

# Schwarmintelligenz selbst programmieren

Woher weiß Netflix, was ich als nächstes gucken sollte?!

# Max Rose

— — —

Gründer, Softwareentwickler, Informatiker

B. Sc. Angewandte Informatik

Interesse: Künstliche Intelligenz

Portfolio: <https://rosemax.de>



# Ziele des Talks

---

1. Grundlagen verstehen
2. Startpunkt bekommen
3. Motivation zur Anwendung

# Gliederung

— — —

1. Relevanz von individuellen Empfehlungen in 2021
2. Vorstellung KNN Algorithmus
3. [Praxis] Film Live-Demo
4. Ethische Fragestellungen
5. Weitere Projektideen

# Relevanz von individuellen Empfehlungen in 2021

# Relevanz von individuellen Empfehlungen

---

# NETFLIX

Netflix, der bekannteste US-Online-Vermieter für Film-Silberscheiben, hat einen hoch dotierten Wettbewerb ausgeschrieben, um seinen DVD-Empfehlungsdienst zu verbessern. Eine Million Dollar soll derjenige erhalten, der den Algorithmus in Sachen Genauigkeit bis 2011 um zehn Prozent verbessert.

# Relevanz von individuellen Empfehlungen

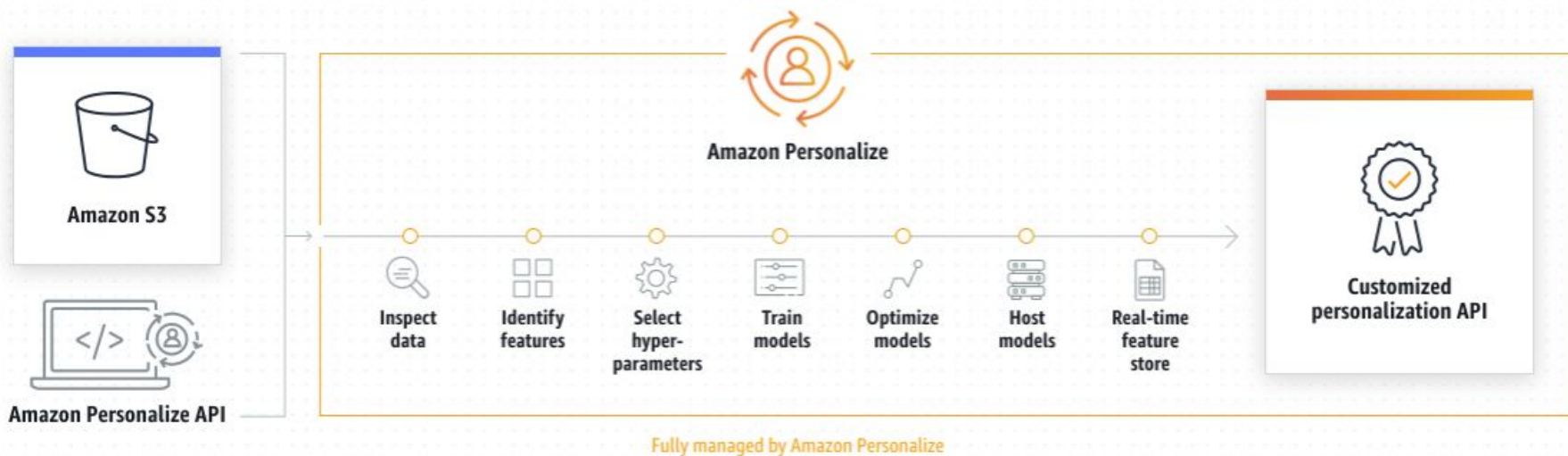
---



35% aller Verkäufe basieren auf Produktempfehlungen

Was 1997 mit einem einfachen Produktvergleich begonnen hat, ist heute Dank freiskalierbarer Rechenkapazität, Machine Learning und Big Data zu einem echten „Monstrum“ geworden, von dem mittel- und langfristig vor allem einer profitiert: Amazon.

