

Which car fits my life?

Mobile.de's approach to recommendations

PyData Berlin, July 1st, 2017

Florian Wilhelm, Arnab Dutta



mobile.de

WELCHER IST DEIN NÄCHSTER?

Introduction



Dr. Arnab Dutta

Data Scientist

mobile.de GmbH

 @kopfhohen

 kraktos



Dr. Florian Wilhelm

Data Scientist

inovex GmbH

 @FlorianWilhelm

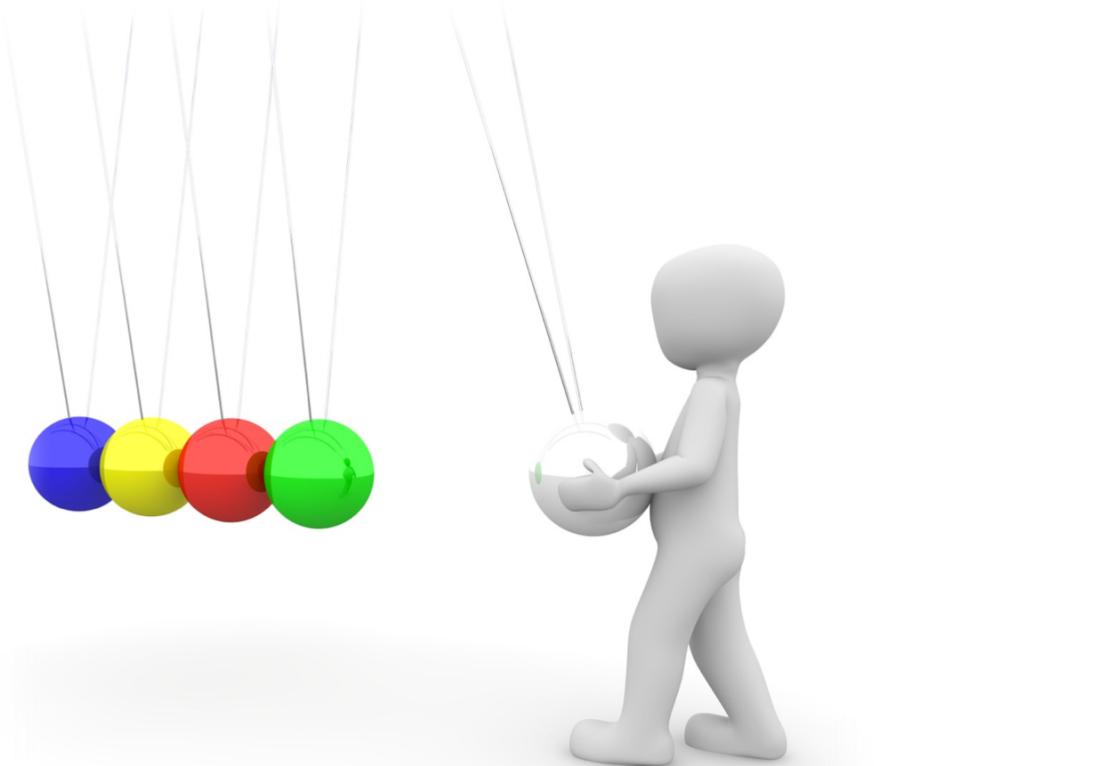
 FlorianWilhelm

 florianwilhelm.info



Outline

- **Introduction**
- Use-cases
- Theory
- Our Approach
- Example
- Outlook



Part of
ebay Tech

MOBILE.DE
GERMAN MARKET
LEADER

13.5 MIO
UNIQUE USER
PER MONTH

1.6 MIO
VEHICLES

290
EMPLOYEES

**DREILINDEN /
FRIEDRICHSHAIN
BERLIN
HEADQUARTERS**



inovex



IT-project house for digital transformation:

- ▶ Agile Development & Management
- ▶ Web · UI/UX · Replatforming · Microservices
- ▶ Mobile · Apps · Smart Devices · Robotics
- ▶ Big Data & Business Intelligence Platforms
- ▶ Data Science · Data Products · Search · Deep Learning
- ▶ Data Center Automation · DevOps · Cloud · Hosting
- ▶ Trainings & Coachings

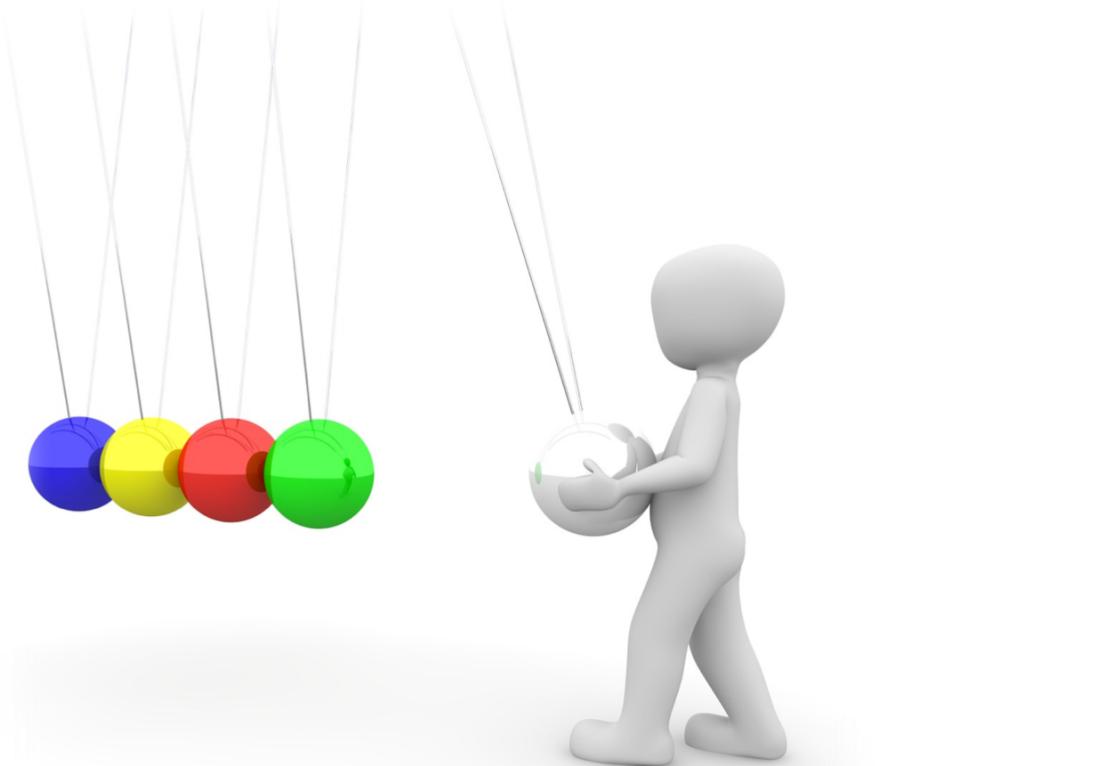
inovex offices in
Karlsruhe · Pforzheim ·
München · Köln · Hamburg.

www.inovex.de

Using technology to inspire our
clients. *And ourselves.*

Outline

- Introduction
- **Use-cases**
- Theory
- Our Approach
- Example
- Outlook



Why Recommendations?



Händlerbereich Kontakt Deutsch Meine Suchen (0) Parkplatz (1) Mein mobile.de

mobile.de Deutschlands größter Fahrzeugmarkt Suchen Verkaufen Informieren

FIAT

DIE FIAT TIPO FAMILIE
AB **99 € MTL.¹** FINANZIEREN
MIT **0% FINANZIERUNG!¹**

3 MODELLE – 1 RATE!

AB **99 € MTL.¹** FINANZIEREN
MIT **0% FINANZIERUNG!¹**

Der Trip Deines Lebens ruft!
Lass ihn nicht warten –
hol Dir Deinen neuen Begleiter.
Jetzt finden

Marke: Beliebig Modell:
Preis bis: Beliebig Erstzulassung ab: Beliebig
Kilometer bis: Beliebig Kraftstoffart: Beliebig ändern
Land: Beliebig Ort oder PLZ eingeben:
Detailsuche: 1.672.918 Treffer
Helfen Sie uns, unsere Suche zu verbessern [Jetzt teilnehmen](#)

+ INKL. 1.260 € BONUS¹
+ INKL. 4 JAHRE FIAT GARANTIE²
+ NUR BIS ZUM 30. JUNI

ZUM ANGEBOT

Interessiert am VW Golf?
Alle Infos aus einer Hand.
[Alles zum VW Golf 7](#)

Welches Auto passt zu mir?
Finde Dein passendes Auto.
[Kaufberatung starten](#)

Fragen rund ums Auto?
Hier gibt es die Antworten.
[Jetzt zu MOTOR-TALK](#)

DER FIAT TIPO 5-TÜRER, KOMBI ODER LIMOUSINE
3 MODELLE – 1 RATE!

Zuletzt gesehen

 Mercedes-Benz C 200 C Autonom./MapPilot/LED ILS/Park-P-/ 32.888 € EZ 03/2017, 3.500 km, Benzin, 135 kW (184 PS), Automatik 5,4 l/100km (komb.), 125 g CO ₂ /km (komb.)*	 BMW X5 3.0 d Vollausstattung 5.000 € EZ 10/2001, 290.000 km, Diesel, 135 kW (184 PS), Automatik ≈ 9,7 l/100km (komb.), ≈ 259 g CO ₂ /km (komb.)	 BMW X5 3.0 d 5.499 € EZ 03/2002, 252.000 km, Diesel, 135 kW (184 PS), Automatik ≈ 9,7 l/100km (komb.), ≈ 259 g CO ₂ /km (komb.)	 Volkswagen Polo 1.4 Automatik*2.Hand*Klima*SH*TÜ V*Sparsam* 5.490 € EZ 01/2006, 113.000 km, Benzin, 55 kW (75 PS), Automatik ≈ 7,5 l/100km (komb.), ≈ 180 g CO ₂ /km (komb.)
--	---	---	--

Meldungen **Tests** **Ratgeber**

Why Recommendations?

