

TECHNOLOGY LAB

Wie sich das Potential von Machine Learning ausreizen lässt

Felix Vetter | InterSystems





Beispiel: Entwickeln eines Empfehlungssystems

**Üblicher Ansatz:
einfach irgendwas vorschlagen**



Probier doch mal Mayo!



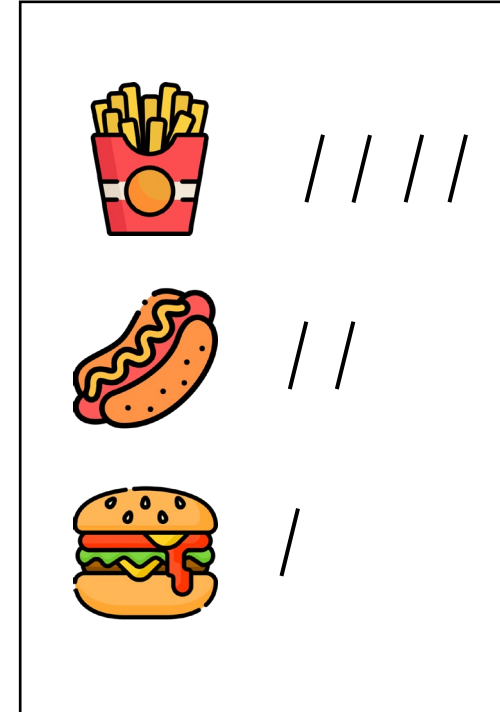


Mit System: Kaufverhalten analysieren





Mit System: Kaufverhalten analysieren



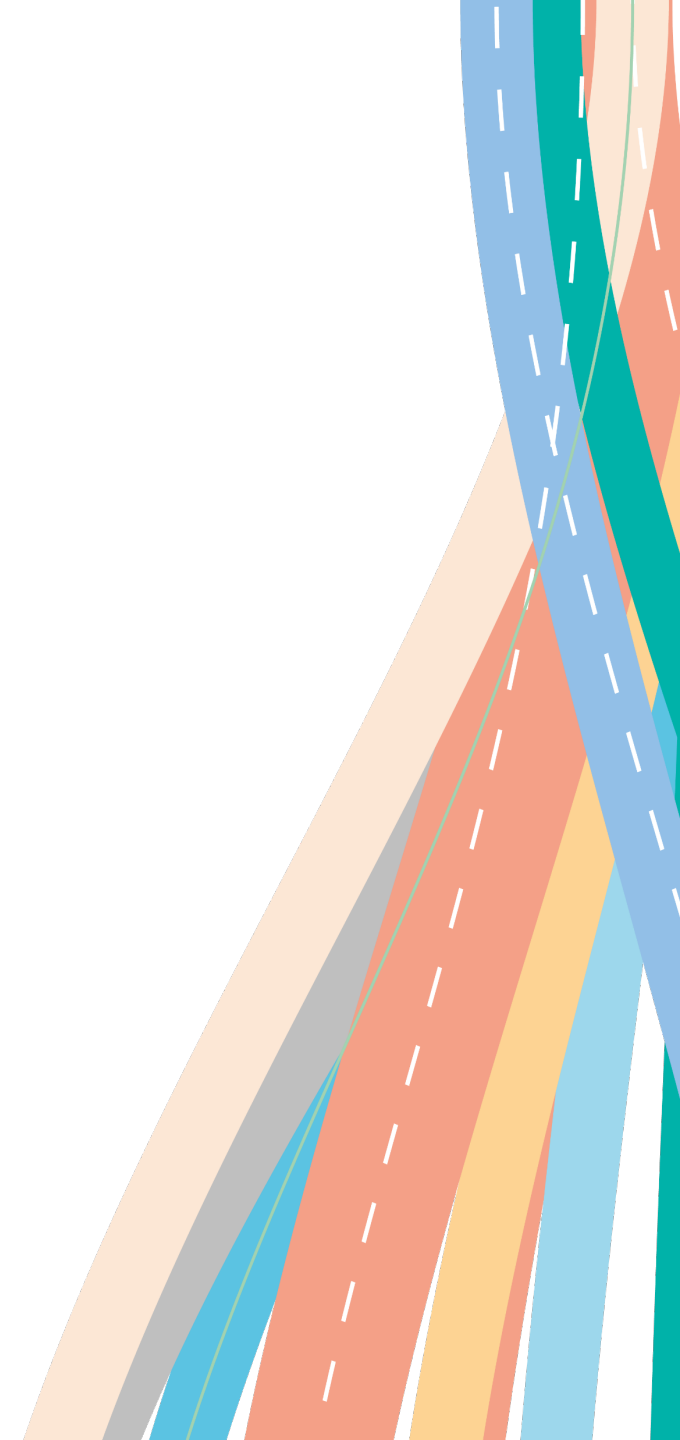


Probier doch mal
Pommes!






