



Cooky

AUTOMATIC RECIPE SUGGESTION



AGENDA

IDEE

ZIEL & FOKUS

AUFBAU

DATENSATZ &
PREPROCESSING

RECOMMENDER

LIVE DEMO

IDEE

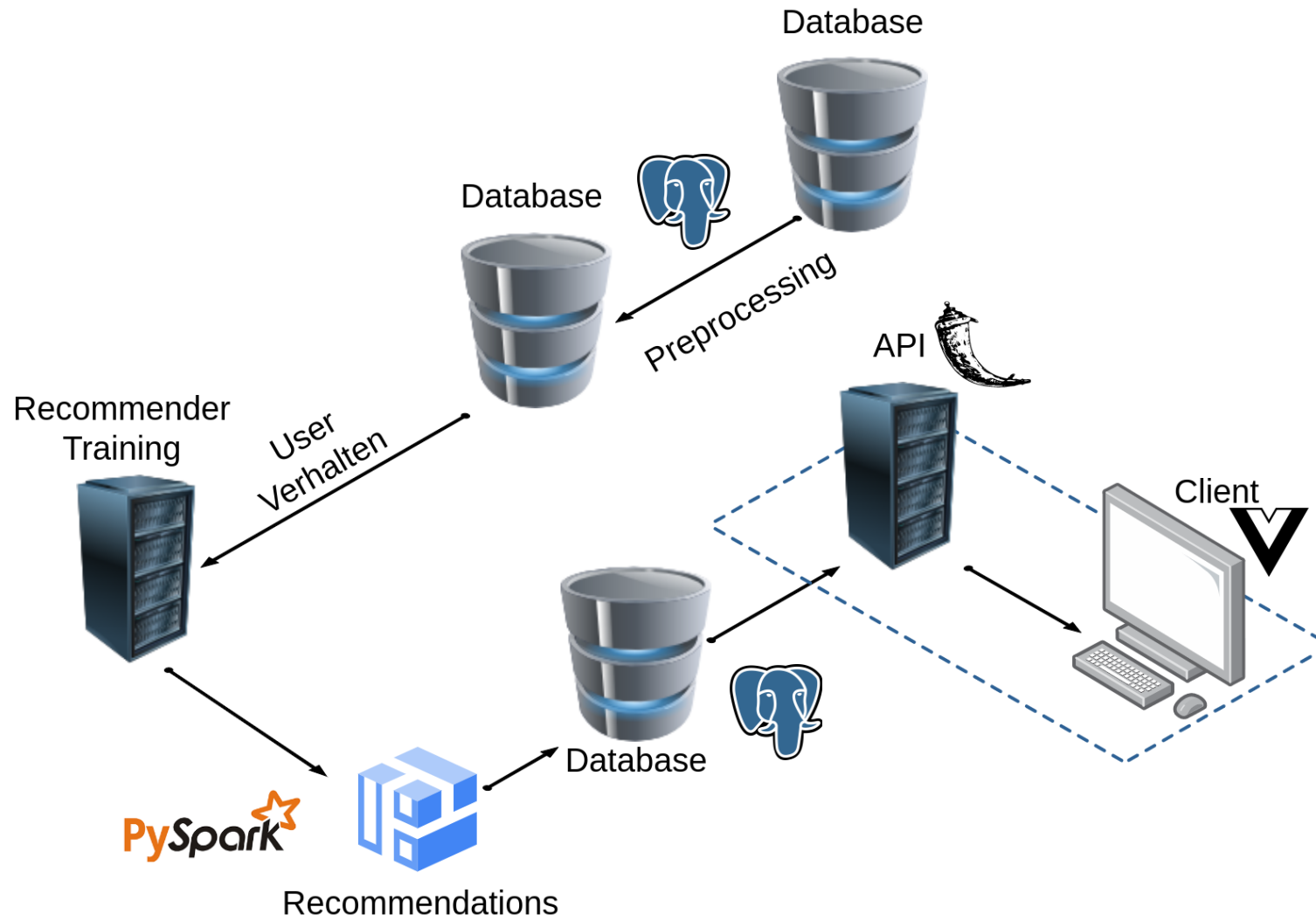
- Kochrezepte für Nutzer online zugänglich machen
- Bewertung dieser ermöglichen
- Rezepte aufgrund des Benutzerverhaltens empfehlen
- Virtuelle Vorratslager für Benutzer zur Optimierung der Empfehlung

FOKUSPUNKTE

- Data Preprocessing and Cleaning (von großen unstrukturierten Daten)
- Deployment (einer Web-App als UI)

