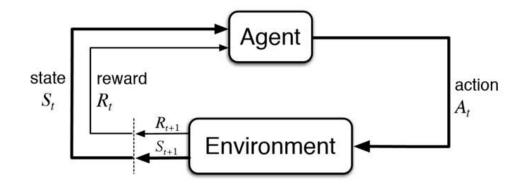
Fine Film

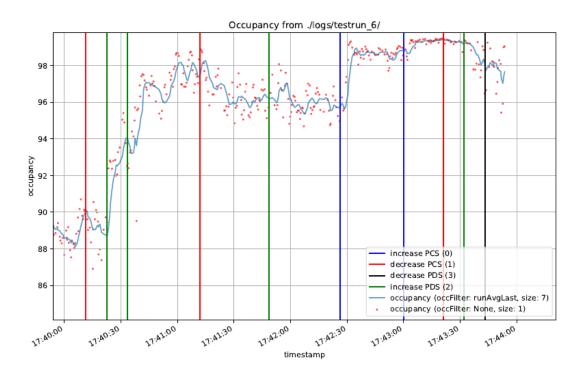
Ansatz

- RL-Agent als Regler für das Walzwerk (Q-Learning)
- Action-Space (diskret):
 - Tempertaturen und Drücke
- State-Space (kontinuierlich):
 - Bildfeatures aus dem FineFilm System
 - Ist-Werte der Temperaturen und Drücke
- Reward:
 - Belegungsgradänderung

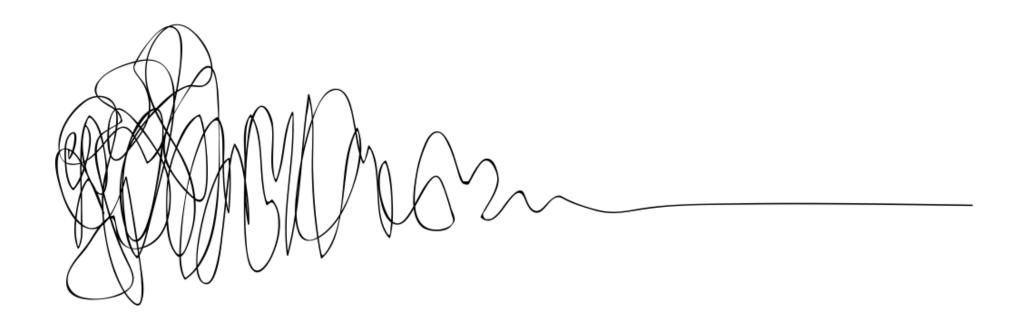
Herausforderungen

- Funktionsapproximation für die aktion-value Funktion.
- Keine Simulation möglich, sondern Lernen nur im Live-Modus möglich.
- Gutes Startmodell nötig.





Empathie und Engineering





Ups...

Gartner Report (2016):

... estimated that 60% of analytics and BI projects fail.

Ein Jahr später korrigiert der Gartner Analyst Nick Heudecker. Die Abschätzung war "too conservative". Der eigentliche Anteil:

"[...] closer to 85 percent "

"The problem isn't technology. It's you."

30. September 2020



Machine Learning Modell ≠ Produkt

Definition Meierhofer et al. (2019).

die Anwendung einer Mischung von Fähigkeiten aus Data Analytics, Engineering und Kommunikation, die darauf abzielt, **aus den Daten selbst** einen Mehrwert zu generieren, um einer anderen Einheit Nutzen zu bringen.

- «Anderer Einheit»: menschlicher User oder ein weiterer Prozess.
- Der Nutzen entsteht indem dem User Informationen (insights) über eine Schnittstelle (interface) zugänglich gemacht werden.





User

Rätsel: Wer ist gemeint?