# Amazon Personalize





# Relevanz von individuellen Empfehlungen

---

The Amazon logo, featuring the word "amazon" in a dark blue, lowercase, sans-serif font. A curved orange arrow starts under the letter 'a' and points towards the letter 'z'.The Netflix logo, consisting of the word "NETFLIX" in a bold, red, uppercase, sans-serif font.The Facebook logo, featuring the word "facebook" in a blue, lowercase, sans-serif font.The Google logo, featuring the word "Google" in its multi-colored, lowercase, sans-serif font: 'G' is blue, 'o' is red, 'o' is yellow, 'g' is blue, 'l' is green, and 'e' is red.

Alle erhöhen seit Jahren den Nutzen für den Kunden  
durch das Sammeln und Auswerten von Daten

# Mit anderen Worten...

---

“Langsam wird's peinlich, wenn man 2021 KEINE Empfehlungs-Algorithmen benutzt, um die Nutzererfahrung zu verbessern!”

# Vorstellung KNN Algorithmus

# Inhalt- vs. nutzerbasiertes Filtern

— — —



z.B.: amazon, Netflix

# Inhalts- vs. nutzerbasiertes Filtern

— — —



z.B.: amazon, Netflix



z.B.: Facebook, Google,  
Netflix

# Inhaltsbasiertes Filtern

— — —



## Problem: Kein Bezug zum Nutzer

“Sie haben Star Wars 1-3 gekauft. Das könnte sie auch interessieren”

- Star Wars 1
- Star Wars 2
- Star Wars 3
- Star Wars: The Clone Wars
- Star Trek

# Feature Vektoren

— — —



?

# Feature Vektoren

— — —



→

Titel, Dauer, Genre,  
Schauspieler,  
Bewertungen

→

“Star Wars 1”  
136 Minuten  
“Sci-Fi/Action”  
SP0001, SP0003,  
...  
4,2 ...



# Feature Vektoren



→  
Titel, Dauer, Genre,  
Schauspieler,  
Bewertungen

→  
[  
"Star Wars 1"  
136 Minuten  
"Sci-Fi/Action"  
SP0001, SP0003,  
...  
4,2 ...  
]

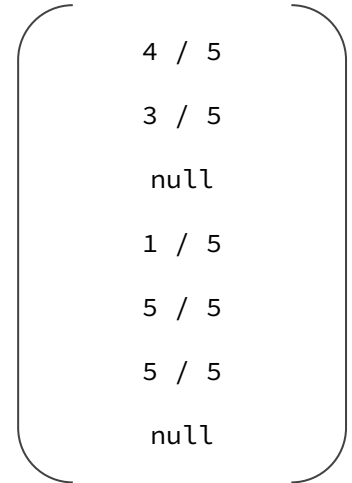
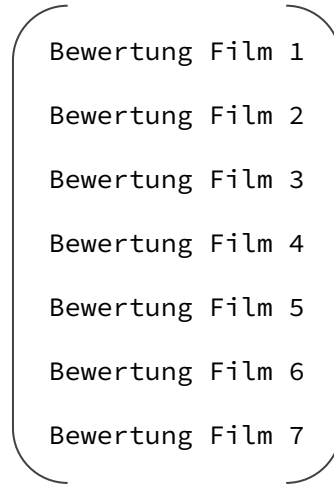


→  
Alter, Geschlecht,  
Verhalten,  
Präferenzen

→  
[  
24  
männlich  
Hat Star Wars 2  
schon 3 Mal  
geguckt  
Hat Star Wars 3  
mit 5 Sternen  
bewertet  
]