Besser: Produktkombinationen








Besser: Produktkombinationen

			
	////	//	///
	//	////	
	///		////

$$\begin{pmatrix} 7 & 2 & 5 \\ 2 & 4 & 0 \\ 5 & 0 & 4 \end{pmatrix}$$

Ei ma Pommes
mit ohne alles!

Willste Pommes
dazu?

$$\begin{pmatrix} 7 & 2 & 5 \\ 2 & 4 & 0 \\ 5 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \\ 5 \end{pmatrix}$$




Noch Besser: Mit Machine Learning

Machine Learning likes square data



X

=

y

id ^	gender ^	date_birth ^	hypertension ^	heart_disease ^	ever_married ^	work_type ^	zip ^	stroke ^
67	Female	10/16/2003	0	0	No	Private	98815	0
77	Female	4/5/2007	0	0	No	children	98671	0
84	Male	5/17/1965	0	0	Yes	Private	99126	0
91	Female	3/29/1978	0	0	No	Private	98382	0
99	Female	12/24/1989	0	0	No	Private	98323	0
121	Female	4/1/1982	0	0	Yes	Private	98122	0
129	Female	8/23/1996	0	0	No	Private	98465	0
132	Female	1/21/1940	0	0	Yes	Govt_job	99016	0

Machine Learning likes square data



X

=

y

id ^	gender ^	date_birth ^	hypertension ^	heart_disease ^	ever_married ^	work_type ^	zip ^	stroke ^
67	Female	10/16/2003	0	0	No	Private	98815	0
77	Female	4/5/2007	0	0	No	children	98671	0
84	Male	5/17/1965	0	0	Yes	Private	99126	0
91	Female	3/29/1978	0	0	No	Private	98322	0
99	Female	12/24/1989	0	0	No	Private	98323	0
121	Female	4/1/1982	0	0	Yes	Private	98122	0
129	Female	8/23/1996	0	0	No	Private	98465	0
132	Female	1/21/1940	0	0	Yes	Govt_job	99016	0

Jede Zeile ist eine einzelne Vorhersage

Machine Learning likes square data



$$f(x)$$

$$= y$$

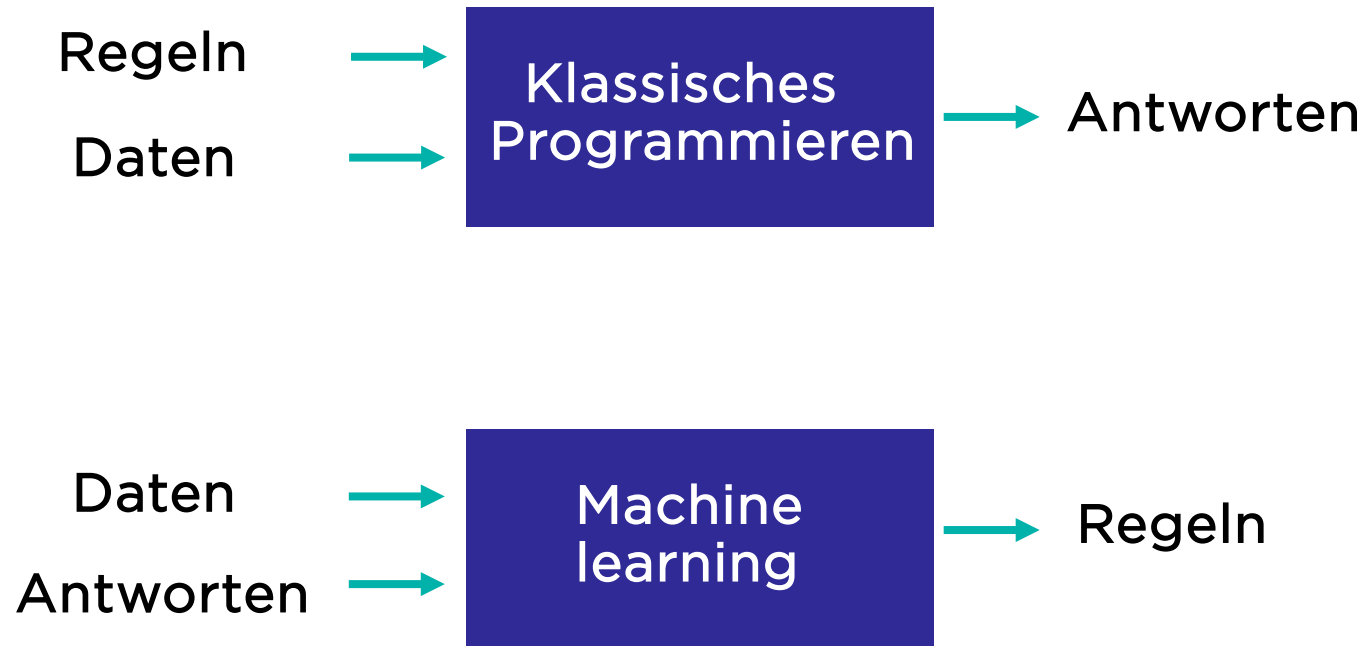
$f(x)$								y
id ^	gender ^	date_birth ^	hypertension ^	heart_disease ^	ever_married ^	work_type ^	zip ^	stroke ^
67	Female	10/16/2003	0	0	No	Private	98815	0
77	Female	4/5/2007	0	0	No	children	98671	0
84	Male	5/17/1965	0	0	Yes	Private	99126	0
91	Female	3/29/1978	0	0	No	Private	98322	0
99	Female	12/24/1989	0	0	No	Private	98323	0
121	Female	4/1/1982	0	0	Yes	Private	98122	0
129	Female	8/23/1996	0	0	No	Private	98465	0
132	Female	1/21/1940	0	0	Yes	Govt_job	99016	0

Jede Zeile ist eine einzelne Vorhersage



**Unterschied zum
Programmieren?**

Unterschied zum Programmieren?