User Benefits

- engagement
- inspiration
- relevance



Business Benefits

- high click-through-rate
- small exit- & bounce-rates

Mobile.de Conversion Funnel



Recommendations on Home

Home

SRP

View

Wishlist

Contact

Buy

The screenshot shows the mobile.de homepage with a search interface for vehicles. On the left, there are filters for vehicle type (cars, motorcycles, vans, trucks), vehicle status (New car, Used car, Registration permit), brand (Any), model (Any), price (Any), registration date (Any), kilometers (Any), fuel type (Any), country (Any), postal code (Any), and radius (Any). Below these filters, a button says "Hier starten". To the right is a map showing the locations of 1,501,153 offers. At the bottom of the page, a section titled "Passend zu Deinen geparkten Fahrzeugen." displays three recommended car models:

Car Model	Variant	Price
Volkswagen Golf	Variant 1.6 TDI DPF Euro 6	€6,999
Volkswagen Golf	1.6 TDI Kombi Klima	€6,248
Volkswagen Golf	VII.6 GTDI Variant	€6,950

An orange arrow points from the text "Recommendations based on preferences of visiting users as an alternative entry point." to this section.

Recommendations based on
preferences of visiting users as an
alternative entry point.

Recommendations on Search Results Page

mobile.de

Home

SRP

View

Wishlist

Contact

Buy

Recommendations based
on similar vehicle make
and model id to present
alternatives



m
vergleichbares Fahrzeug
BMW 3er Reihe

[zu Suche hinzufügen](#)

 Audi A4 Avant 2.5 TDI EZ 05/2003, 279.000 km, 120 kW (163 PS) Kombi, Diesel, Schaltgetriebe, 4/5 Türen ≈ 6,9 l/100km (komb.), ≈ 186 g CO ₂ /km (komb.) DE-51145 Köln, Privatanbieter Finanzierung , Versicherung Parken
 Neu Audi A4 Avant 1.8 T EZ 06/2001, 236.800 km, 110 kW (150 PS) Kombi, Unfallfrei , Benzin, Schaltgetriebe, HU 09/2018, 4/5 Türen ≈ 8,2 l/100km (komb.), ≈ 197 g CO ₂ /km (komb.) DE-49716 Meppen, Privatanbieter Finanzierung , Versicherung Parken
 Audi audi a4 b6 8E 1.8T 200ps EZ 11/2001, 234.000 km, 147 kW (200 PS) Limousine, Benzin, Schaltgetriebe, HU 11/2017 DE-94481 Grafenau, Privatanbieter Finanzierung , Versicherung Parken

11



Recommendations on View Item Page

Home

SRP

View

Wishlist

Contact

Buy

Volkswagen Golf GTI Clubsport S

59.999 € (Brutto)

50.419 € (Netto) 19,00% MwSt.

Anzeige: Kfz-Versicherung: Jetzt berechnen und bis zu 320 € sparen

[Finanzierungsvergleich](#)

Technische Daten

Fahrzeugzustand

Kategorie

Herkunft

Kilometerstand

Hubraum

Leistung

Kraftstoffart

Verbrauch

CO2-Emissionen

Zugr.-Igd. Treibstoffart

Energieeffizienzklaasse

CO2-Effizienz

Unfallfrei

Sportwagen / Coupé, Tageszulassung

Deutsche Ausführung

5 km

1.984 cm³

228 kW (310 PS)

Benzin

6,0 l/100km (kombiniert)*

7,5 l/100km (innerorts)*

5,1 l/100km (außerorts)*

139 g/km (kombiniert)*

Benzin Super

E

Auf der Grundlage der gemessenen CO2-Emission unter Berücksichtigung der Masse des Fahrzeugs ermittelt.



Ähnliche Fahrzeuge



Volkswagen VOLKSWAGEN
Golf GTI Clubsport 2.0 S -

Volkswagen Golf VII GTI
Clubsport DSG ABT 340PS

Volkswagen Golf GTI
Clubsport S Nr.033/400

Recommendations based on the specific make and model a user is viewing to present alternatives

Recommendations on your Wishlist

mobile.de

Home

SRP

View

Wishlist

Contact

Buy

Mein Parkplatz (3)

Nicht mehr verfügbar
Audi A3 Sportback 2.0 TDI 2x S line Navivorb Ambition
EZ 11/2015, 17.600 km, Diesel, Schaltgetriebe, 4/5 Türen, Limousine
Geparkt am 2.5.2016 für 31.980,00 €

Ähnliche Fahrzeuge
Audi A3 1.6TDI Sportback Ambition S tronic
EZ 12/2014, 42.800 km, Diesel, 81 kW (110 PS), Automatik, 4/5 Türen, Kombi
= 4.0 l/100km (komb.), ≈ 105 g CO₂/km (komb.)

Nicht mehr verfügbar
Volkswagen Golf VI Variant Trendline BlueMotion
EZ 04/2012, 176.000 km, Diesel, Schaltgetriebe, 4/5 Türen, Kombi
Geparkt am 14.11.2016 für 6.699,00 €

Audi A3 Sportback 2.0 TDI Quattro S-Line Xenon+ B&O
EZ 10/2014, 141.540 km, Diesel, 135 kW (184 PS), Automatik, 4/5 Türen, Limousine
≈ 124 g CO₂/km (komb.)
Geparkt am 27.2.2017 für 21.480,00 €

Andere Nutzer fanden auch folgende Fahrzeuge interessant
Audi A3 1.6 TDI Limousine (clean diesel) S tronic...
Audi A3 Sportback 2.0 TDI Ambition LED Navi PDC B&O S
Audi A3 2.0 TDI Sportback 2x S-Line Bi-Xenon Navi Pan
Audi A3 2.0TDI Sportback quattro Aut S line Sportpaket



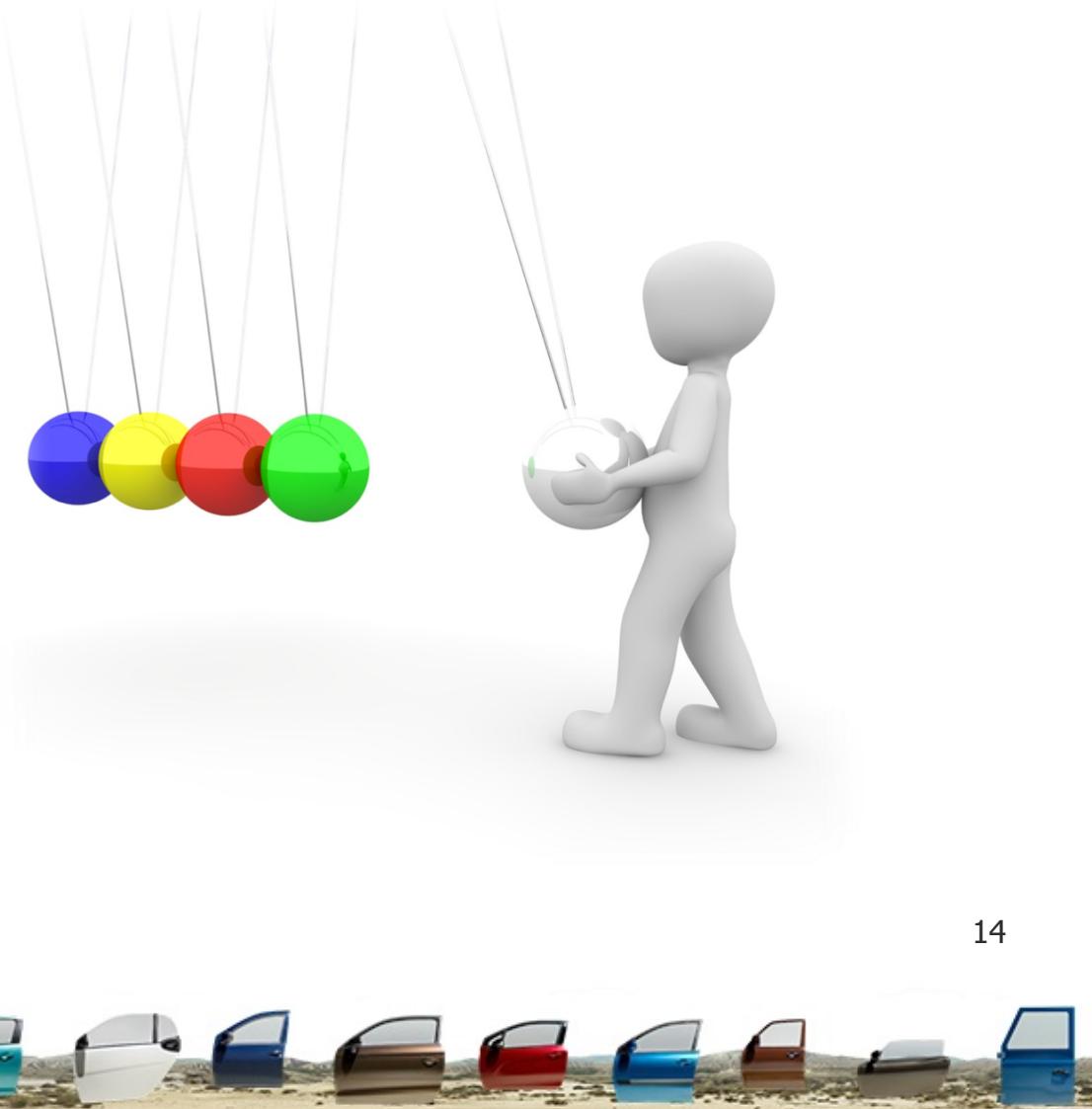
P Parken

Recommendations based on the specific **make and model** of a **deleted ad** to provide almost identical recommendations