Person 1

- Männlich
- Geboren 1948
- Aufgewachsen in England
- Verheiratet
- Hat Kinder
- Ist Erfolgreich
- Ist Reich
- Verbringt den Winterurlaub in den Alpen
- Mag Hunde

Person 2

- Männlich
- Geboren 1948
- Aufgewachsen in England
- Verheiratet
- Hat Kinder
- Ist Erfolgreich
- Ist Reich
- Verbringt den Winterurlaub in den Alpen
- Mag Hunde





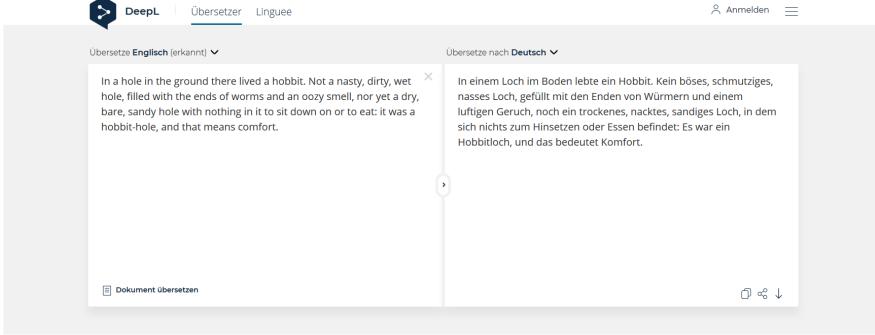
Swiss Open Data



- Open Government Data (Demographie, Umwelt, Arbeit, Politik, ...)
- Insights: Vorverarbeitete Rohdaten: Aggregation (Datenschutz), Cleansing, Transformation in ein Standardformat, ...
- Interfaces: Application Programming Interface (API):
 - HTTP GET und POST
 - Resultate im JSON Format.



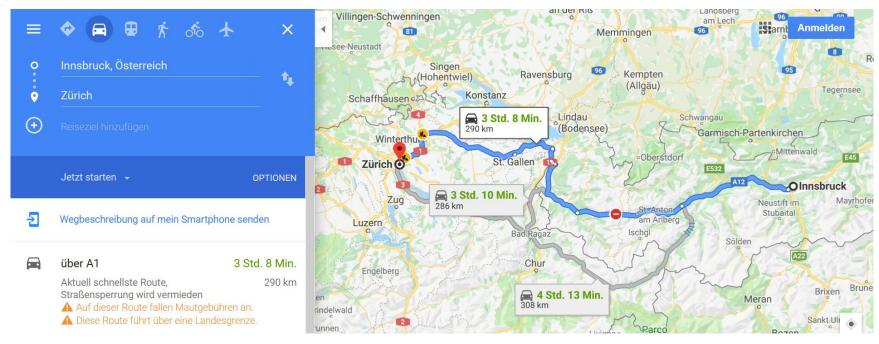
DeepL



- Online Übersetzungstool über Web Zugriff (free), API Zugriff (pro)
- Insights: Vollautomatische Übersetzung zwischen 9 Sprachen.
- Interfaces:
 - Interaktiver Web Zugriff (freie Version)
 - Rest API (in der pro Version) f
 ür die Integration in eigene Produkte (Ear-Pods, ...)



Google Maps



- Interaktive Online Karte
- Insights: Routenplanung, Stauwarnung, Baustellen, 3D-Geländemodell
- Interfaces:
 - Interaktiver Web Zugriff mit intuitiver Navigation
 - Mit Google Konto Zugriff auf weitere Inhalte