Machine Learning Problemtypen



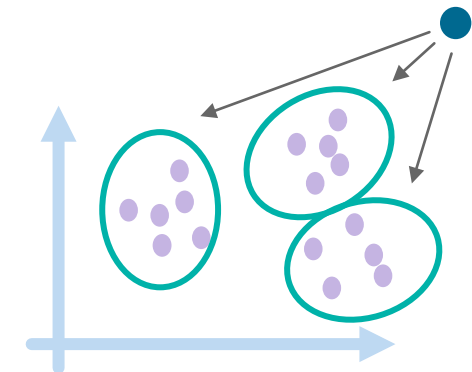
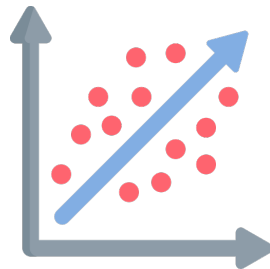
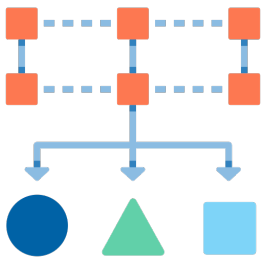
Supervised learning
(labeled Data)

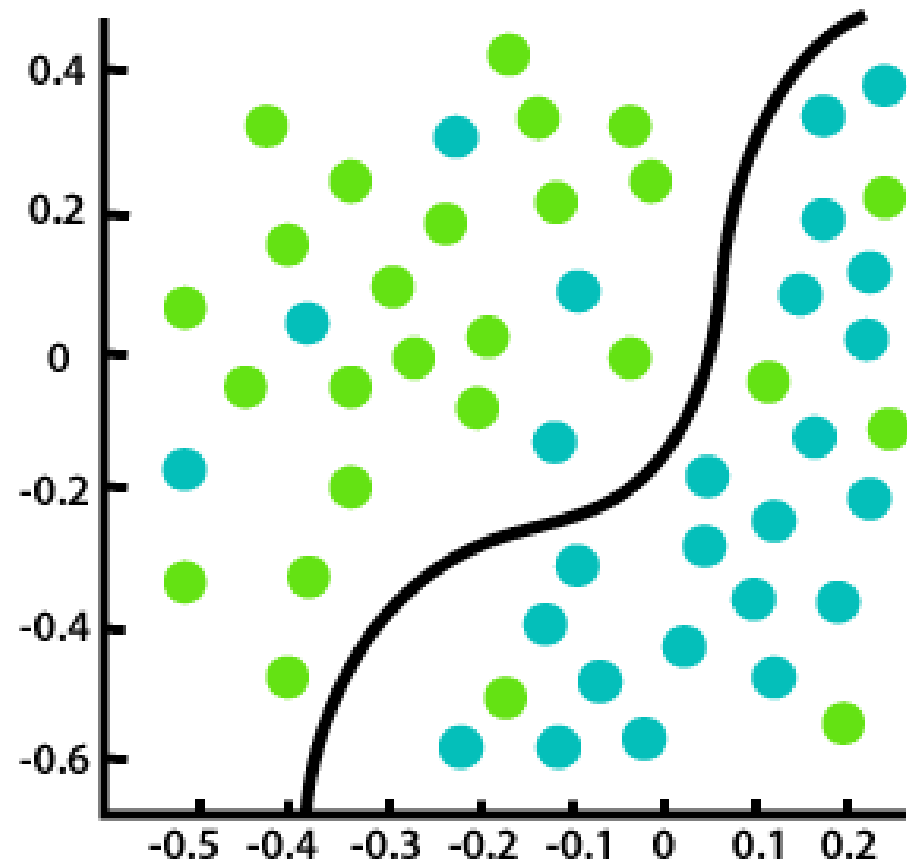
Unsupervised learning
(unlabeled data)