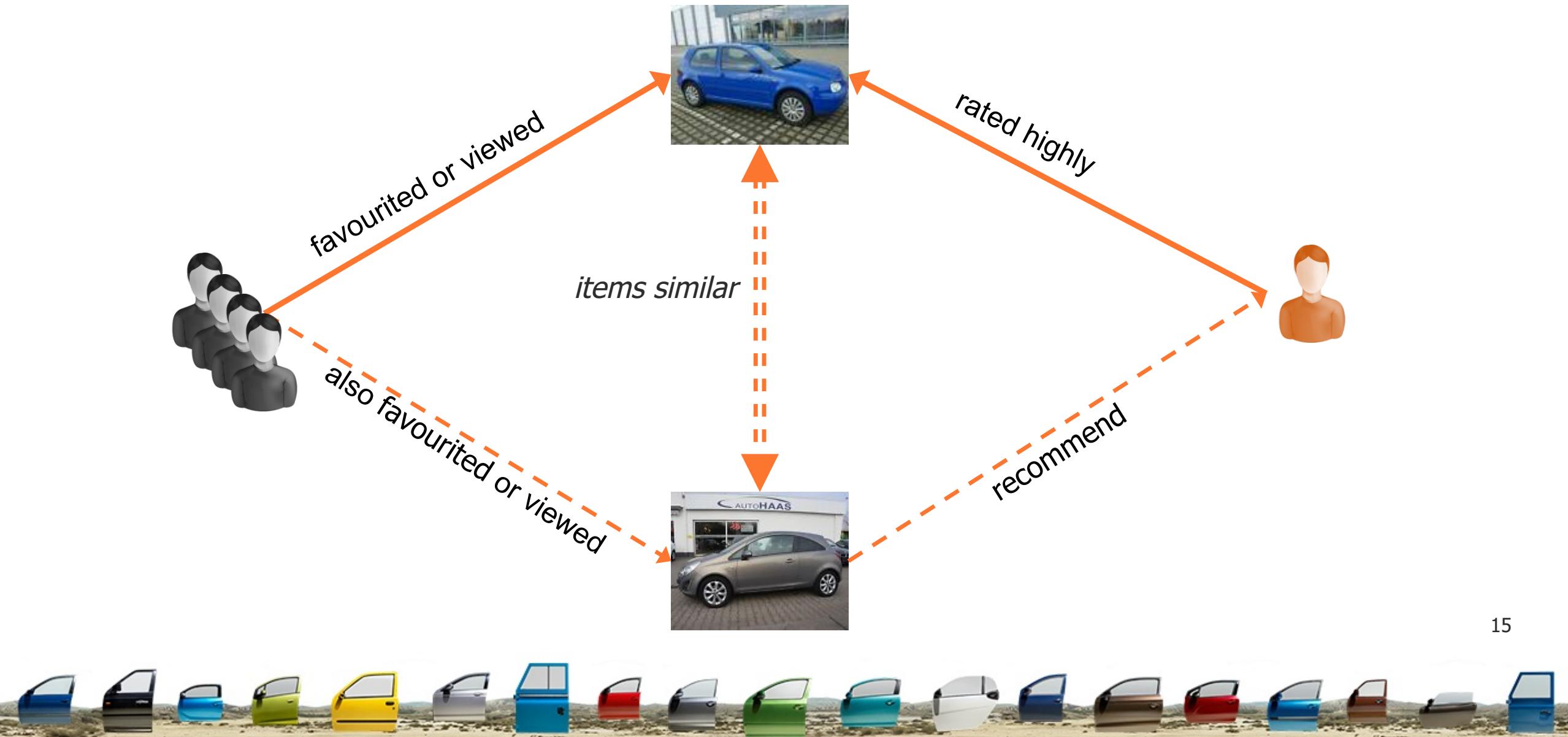
Recommendations based on the **users car preferences** and the **parking lot items**.



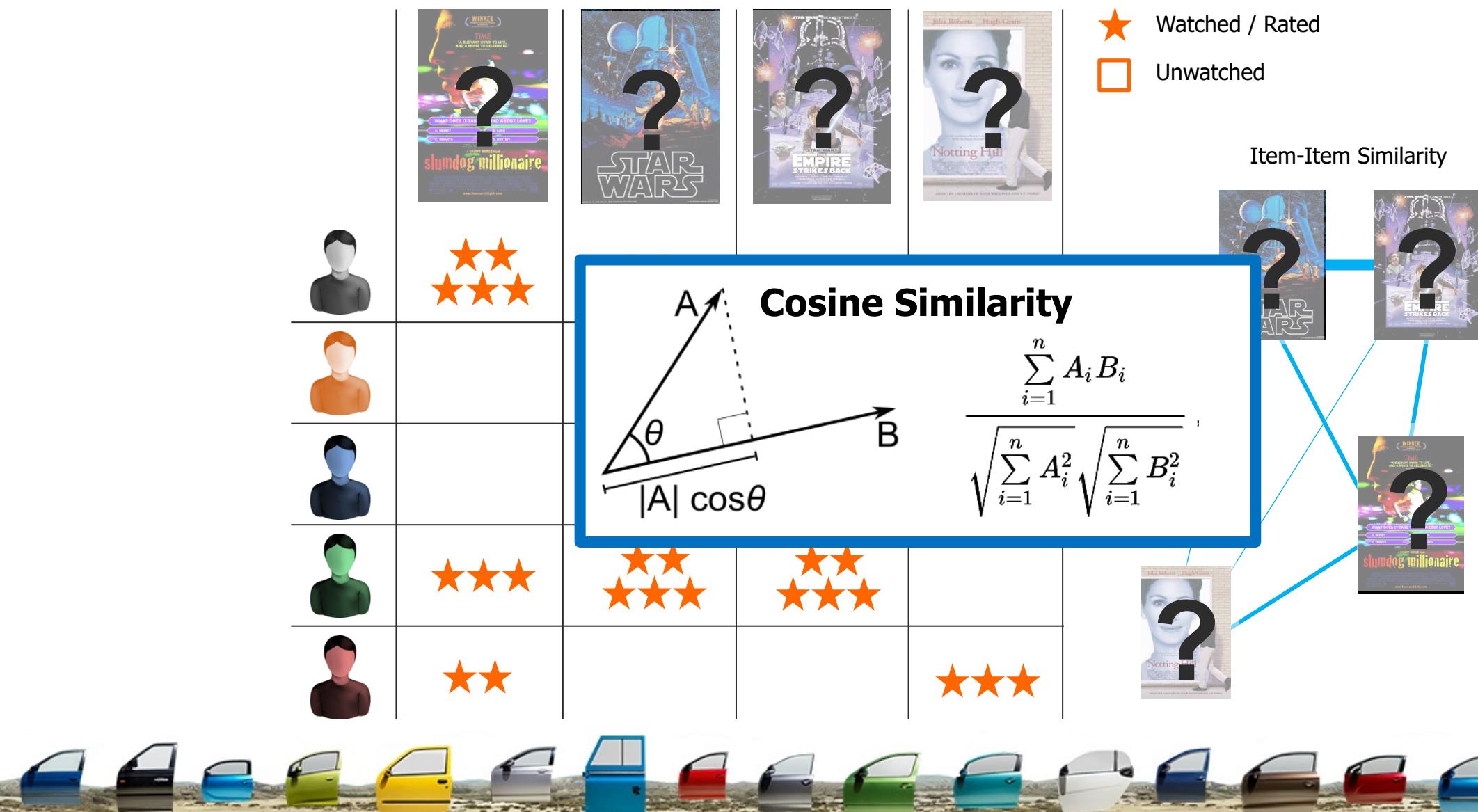
- Introduction
- Use-cases
- **Theory**
 - Collaborative Filtering
 - Content Based
- Our Approach
- Example
- Outlook