# Feature Vektoren

— — —



# Ähnliche Nutzer finden

---

$$\text{similarity}(\text{👤}, \text{👤}) = 1.0$$

# Ähnliche Nutzer finden

---

$$\text{similarity}(\text{👤}, \text{👤}) = 1.0$$

$$\text{similarity}(\text{👤}, \text{👤}) = 0.76$$

↖ Martin

# Ähnliche Nutzer finden

---

$$\text{similarity}(\text{stick figure with Martin's head}, \text{stick figure with Martin's head}) = 1.0$$

Martin


$$\text{similarity}(\text{stick figure with Martin's head}, \text{stick figure with Basti's head}) = 0.76$$

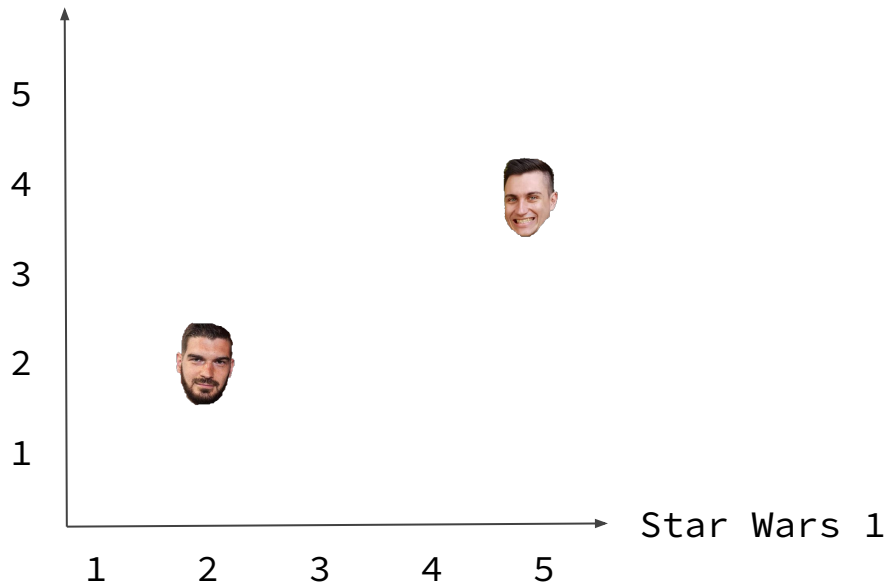
Basti


$$\text{similarity}(\text{stick figure with Basti's head}, \text{stick figure with Martin's head}) = 0.83$$