Classification

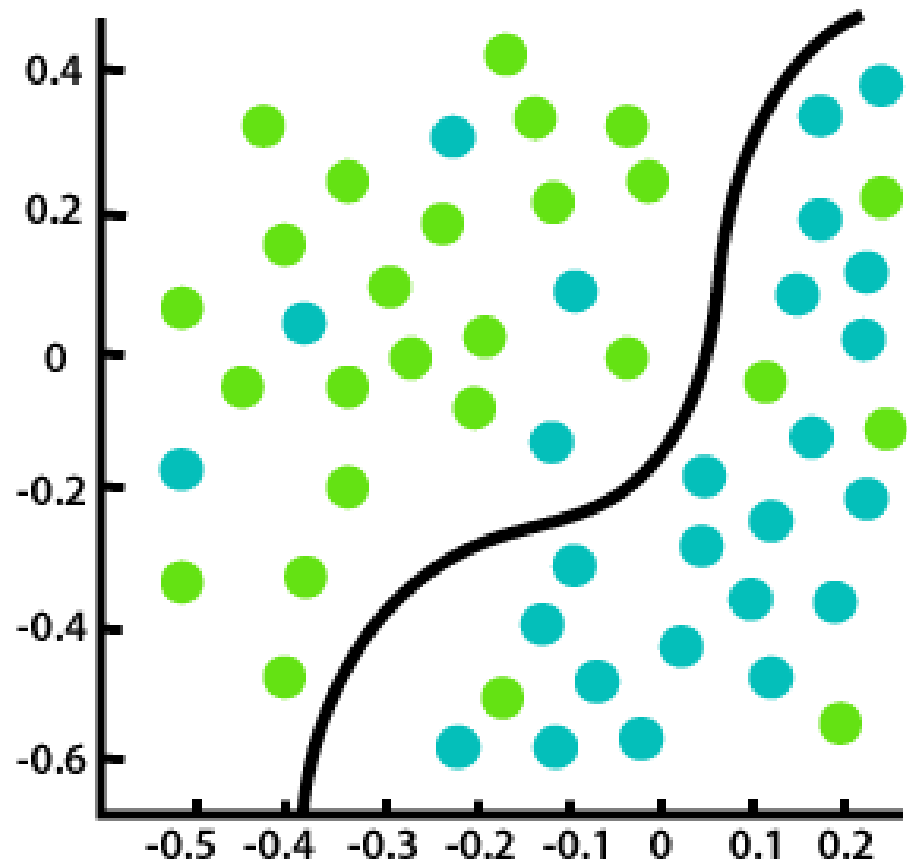
Regression

Clustering

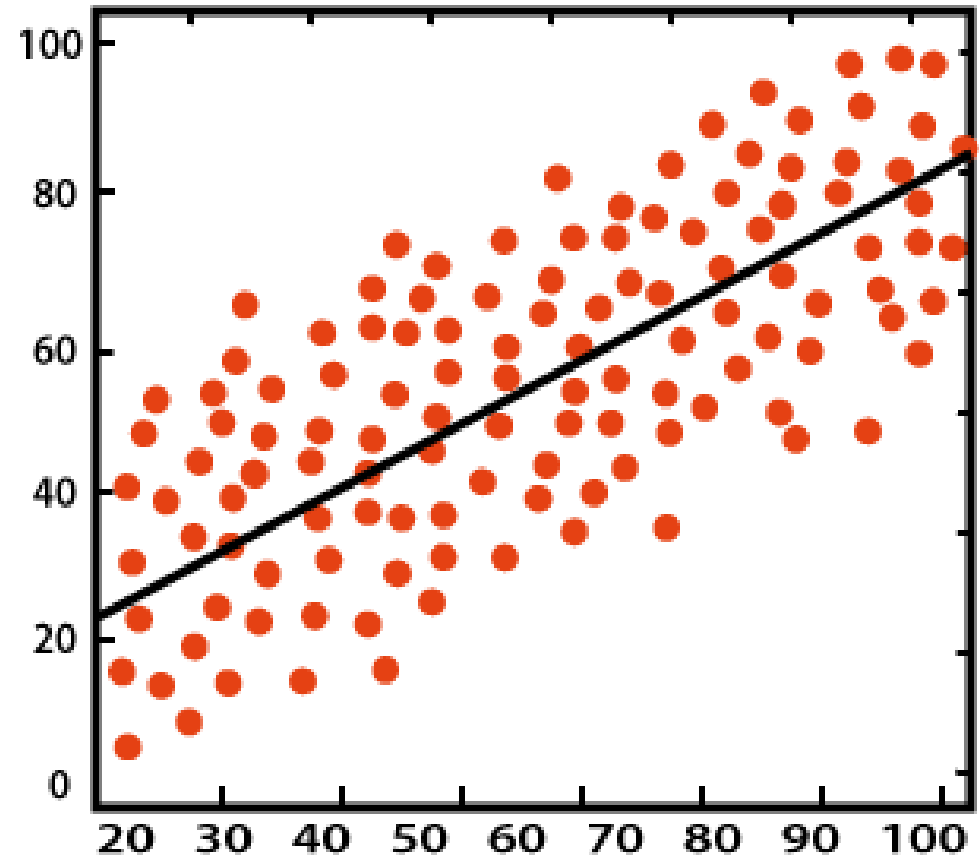




Classification

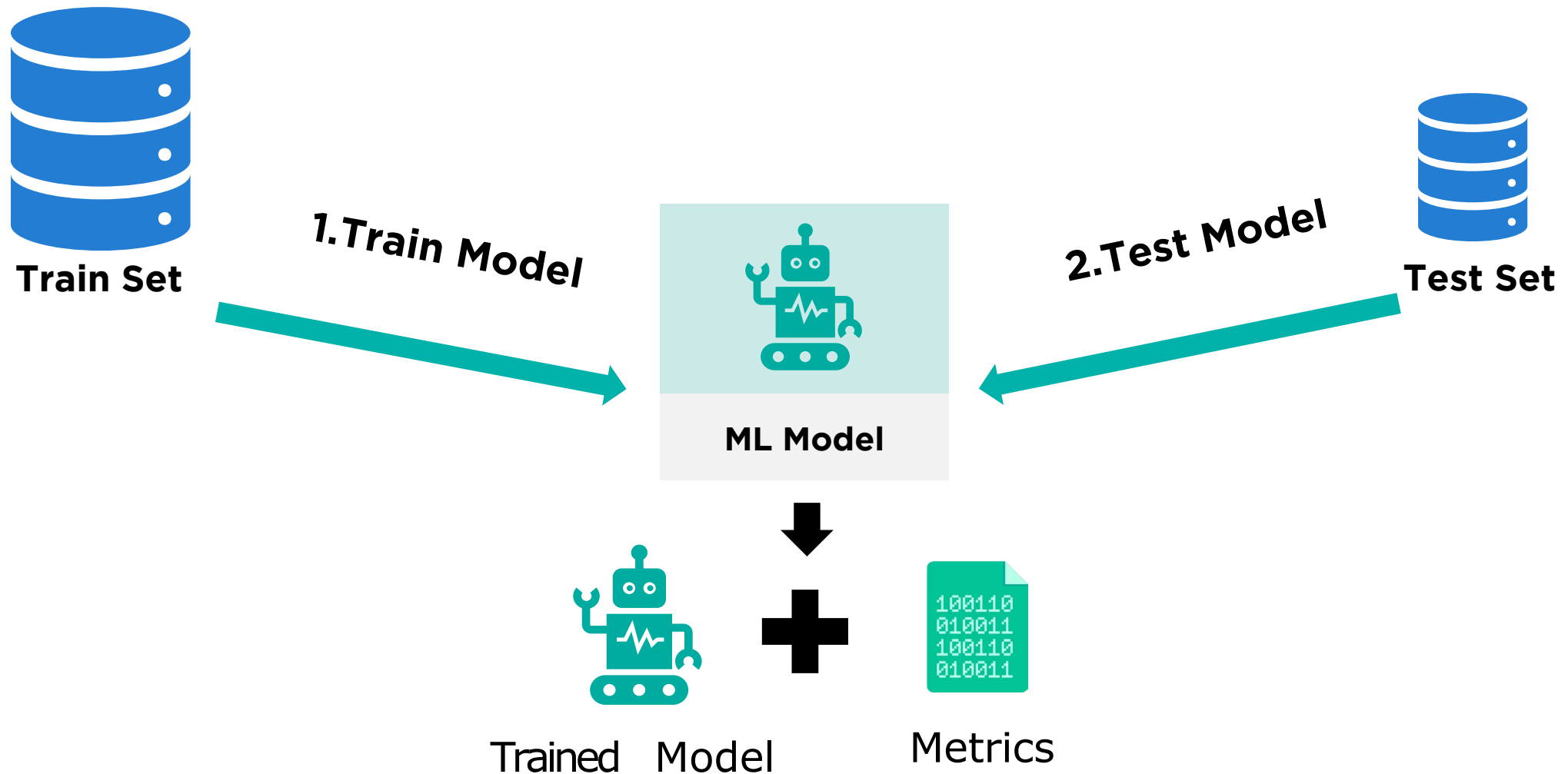


Classification

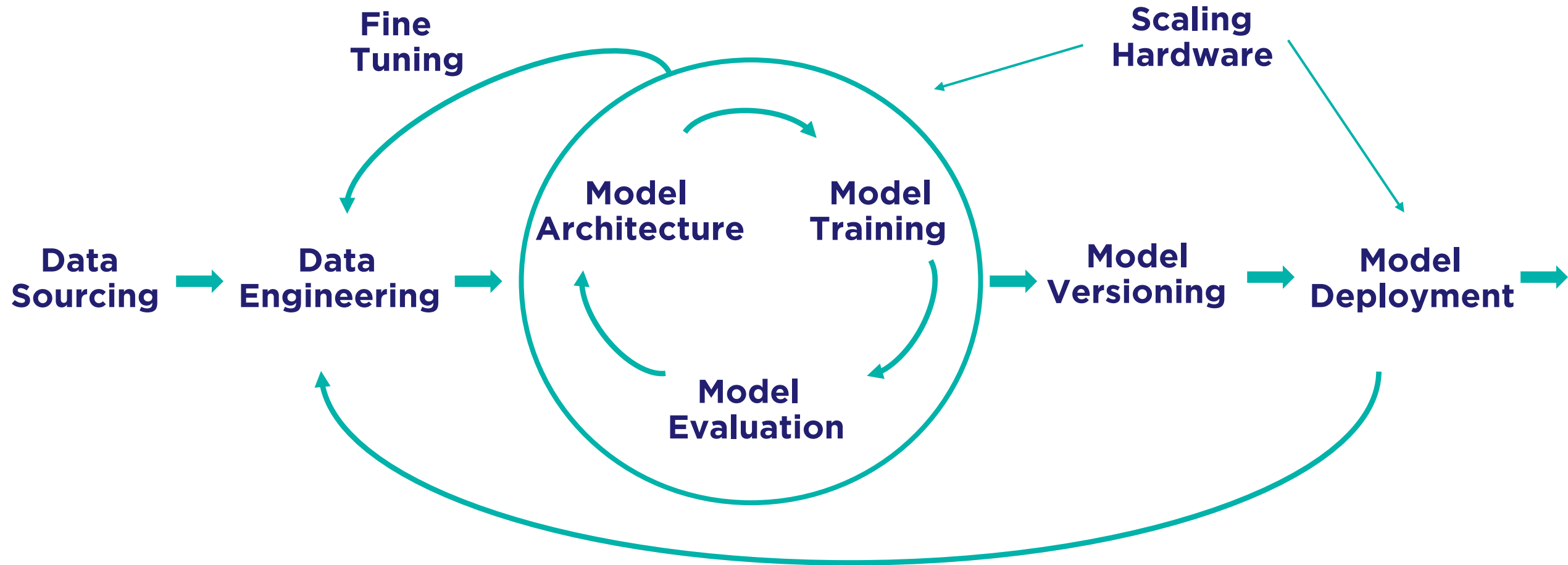


Regression

Ablauf „supervised Machine Learning“



ML Zyklus