ZIEL

- Responsive Webapp für Benutzer
- Vollständiges Recommendersystem und Datenbank



AUFBAU

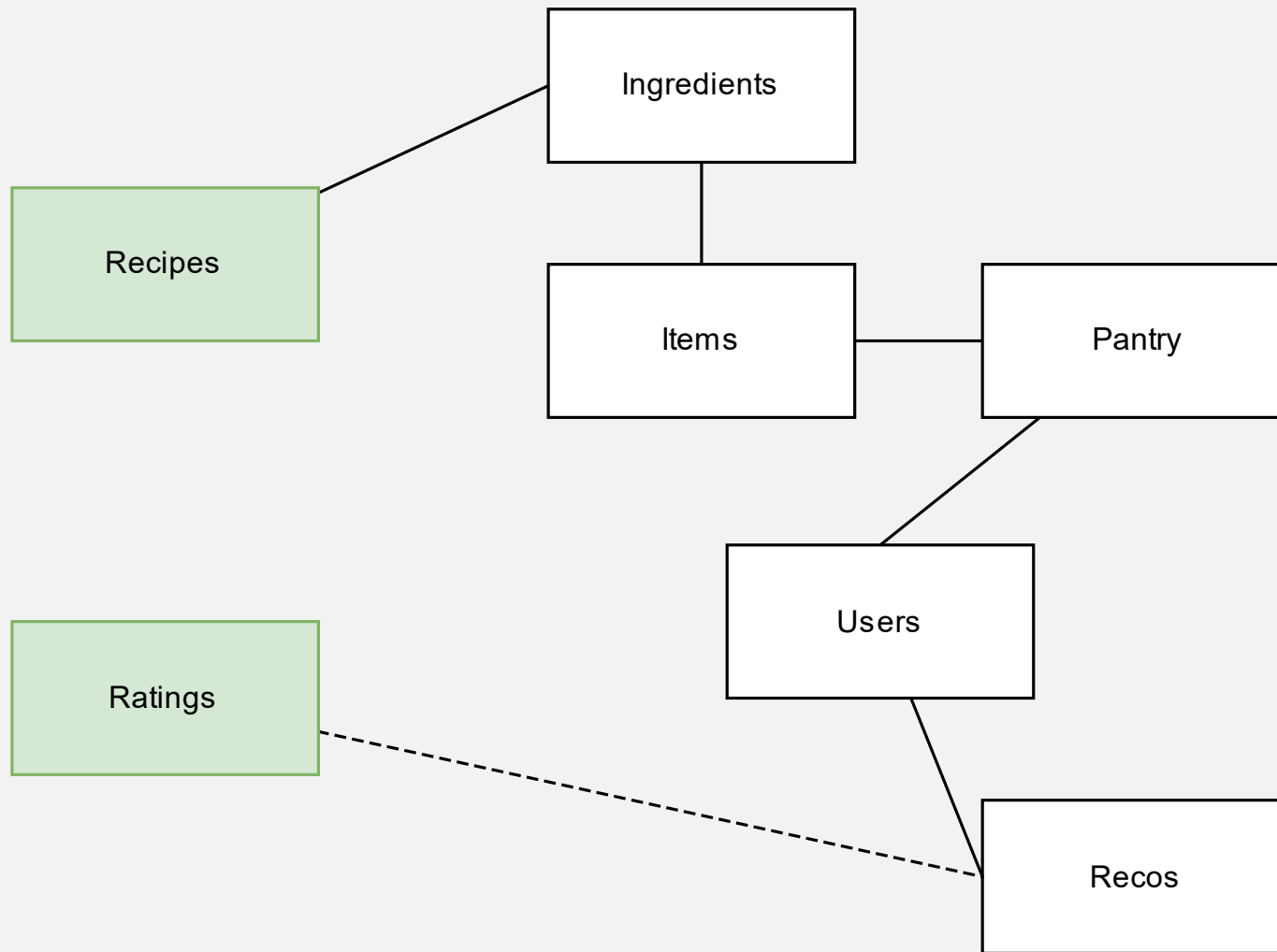
- Rezepte Datensatz
 - 2,2 Mio Rezepte

Title	Ingredients	Directions	NER
Nolan'S Pepper Steak	"["1 1/2 lb. round steak (1-inch thick)," "1 can drained tomatoes, ...]"	"["Roll steak strips in flour."", ...]"	[„round steak“, „tomatoes“, ...]