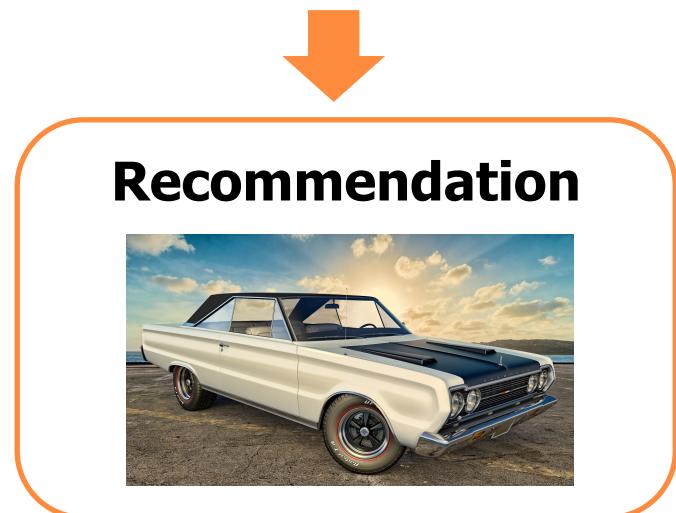
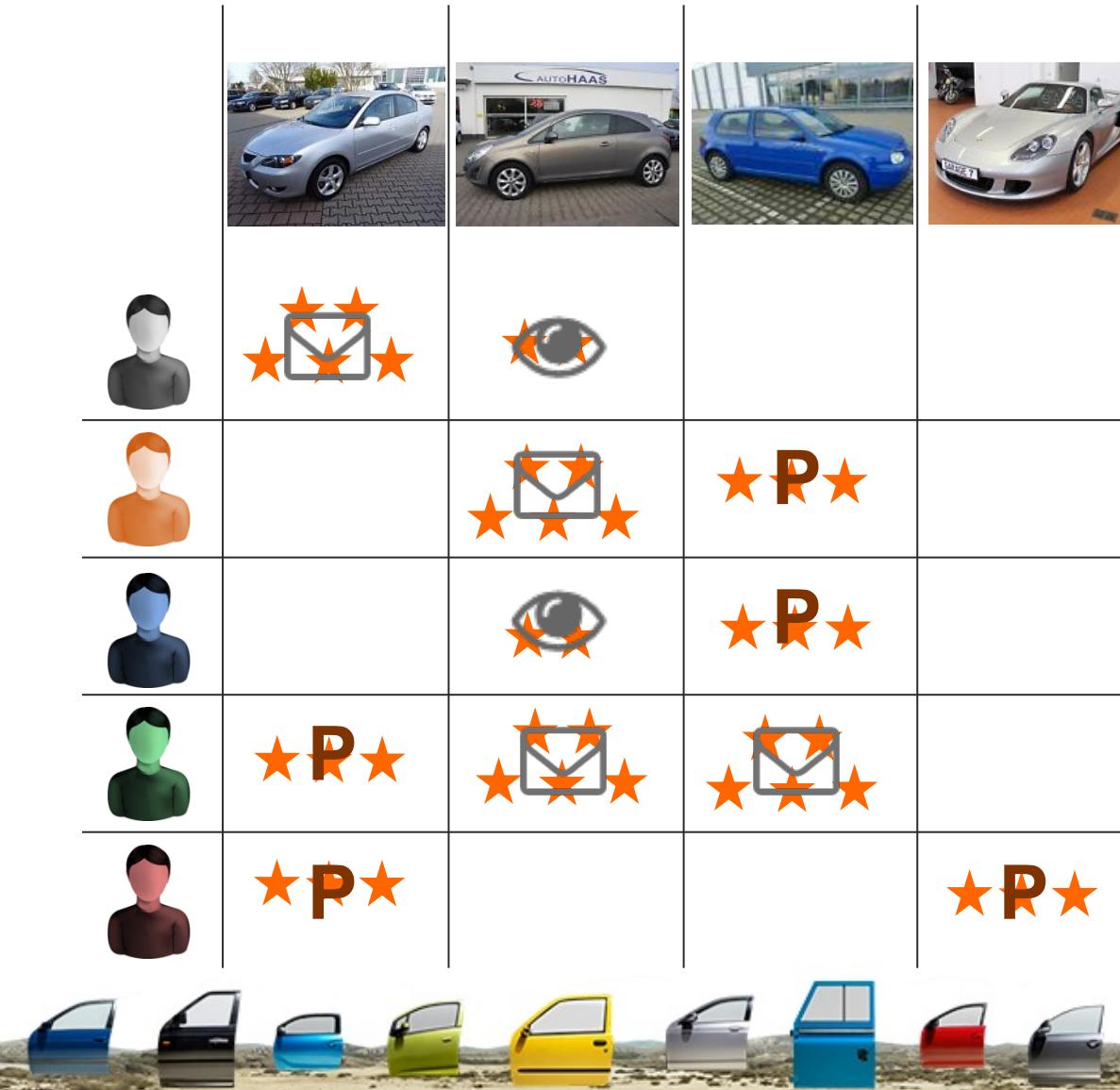
Collaborative Filtering



Item-based Recommendations



Item-based Recommendations



Summary of Collaborative Filtering

- ✓ Collective behaviour of users
- ✓ Standard-Method (it works, it's reliable etc.)
- ✗ **Cold Start Problem:** New listings need a certain number of clicks to be recommended.
- ✗ **Sparsity problems:** lot fewer interaction data points than total items and users.
- ✗ Content agnostic
- ✗ Only “batch-based” learning



Content-based Filtering: User Preferences



mobile.de User Preferences

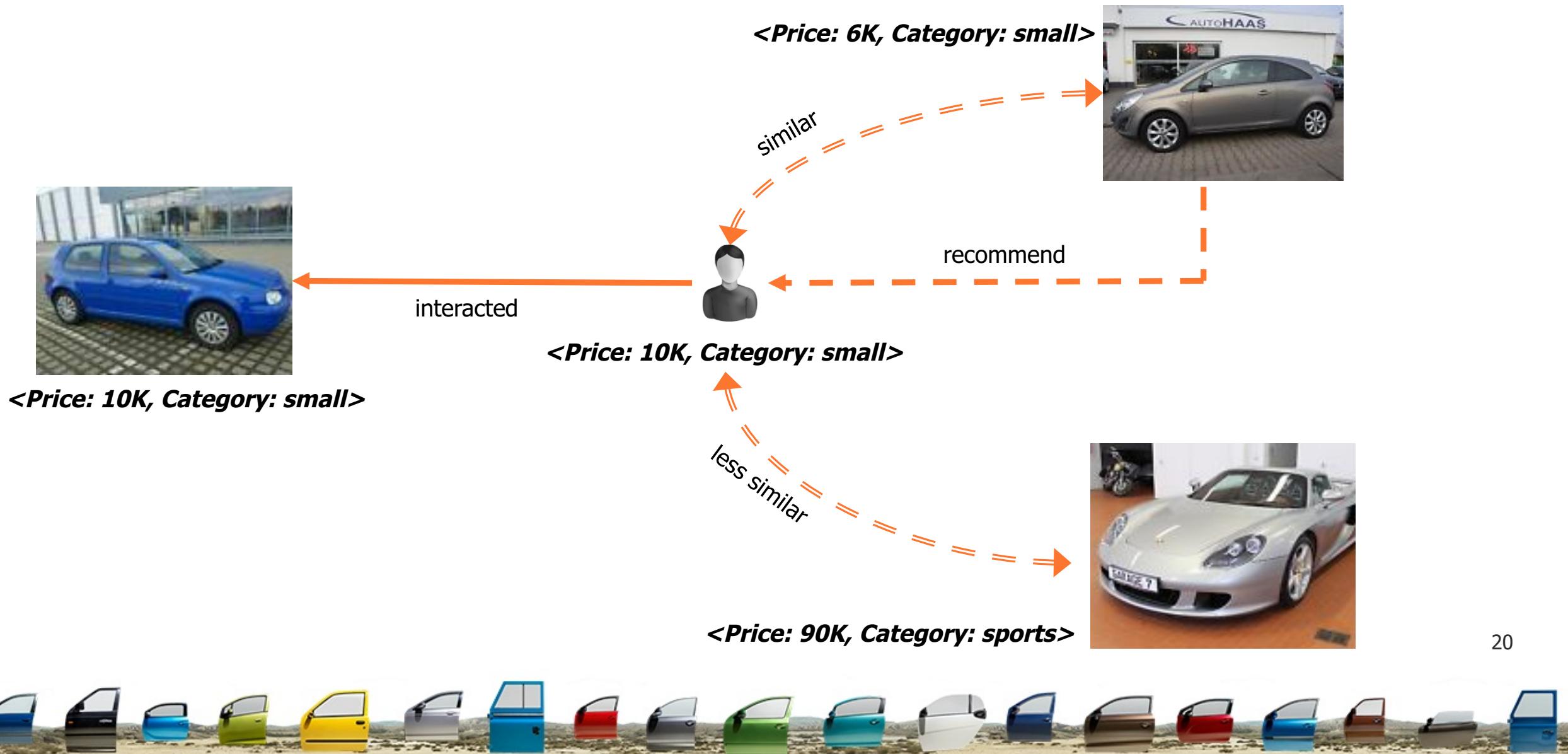


Anonymous

Looking For:	Used Car (100%)
Prefers (Make):	BMW (50%), Audi (50%)
Prefers (Model):	Audi A3 (25%), Audi A4 (25%), BMW 318 (50%)
Searching In:	lat 52.5206, lon 13.409
Search Radius:	300km
Preferred Price:	20 000€ ± 1500€
Preferred Mileage:	10 000km ± 5000km