Wie entstehen Metriken?



- Die wichtigste Quelle für die Schätzung der **Accuracy** bei **classification** ist die **Confusion Matrix**

Predicted Class	True Class	
	P	N
	P	N
P	True Positives	False Positives
N	False Negatives	True Negatives

- Accuracy**
- Precision**
- Recall**

Object
Classifier?

Perzeptron?

XGB
Regressor?

KI? Naïve Deep
Bayes? Learning?

RNN? Tensor ChatGPT? Keras? Python?

Flow? Gradienten
Random Wo liegen eigentlich die Unterschiede?

Forest? Abstieg? Neuronale Sckit
Netze

PMML? CNN? Ensemble

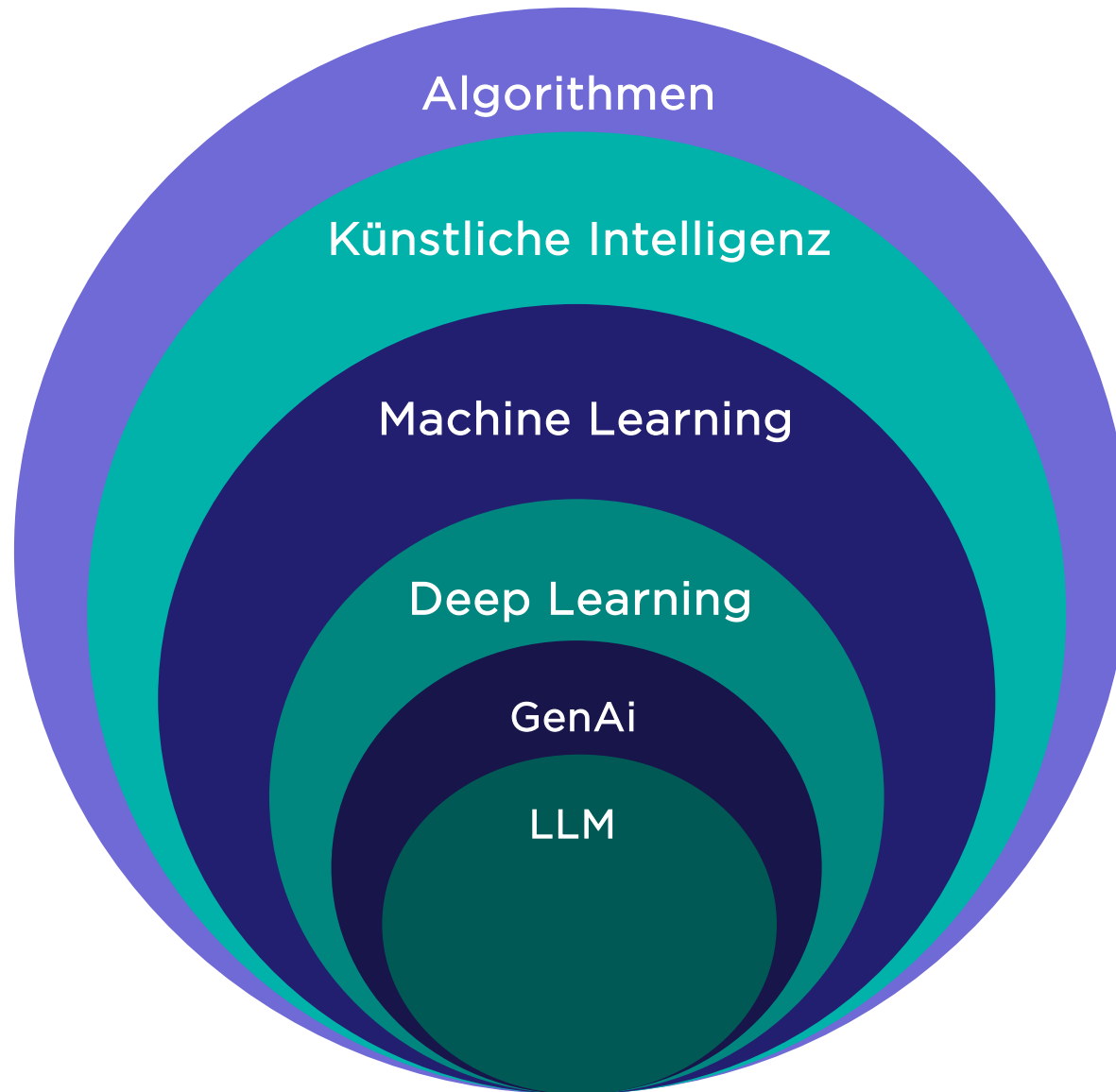
Machine Learning? NLP? Hidden Layer? Image Recognition?