- Recommender Trainings-Datensatz
- 1,1 Mio Bewertungen

User-ID	ISBN	Book-Rating
276725	0155061224	5

DATENSATZ



DATENBANK

PREPROCESSING

VORGEHEN

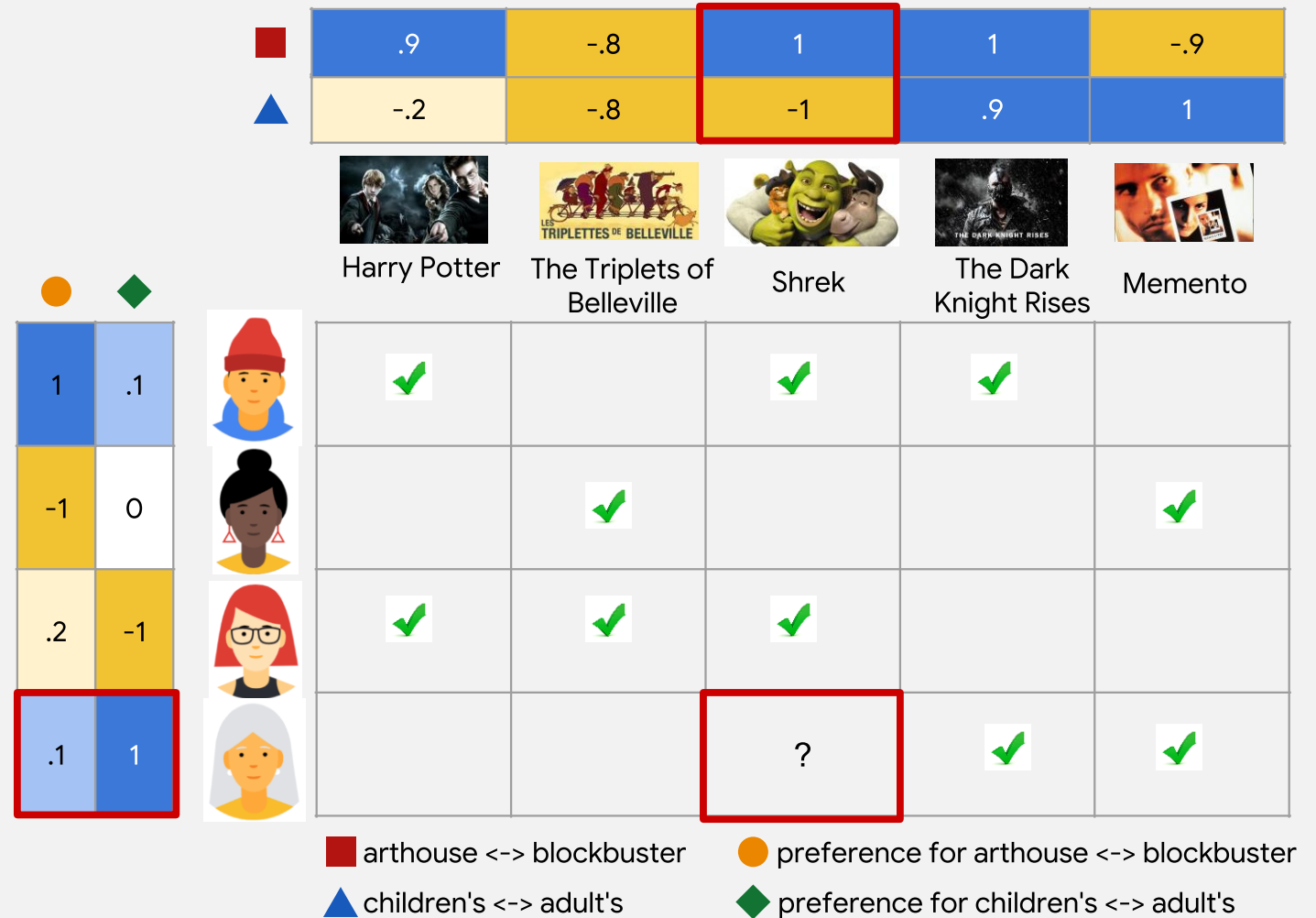
1. **Rezepte speichern**
2. **Extrahiere Zutaten (Variante A)**
 - mit FoodBERT (NLP Modell von HuggingFace)
 - Ermittle Einheit und Menge mit Regular Expression→ Nicht Zuverlässig
3. **Extrahiere Zutaten (Variante B)**
 - Mit NER (gegeben durch Datensatz)
 - Keine Mengeneinheiten
4. **Erzeuge einzigartige Items**
5. **Generiere synthetische Nutzerdaten**