# Ähnliche Nutzer finden

— — —

Star Wars 2



$$= \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

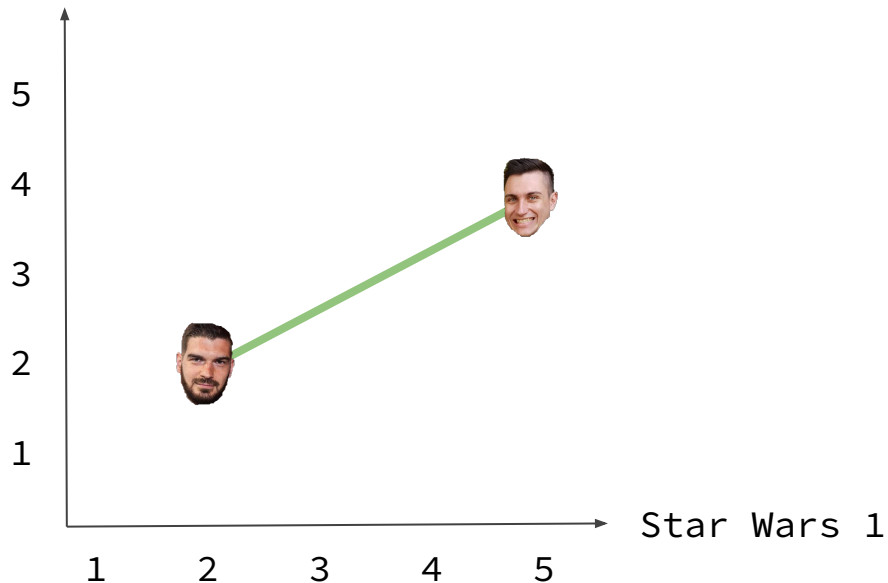


$$= \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

# Ähnliche Nutzer finden

— — —

Star Wars 2



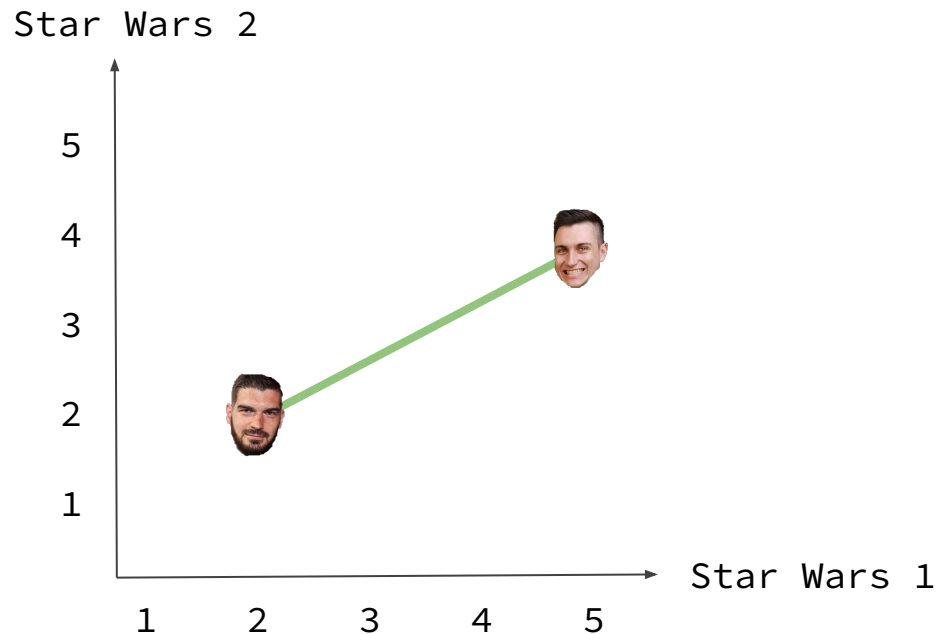
$$\text{Smiling Man} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \quad \text{Bearded Man} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$d = \sqrt{(5 - 2)^2 + (4 - 2)^2}$$

$$d \approx 3.6$$

# Ähnliche Nutzer finden

— — —



$$\text{Smiling Man} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \quad \text{Bearded Man} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$d = \sqrt{(5 - 2)^2 + (4 - 2)^2}$$

$$d \approx 3.6$$

$$\text{sim} = \frac{1}{1 + d}$$

$$\text{sim} = 0.217$$



# Funktionsweise von KNN

— — —

**KNN (int k, FeatureVector F):**

L = []

Für alle Nutzer N:

L[N] = *similarity*(F, N)

Sortiere Liste L absteigend

**return** k ersten Nutzer aus L

# Funktionsweise von KNN

— — —

**KNN** (int k, FeatureVector ):

L = []

Für alle Nutzer N:


L[N] = *similarity*(; N)


Sortiere Liste L absteigend

**return** k ersten Nutzer aus L


# Vorschläge erzeugen

— — —

 $= \begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \end{pmatrix}$

 $= \begin{pmatrix} 4 \\ 5 \\ 2 \\ 5 \end{pmatrix}$

↩ Basti

 $= \begin{pmatrix} 5 \\ 5 \\ ? \\ ? \end{pmatrix}$

# Vorschläge erzeugen

— — —

$$\text{Face 1} = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \end{pmatrix} \quad \text{Face 2} = \begin{pmatrix} 4 \\ 5 \\ 2 \\ 5 \end{pmatrix}$$