Kurz durchatmen

- **3 sek EIN**
- **3 sek AUS**

Der KI Jungle mal einfach







You

Erstelle mir einen Vortrag mit Inhalt mit dem Titel "Wie sich das Potential von Machine Learning ausreizen lässt"



ChatGPT

Nein.



Message ChatGPT...



eXplainableAi

Ihr erstes Machine Learning Modell



Herd- platte	Angeschaltet	Herdplatte leuchtet	TopfAuf Herd	Induktion	Reissack FälltUm	Essen dampft	Herd Berühren
1							1
1	1	1	1				0
1	1	0	0	1			1
1	45	Kiwi	1	0			-(0)
1	1	0	0	1	1		1
1	0	0	0	-	0	1	-(0)

Hohe Datenqualität

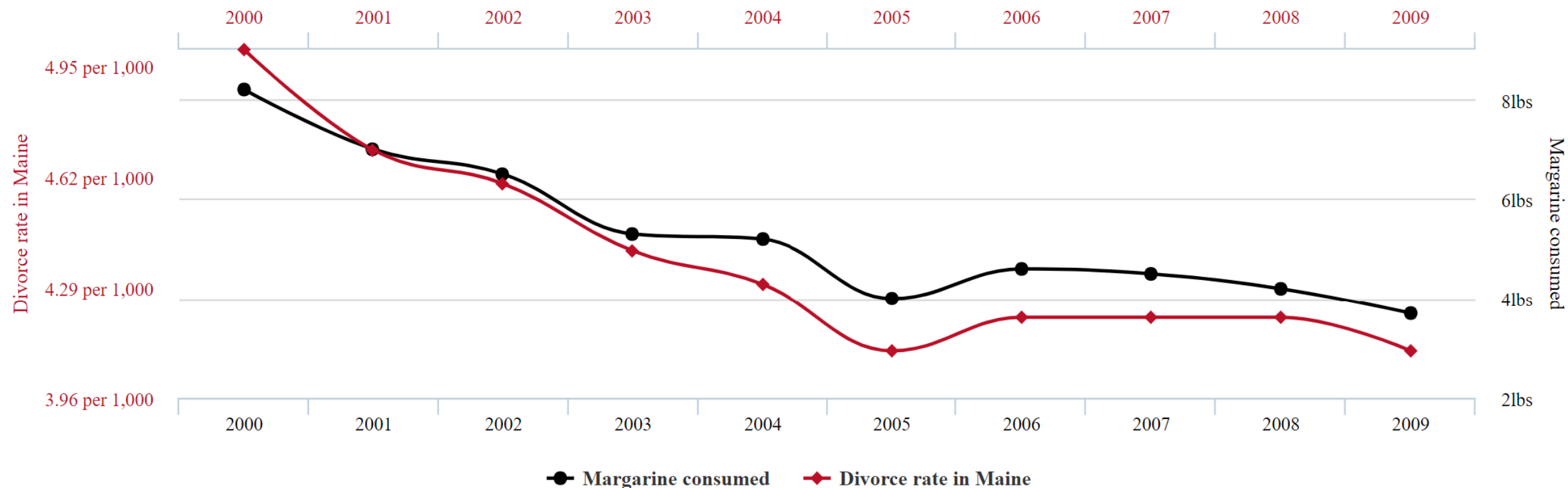


Divorce rate in Maine

correlates with

Per capita consumption of margarine

Correlation: 99.26% ($r=0.992558$)



tylervigen.com

Data sources: National Vital Statistics Reports and U.S. Department of Agriculture