Content-based Filtering

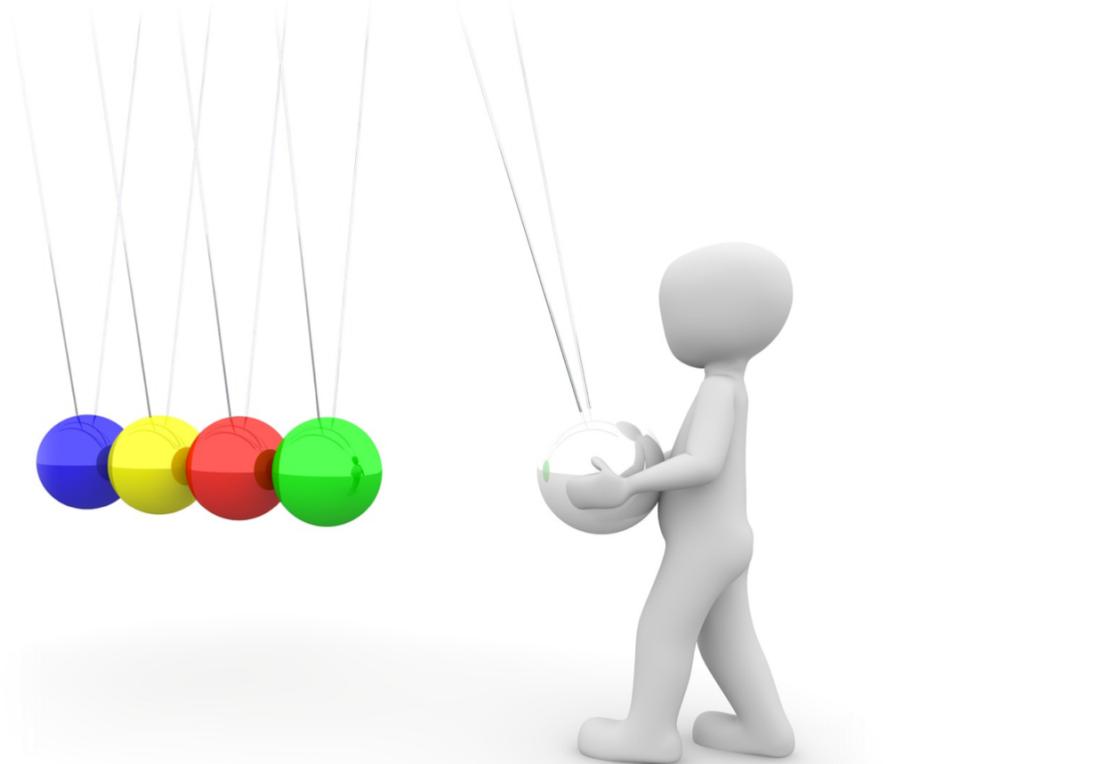


Summary of Content-based

- ✓ Works even if there are no other users
- ✓ **content-based** preferences of users based on a weighted vector of item features
- ✗ Hard to do recommendations for new users (cold start problem)
- ✗ Non-applicable for heterogeneous content types
- ✗ Low diversity, i.e. more of the same



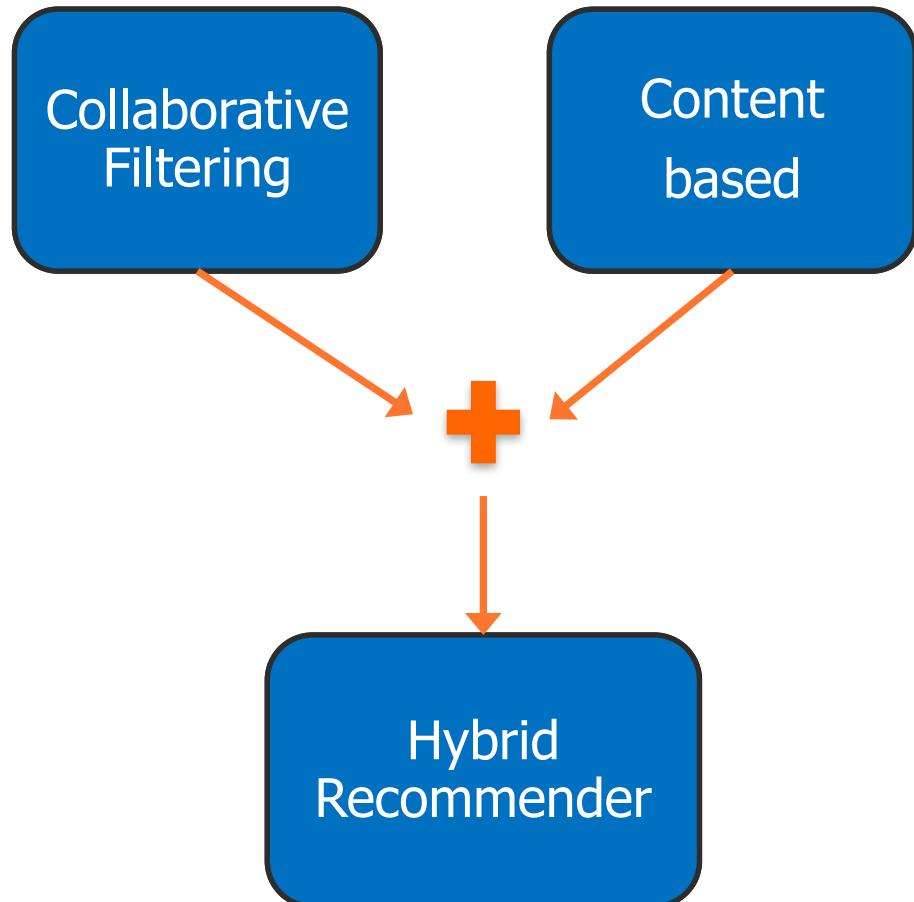
- Introduction
- Use-cases
- Theory
- **Our Approach**
- Example
- Outlook