BEISPIEL

- „Nolan'S Pepper Steak“ mit Recipe ID 3
- „I ½ lb. steak“ → „steak“
- „I ½ lb. steak“ → „lb.“ und „1,5“
- Steak → Item ID 15
- ISBN → Recipe ID (gegen Cold-Start Problem)

RECOMMENDER: COLLABORATIVE FILTERING

- **Intuition:** Ähnliche Nutzer mögen ähnliche Gerichte
- Matrix Faktorisierung
 - Aus Benutzern und Produktauswahl Empfehlungen ziehen





Harry Potter



The Triplets of
Belleville



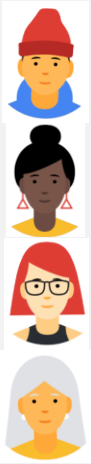
Shrek



The Dark
Knight Rises



Memento



✓		✓	✓	
	✓			✓
✓	✓	✓		
			✓	✓

≈

1	.1
-1	0
.2	-1
.1	1

.9	-1	1	1	-.9
-.2	-.8	-1	.9	1

.88	-1.08	0.9	1.09	-0.8
-0.9	1.0	-1.0	-1.0	0.9
0.38	0.6	1.2	-0.7	-1.18
-0.11	-0.9	-0.9	1.0	0.91

RECOMMENDER: ALS

- Optimizer ALS (Alternating Least Squares)
→ Löst Optimierung mit 2 unbekannten
- RMSE = 5.74
(Rating weicht also bis zu +- 5.74 ab)

```
from pyspark.ml.recommendation import ALS

(train, test) = pdf_ratings.randomSplit([0.7, 0.3], seed=123)
als = ALS(
    rank=10,
    maxIter=10,
    regParam=0.01,
    alpha=1,
    userCol="n_user_id",
    itemCol="n_recipe_id",
    ratingCol="n_rating",
    coldStartStrategy="drop",
    nonnegative=True,
    implicitPrefs=False,
    seed=123
)

model = als.fit(train)

pred = model.transform(test)
eval = RegressionEvaluator(metricName="rmse", labelCol="n_rating",
    predictionCol="prediction")
rmse = eval.evaluate(pred)

user_recos = model.recommendForAllUsers(10)
```

DEPLOYMENT

Interfaces:

API	Expect	Returned	Uses
/login	Session	Success, 200	check_user()
/pantry	Session	{all:{...},pantry:{...}}, 200	get_current_stock(), get_all_items()
/pantry/delete	Session, Item-ID, Quantity	Success, 202	reduce_stock()
/pantry/add	Session, Item-ID, Quantity	Success, 202	add_item2stock()
/pantry/cook	Session, Recipe-ID	Success, 202	cook_meal()
/explore	Session	Success, 200	Meal_reco_without_pantry(), meal_reco_by_pantry(), get_recipes()
/explore/rating	Session, Recipe-ID	{rating: float}, 200	get_avg_rating()

Live Demo

Appendix

Business Model Canvas

Designed for:

Cooky










Designed by:

Date:

27-07-22

Version:

1.0

Key Partners  <ul style="list-style-type: none"> - Customers - Cloud Anbieter - Werbepartner MOTIVATIONS FOR PARTNERSHIPS: Nutzen- und Gewinnmaximierung	Key Activities  <p>Bereitstellen von individuellen / passenden Kochrezepten. Überwachung & Management des Vorratspeichers</p> <p>CATEGORIES:</p> <ul style="list-style-type: none"> - Empfehlungen - Management 	Value Propositions  <p>Einfach passendes Kochrezept basierend auf individuellen Verhaltensmuster und in der Vorratskammer vorhandenen Zutaten, zu erhalten</p> <p>CHARACTERISTICS: Sehr benutzerfreundlich und intuitiv</p>	Customer Relationships  <p>Persönliche Empfehlungen und Hinweise über die Webapp</p>	Customer Segments  <ul style="list-style-type: none"> - Personen die nicht gerne Planen - Für spontane Köche - Für Anfänger/ Interessierte (App als Unterstützung)
	Key Resources  <p>Datensatz Algorithmen Webapp</p> <p>TYPES OF RESOURCES: data + knowledge</p>		Channels  <p>Webapp, Mail, (Website)</p>	
Cost Structure  <p>Hohe Kosten für Entwicklung -> Weiterentwicklung der App</p> <p>Geringe Kosten für Betrieb und Instandhaltung</p> <p>Value Driven (focused on value creation, premium value proposition).</p>		Revenue Streams  <ul style="list-style-type: none"> - Free Basic Version with simple features - Werbung (Sponsored Recipes) - Premium Version (Licence Business Model monthly/ yearly) mit Ansprechpartner und mehr Funktionen 		