↖ Basti

$$\text{Stick Figure} = \begin{pmatrix} 5 \\ 5 \\ ? \\ ? \end{pmatrix}$$

$$\text{KNN}(2, \text{Face 1}) \rightarrow \{ \text{Face 1}, \text{Face 2} \}$$

# Vorschläge erzeugen

---

$$\text{[Face 1]} = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \end{pmatrix} \quad \text{[Face 2]} = \begin{pmatrix} 4 \\ 5 \\ 2 \\ 5 \end{pmatrix}$$

↖ Basti

$$\text{[Stick Figure]} = \begin{pmatrix} 5 \\ 5 \\ ? \\ ? \end{pmatrix}$$

$$\text{KNN}(2, \text{[Face 1]}) \rightarrow \{ \text{[Face 1]} \text{ [Face 2]} \}$$

$$\frac{\begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 2 \\ 5 \end{pmatrix}}{2} = \begin{pmatrix} 4.5 \\ 4.5 \\ 2.5 \\ 4.5 \end{pmatrix}$$

# Vorschläge erzeugen

---

$$\text{[Face 1]} = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \end{pmatrix} \quad \text{[Face 2]} = \begin{pmatrix} 4 \\ 5 \\ 2 \\ 5 \end{pmatrix}$$

$$\text{KNN}(2, \text{[Face 1]}) \rightarrow \{ \text{[Face 1]} \text{ [Face 2]} \}$$

$$\text{[Stick Figure]} = \begin{pmatrix} 5 \\ 5 \\ ? \\ ? \end{pmatrix}$$

$$\frac{\begin{pmatrix} 5 \\ 4 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 2 \\ 5 \end{pmatrix}}{2}$$

$$= \begin{pmatrix} 4.5 \\ 4.5 \\ 2.5 \\ \mathbf{4.5} \end{pmatrix}$$

Film 4  
zuerst  
vorschlagen

# [Praxis] Film Live-Demo

# Ethische Fragestellungen



# Ethische Fragestellungen

---

1. Hat das überhaupt einen Nutzen?
2. Würde ich meinen Dienst selbst nutzen?

# Ethische Fragestellungen

---

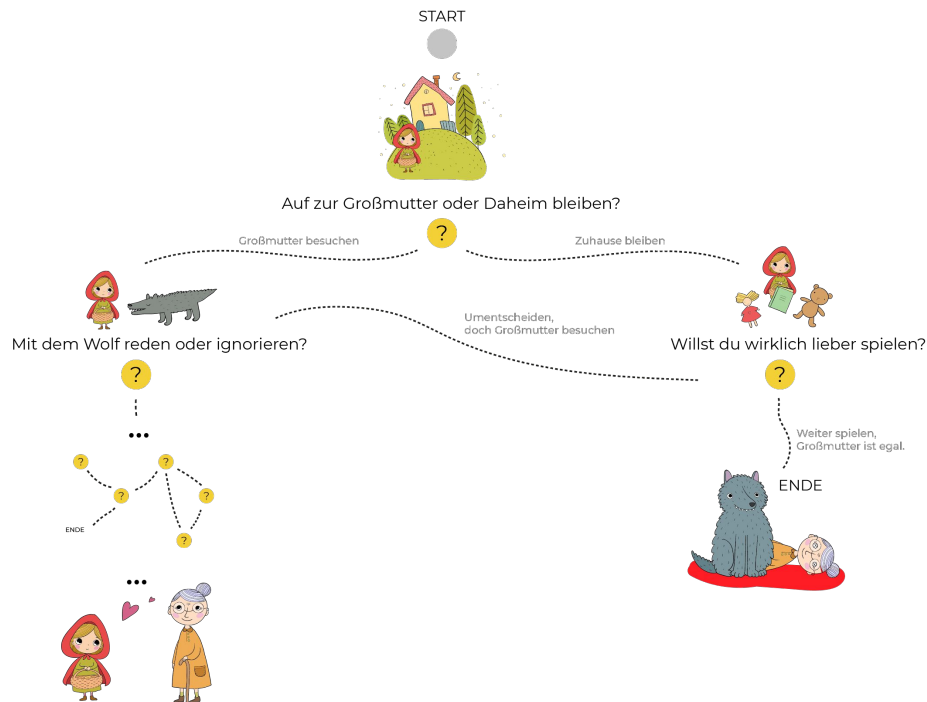
Welches Ziel verfolgst du?