Find the car that perfectly fits your life



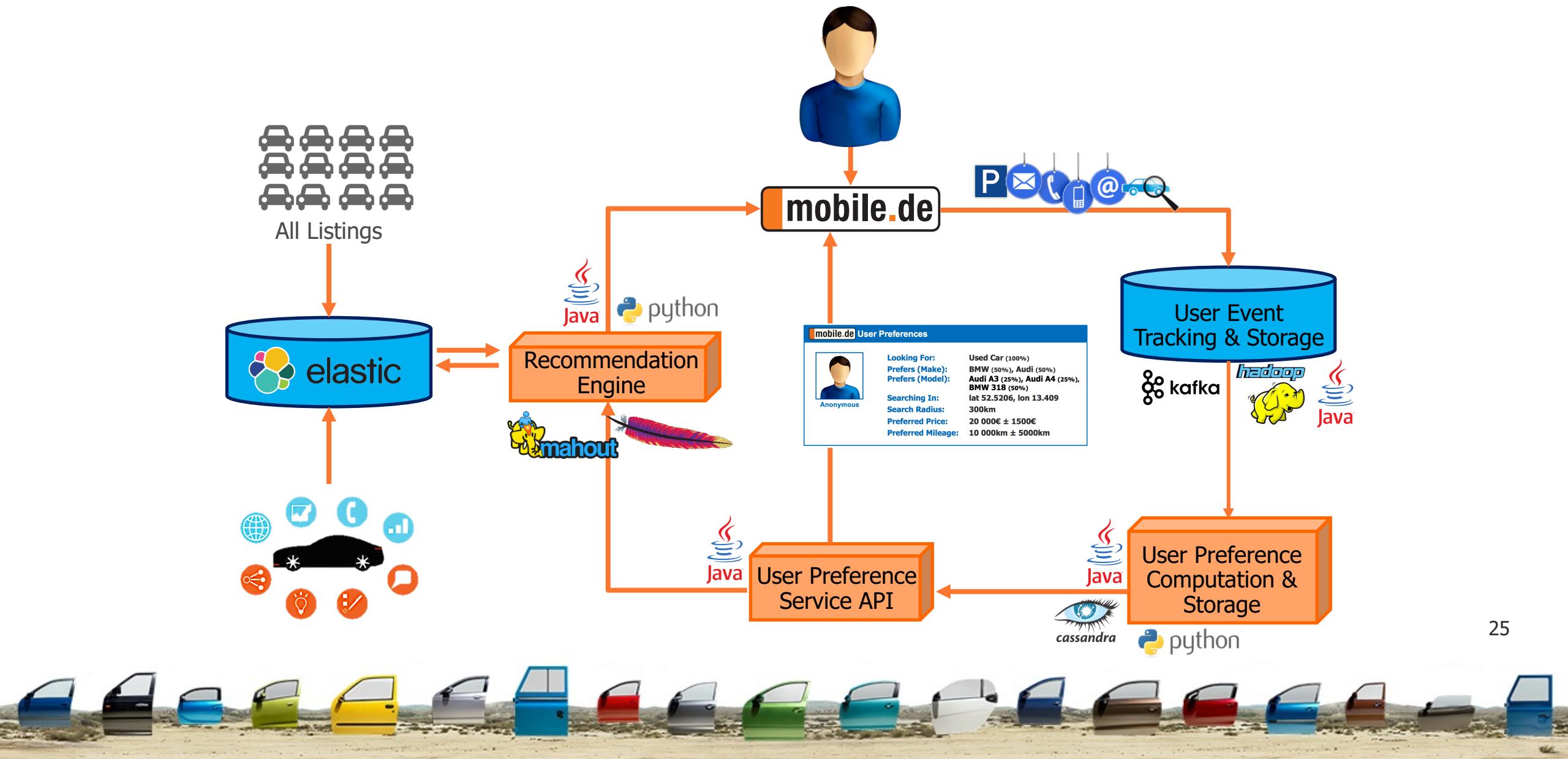
Hybrid Recommender



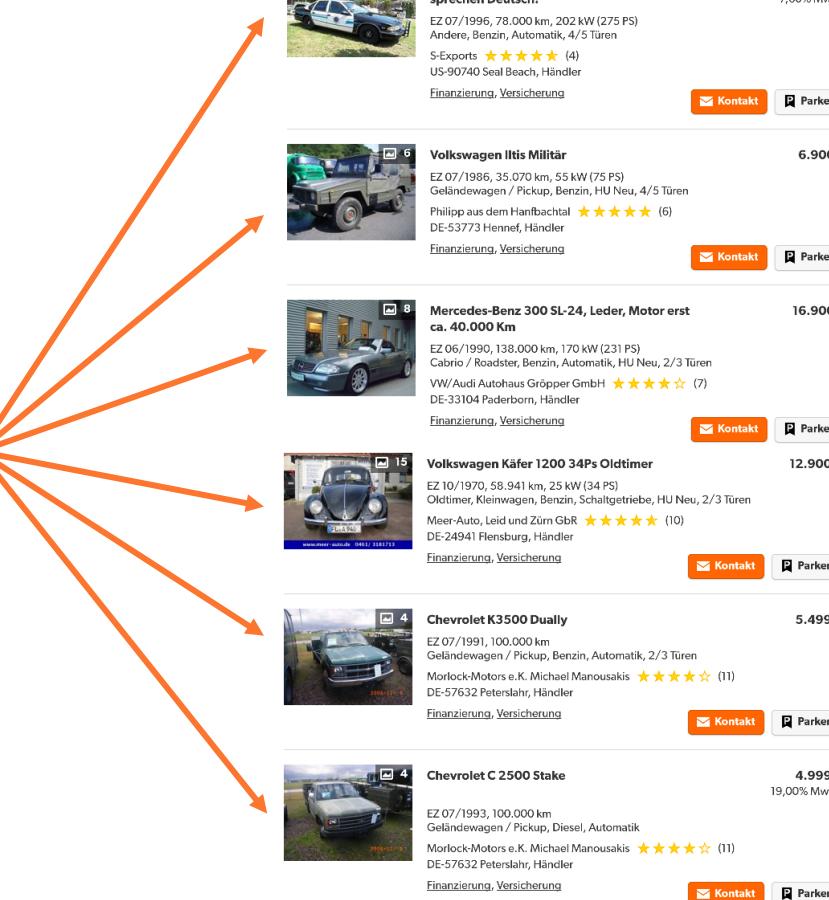
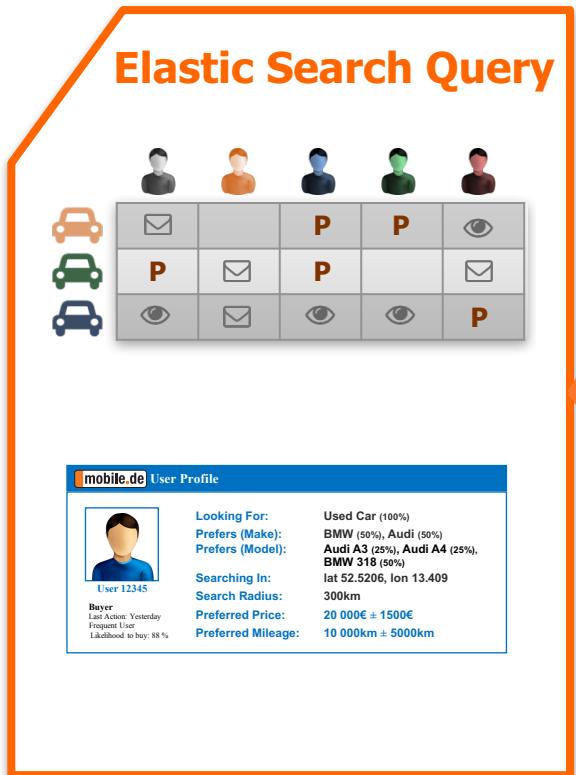
- no cold-start problem of new items
- integrate new user events in real-time
- robust and reliable concepts
- easy to tune for different use-cases
- comprehensible and debuggable



User Preference+ Recommendation Architecture



Hybrid Recommender Concept



Score

0.8

0.3

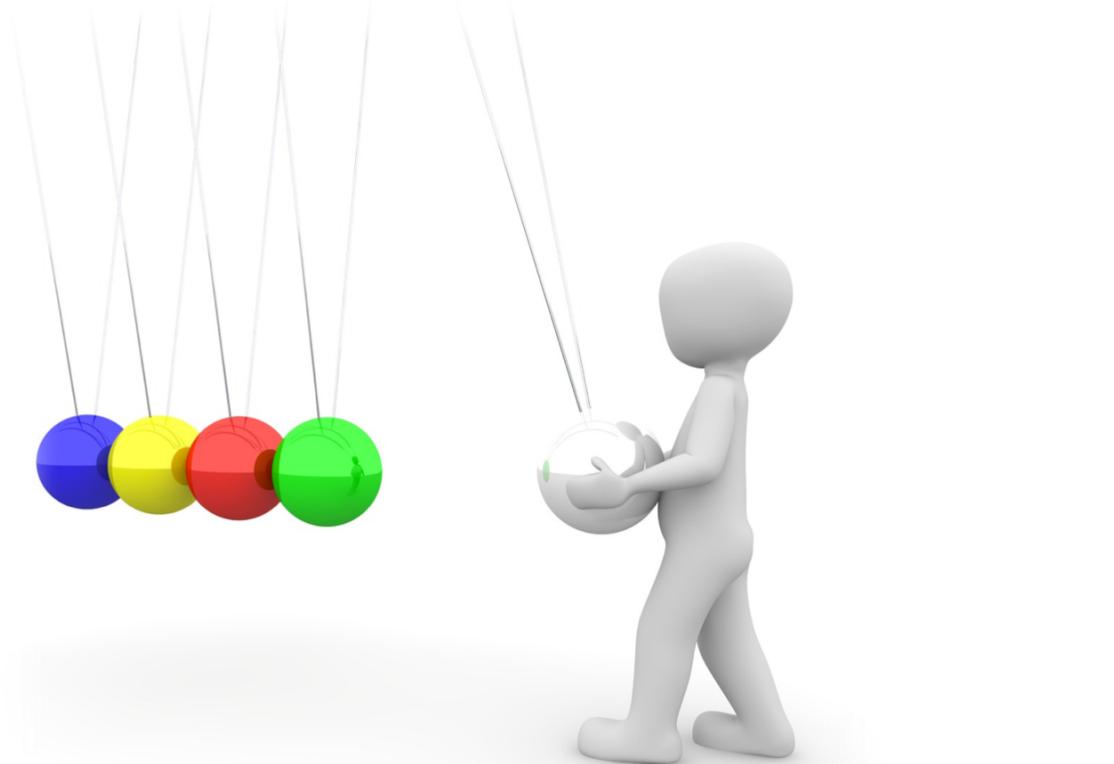
1.7

3.2

1.1

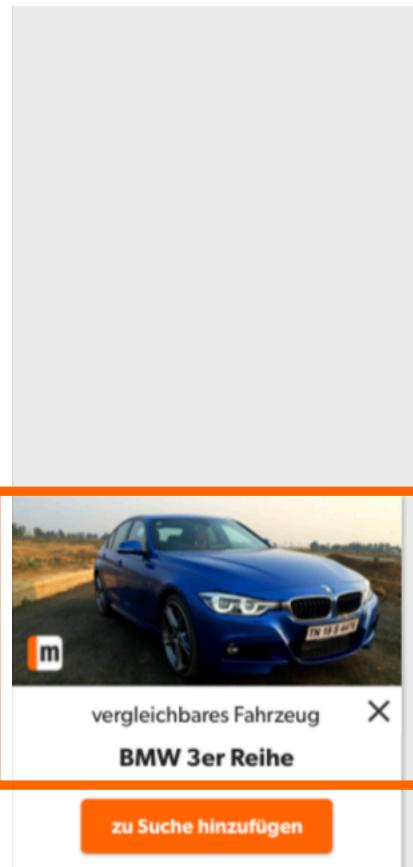
0.9

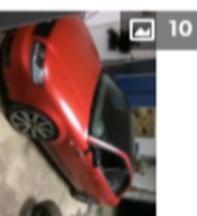
- Introduction
- Use-cases
- Theory
- Our Approach
- **Example**
- Outlook



Finding similar make/models

- **Users** are often *uncertain* with their choices in orientation phase
- Help users *explore similar* models to make informed decisions
- Exploit the **collaborative** aspect in defining the concept of similarity



 EZ 05/2003, 279.000 km, 120 kW (163 PS) Kombi, Diesel, Schaltgetriebe, 4/5 Türen ≈ 6,9 l/100km (komb.), ≈ 186 g CO ₂ /km (komb.) DE-51145 Köln, Privatanbieter Finanzierung , Versicherung 	Audi A4 Avant 2.5 TDI 2.500 €
 EZ 06/2001, 236.800 km, 110 kW (150 PS) Kombi, Unfallfrei , Benzin, Schaltgetriebe, HU 09/2018, 4/5 Türen ≈ 8,2 l/100km (komb.), ≈ 197 g CO ₂ /km (komb.) DE-49716 Meppen, Privatanbieter Finanzierung , Versicherung 	Audi A4 Avant 1.8 T 3.700 €
 EZ 11/2001, 234.000 km, 147 kW (200 PS) Limousine, Benzin, Schaltgetriebe, HU 11/2017 DE-94481 Grafenau, Privatanbieter Finanzierung , Versicherung 	Audi audi a4 b6 8E 1.8T 200ps 4.500 €



Make/Model Recommender with LightFM



Audi A4



similar vehicles

BMW 645

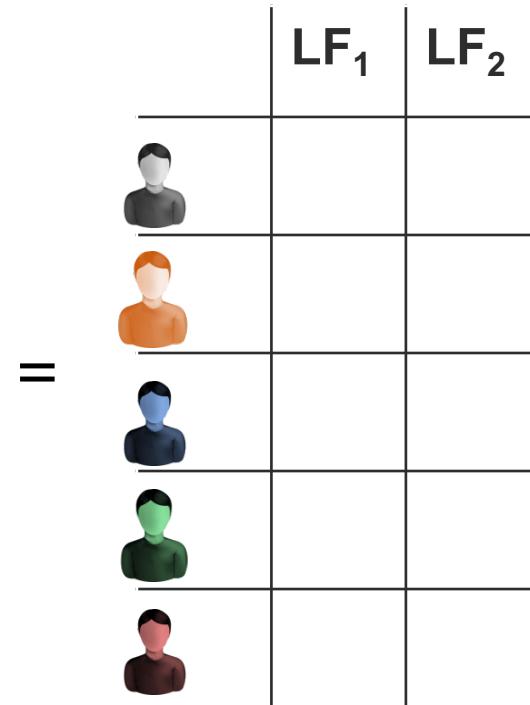


LightFM:

- Matured and well documented Python package
- Optimized and parallelized with Cython
- Hybrid recommender based on matrix factorisation
- Supports Learning-to-Rank objectives (BPR, WARP)



Non-negative Matrix Factorisation (NMF)



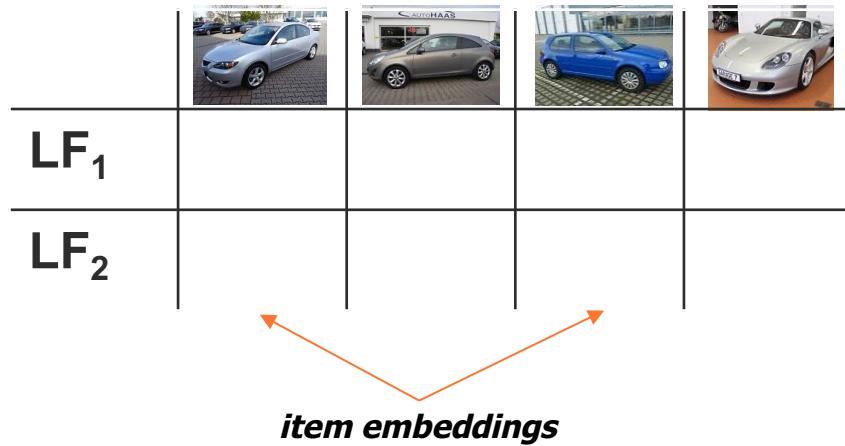
X

LF ₁			
LF ₂			

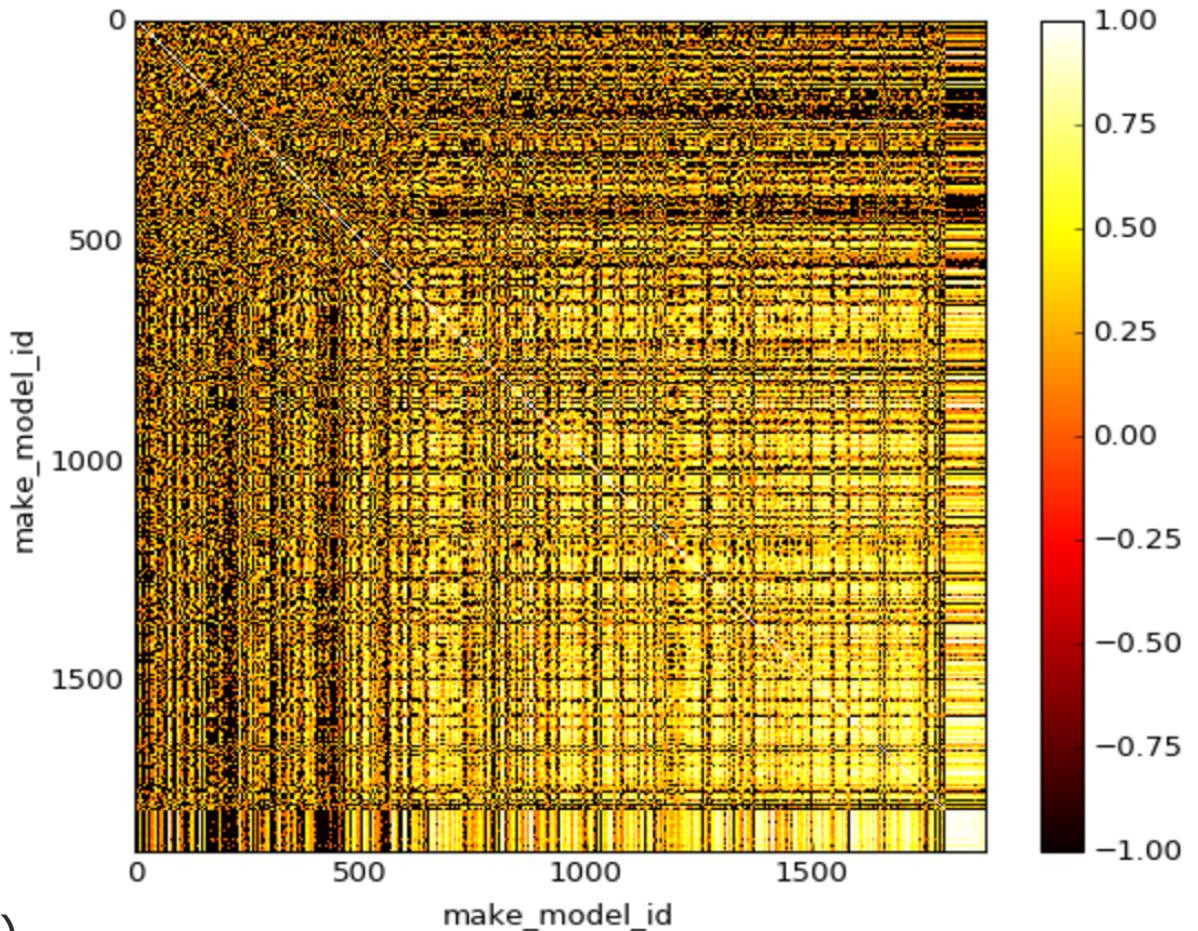
$$M(|U| \times |I|) = L(|U| \times |LF|) \times R(|LF| \times |I|)$$



NMF: Embeddings to Similarities



- car is represented as an **item embedding**
- Given 2 embeddings, \mathbf{LF}_{item_i} and \mathbf{Lf}_{item_j}
- Compute $sim(\mathbf{LF}_{item_i}, \mathbf{LF}_{item_j})$
- Find pairwise values for all item pairs ($|I| \times |I|$)



NMF: Persisting Item similarities

```

def model_fitting(matrix, loss, no_comp, epochs,
                  num_threads, item_alpha):
    .....
    Args:
        matrix: sparse matrix
        loss: loss function to be employed
        no_comp: dimension of the embeddings
        epochs: #passes over the data
        num_threads: parallel threads to run
        item_alpha: L2 regularization

    Returns:
        fitted Matrix factorization (MF) model
    .....
model = LightFM(loss=loss,
                 item_alpha=item_alpha,
                 no_components=no_comp)

return model.fit(matrix, epochs=epochs,
                 num_threads=num_threads)

```

```

{
    "_id" : ObjectId("588f1a8ca2418eeb98367fb8"),
    "itemName" : "bmw-645",
    "makeId" : 3500,
    "makeModelId" : "3500-31",
    "topKResults" : [
        {
            "itemName" : "audi-a4",
            "makeId" : 1900,
            "makeModelId" : "1900-9",
            "score" : 0.67208,
            "modelId" : 9
        },
        {
            "itemName" : "mercedes-benz-cla 180",
            "makeId" : 17200,
            "makeModelId" : "17200-225",
            "score" : 0.65011,
            "modelId" : 225
        }
    ]
}