Datenherkunft

Mehrere Datenquellen

- **Hohe Datenqualität**
- **Datenherkunft**
- **Mehrere Datenquellen**





**“KI
ist weder böse noch
gut, es ist das was
der Mensch daraus
macht”**



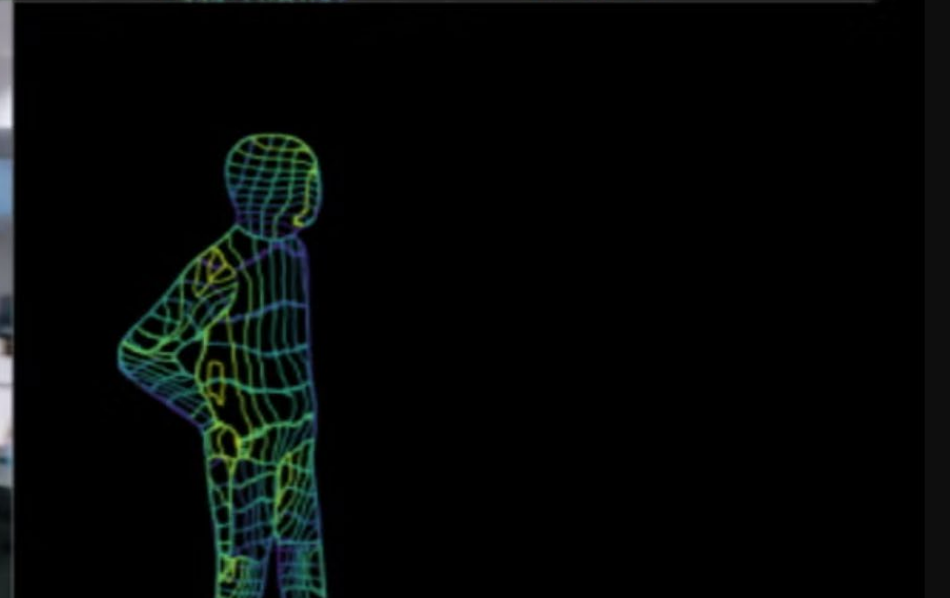
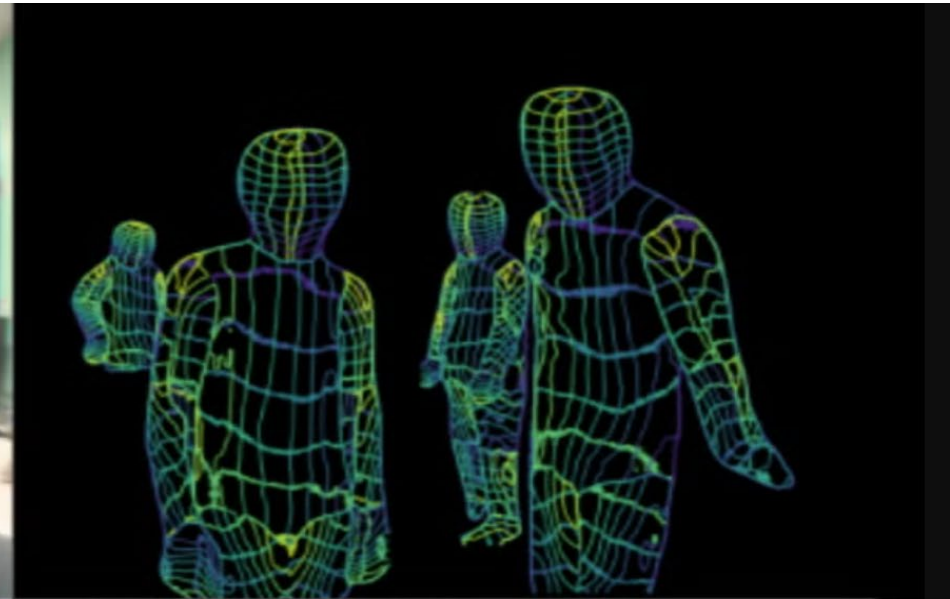
**“Gentechnik
ist weder böse noch
gut, es ist das was
der Mensch daraus
macht”**



**“Kernspaltung
ist weder böse noch
gut, es ist das was
der Mensch daraus
macht”**



**“KI
ist weder böse noch
gut, es ist das was
der Mensch daraus
macht”**



Was ist wichtig für die Arbeit mit KI?

- Was soll mit KI erreicht werden
- Welche Datenquellen brauche ich dafür
- Welche Datenmodelle brauche ich dafür
- Ethische Erwägung und Test auf Bias
- Dauerhafte Modellpflege und Überwachung der Metriken

Call to action:
Sprechen Sie uns an!



„ML ist wie Sex im Teenageralter:

- **Jeder redet darüber**
- **niemand weiß wirklich, wie man es macht**
- **jeder denkt, dass alle anderen es tun**
- **also behaupten sie, dass sie es auch tun.“**





Let's Connect

Felix Vetter | InterSystems

Sales Engineer Intern

+49 1744797827

felix.vetter@intersystems.com