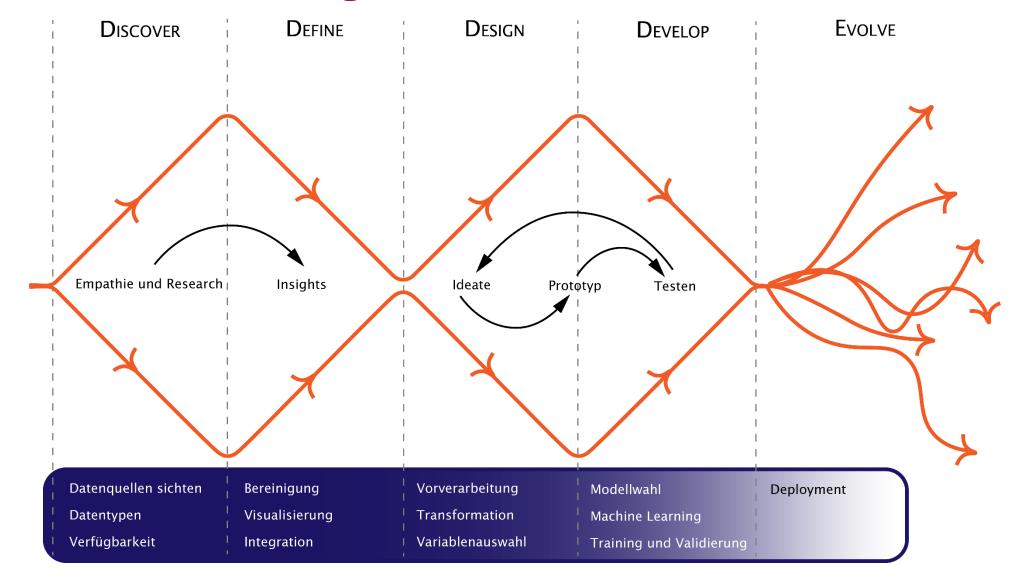
Matrix der Datenprodukte

Intelligenz beim User

Intelligenz beim System

		Rohdaten	Abgeleitete Daten	Algorithmen	Entscheidungs- unterstützung	Automatische Entscheidungs- systeme	Nach Simon O'Regan
technisch	APIs	Swiss Open Data		DeepL (pro)			gan
	Dashboards			Kompressor- Monitoring	Dividella-App		
intuitiv	Web Elements			DeepL (free)	Digitale Pathologie	Amazon's Recommender System	

Data Service Design

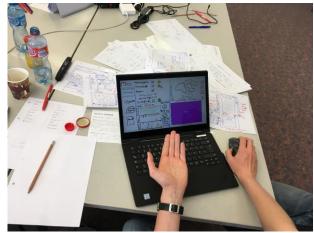




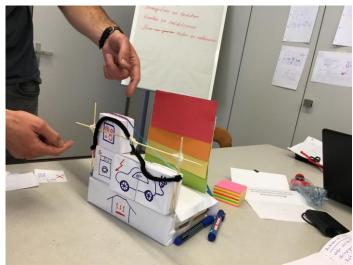
Data Service Design

Kombination von Design Thinking und Data Science

- Anwendung von Methoden des Design Thinking zur Identifikation von Kundenbedürfnissen.
- Verknüpfung mit Machine Learning zur schnellen Entwicklung von Prototypen für Daten Produkte und Services
- Gemeinsame Konzept-Entwicklung mit Gregor Kreuzer (CoWorking Gelbe Fabrik)
- Projektbegleitung bzw. Teaching









Fazit

Zurück zu Turings Frage





Fazit

Können Maschinen denken?

- Nein.
- Al in der moderene Ausprägung ist narrow Al: Spezifische Tasks, die von Maschinen aus Daten gelernt werden können.
- Trotzdem: Al und ML-basierte Entscheidungssysteme haben großes Potential für Unternehmen.
- IT-Hardware und Data-Pipelineing Grundlagen (IoT, Cloud-Computing, Datenbanksysteme, ...) sind notwendig und typischerweise eine grosse Herausforderung.
- Keep it simple. Beginnen mit einfachen Methoden, die interpretierbar und robust sind.
- Kombination von Design Thinking Methoden mit Data Science ermöglicht es,
 Kundenbedürfnisse zu erfassen und Use Cases besser zu definieren.
- · Weiterbildungsangebote nützen bzw. Mitarbeiter ausbilden.



Fragen?