```

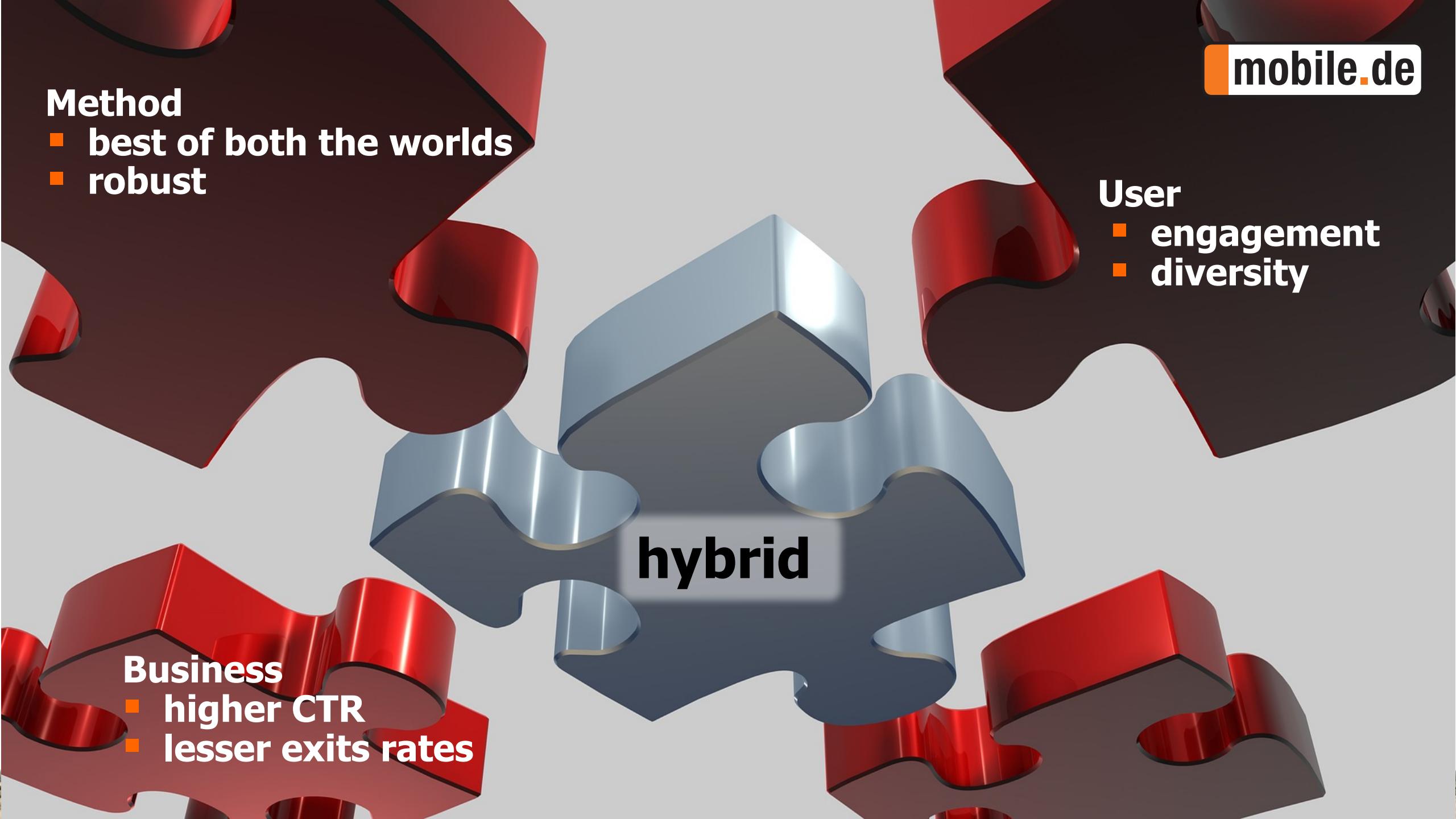


Method

- best of both the worlds
- robust

User

- engagement
- diversity

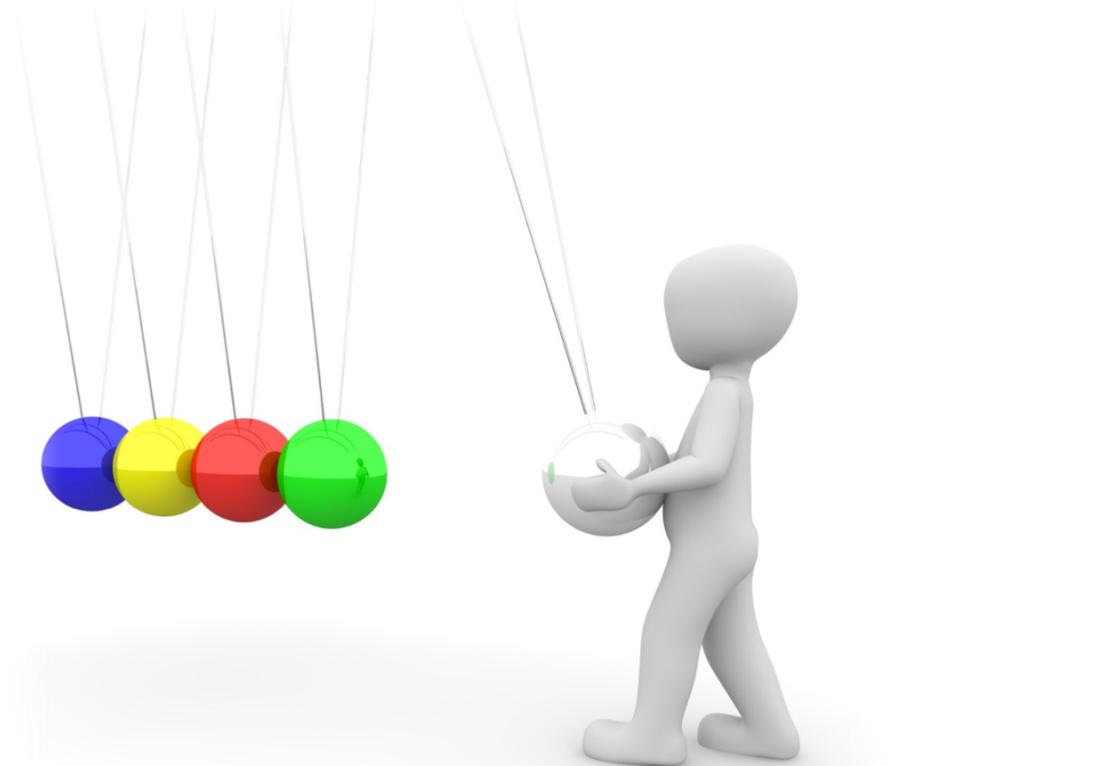


hybrid

Business

- higher CTR
- lesser exits rates

- Introduction
- Motivation & Use-cases
- Theory
- Implementation
- **Outlook**



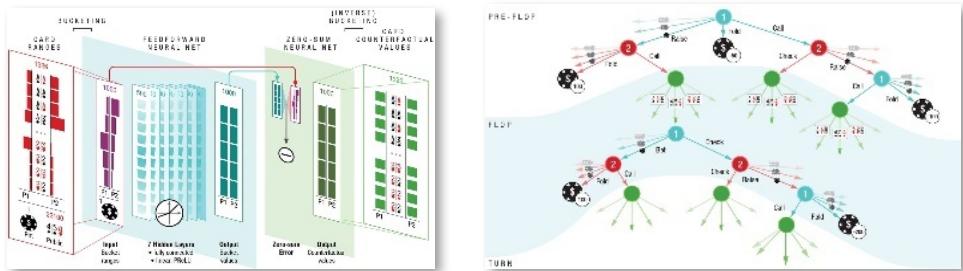
Deep Learning

Recent Breakthroughs in Deep Learning



Dermatologist-level classification of skin cancer with deep neural networks
(Nature 542, 115-118, February 2017)

DeepStack: Expert-level artificial intelligence in heads-up no-limit poker
(Science, March 2017)

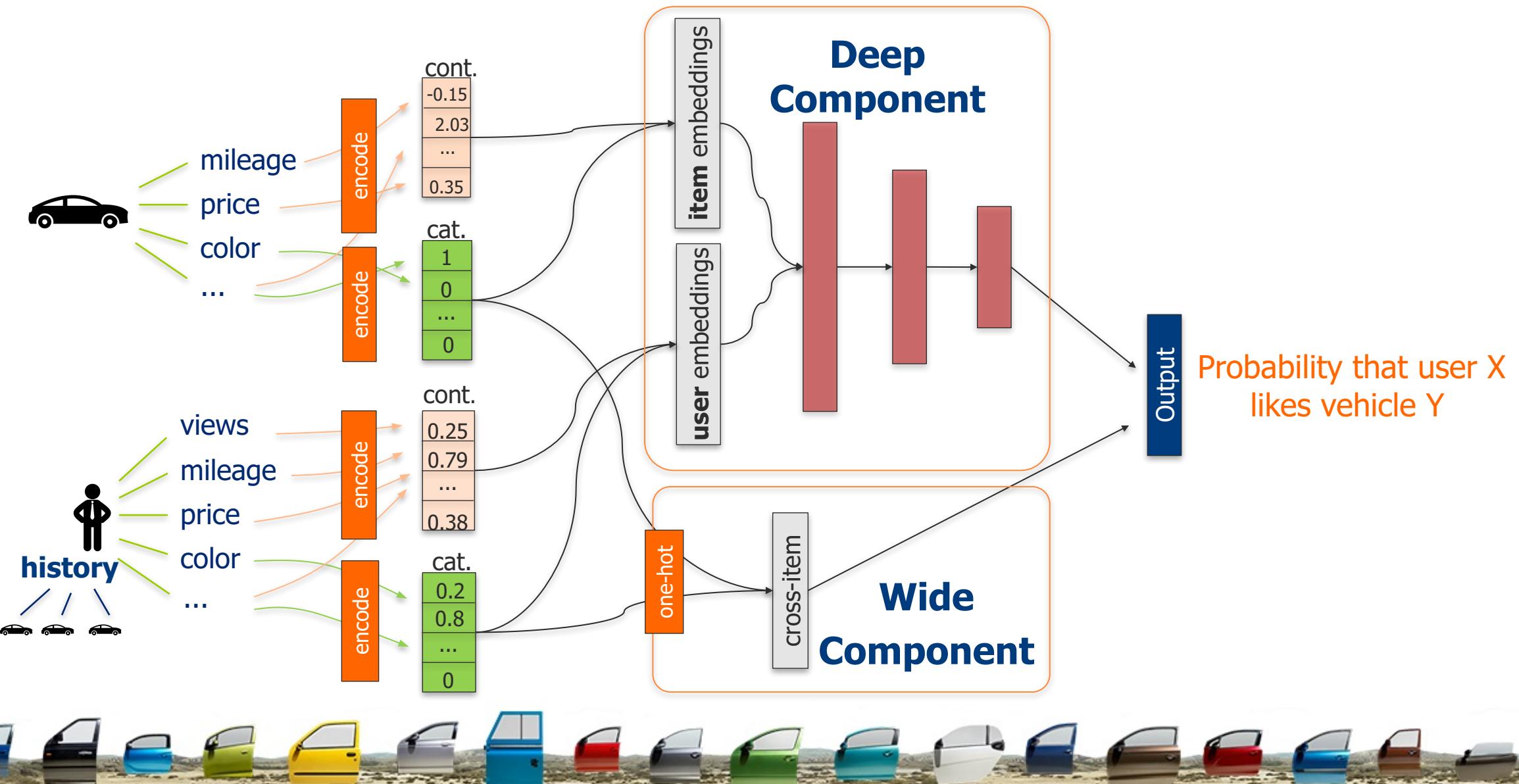


Reasons for Deep Learning

- captures nonlinear relations
- holistic approach, i.e. reduces number of components possibly
- less feature engineering
- possibly improved quality



Approach: Wide and Deep Model





Recommendations support users to find the perfect vehicle based on their preferences and by collaboration

Any questions?