- Gewinnmaximierung
- Nutzen für den Kunden erhöhen z.B.: Fehlkäufe vermeiden

# Weitere Projektideen

# audory - Deine Plattform für interaktive Hörbücher

— — —



# Empfehlungssystem bei audory

— — —



# Empfehlungssystem bei audory



# Empfehlungssystem bei audory

---

Nutzer die sich ähnlich entscheiden, bekommen ähnliche Hörbücher vorgeschlagen

Dieses Sammeln von Daten kann von jedem Nutzer zu jeder Zeit beendet werden!

Wir nutzen diese Daten ausschließlich zur Vermeidung von Fehlkäufen

# Beispiele wie Instagram, Tinder, Amazon, Spotify

---

## **Wie sieht der Feature Vektor aus?**

Instagram: Freunde, Likes, Kommentare, Watchtime

Tinder: Matches, Swipes, Wen findest du attraktiv

Amazon: Käufe, Wunschzettel, Was guckst du dir an

Spotify: Welche Künstler, Genre & Songs hörst du am häufigsten



# Diskussion



max.rose

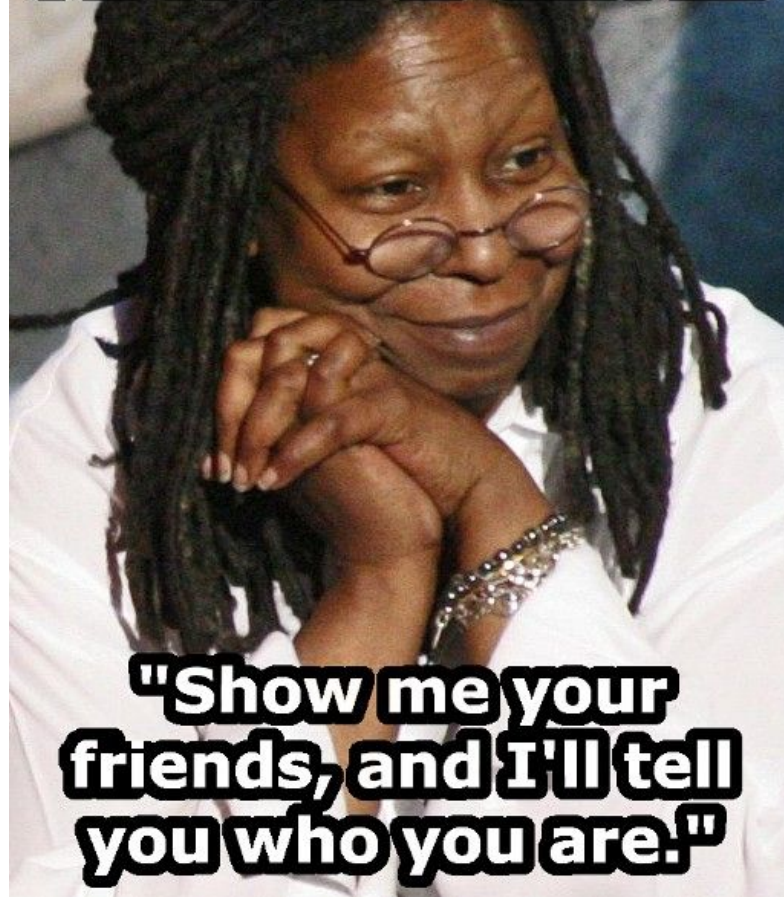


xamesor



Max Rose

**KNN BE LIKE**



**"Show me your  
friends, and I'll tell  
you who you are."**