



## Session 6: **Ausblick**





# Agenda

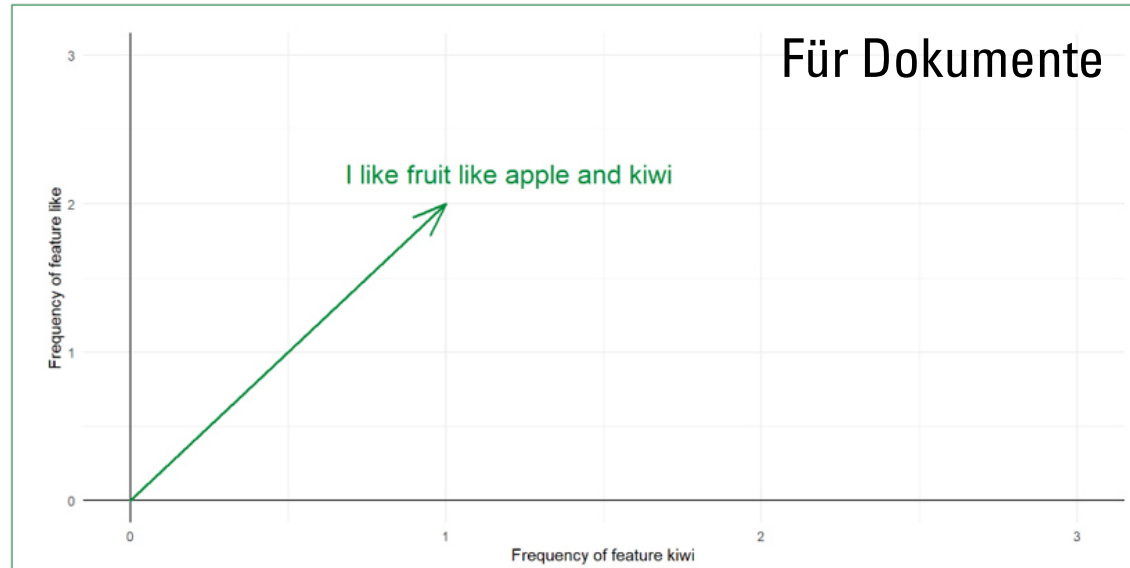
---

1. Weitere Verfahren
2. Ausblick

# 1. Weitere Verfahren

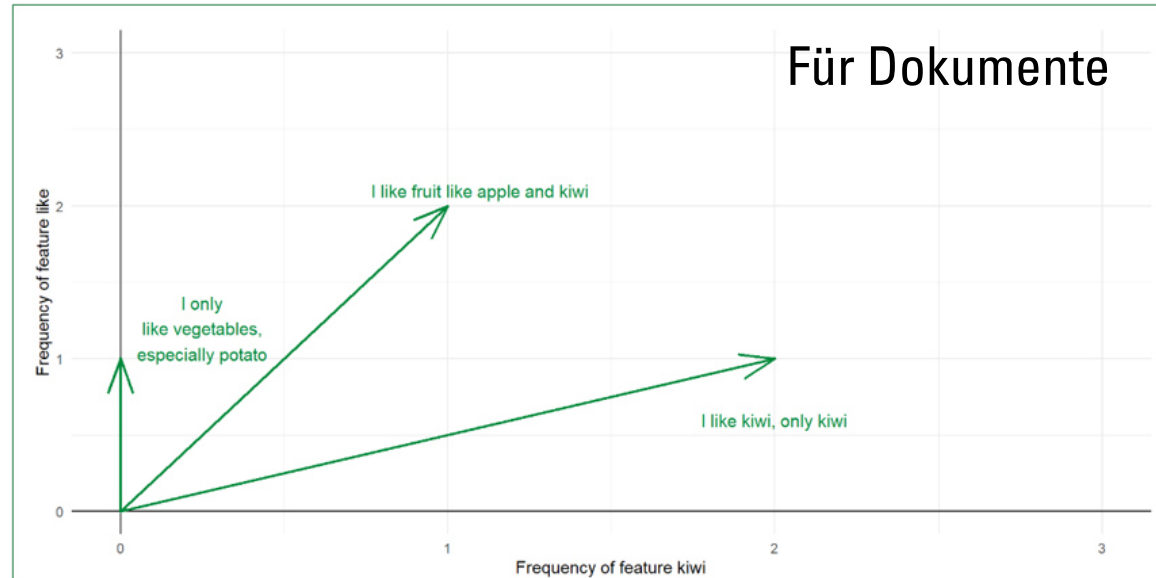
# Trend I: Mehr als Bag-of-Words

- **Vector Space Modelle (VSM):** Repräsentieren Wörter/Dokumente als Vektoren in einem Vektorraum. Vektoren, die einander nahe sind, sind sich semantisch nahe.



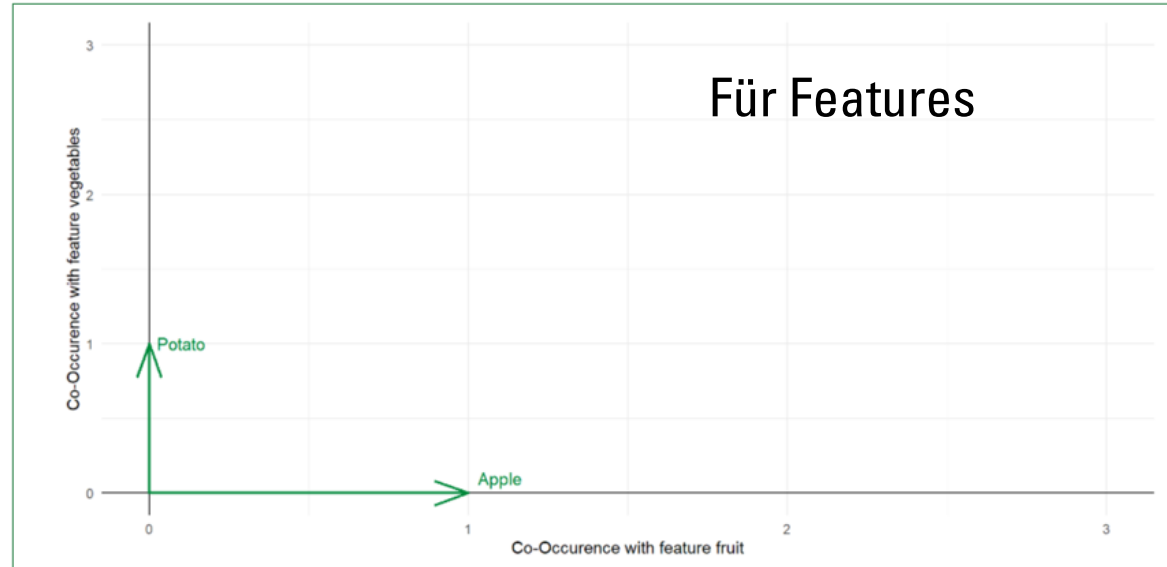
# Trend I: Mehr als Bag-of-Words

- **Vector Space Modelle (VSM):** Repräsentieren Wörter/Dokumente als Vektoren in einem Vektorraum. Vektoren, die einander nahe sind, sind sich semantisch nahe.



# Trend I: Mehr als Bag-of-Words

- **Word Embeddings:** Repräsentieren z. B. Features als Vektoren in einem  $n$ -dimensionalen Raum. Wir erklären die Bedeutung eines Features über die Features in dessen Kontext!





# Trend I: Mehr als Bag-of-Words

---

- Word Embeddings...
  - “predict a focal word as a function of the other words that appear within a small window of that focal word” ([Rodriguez et al., 2023, p. 3](#))
  - “represent a word as a dense vector in a low-dimensional space learned from unlabeled data” ([Grimmer et al., 2022, p. 79](#))
- Wir erklären die Bedeutung eines Wortes durch Wörter um dieses herum!
- Beispiel [hier](#)

# Trend II: Mehr als Text

- Andere Formen von Inhalt analysieren, z.B.
  - Bilder
  - Videos
  - Audio



Quelle: Tagesschau, 2023



# Trend III: Mehr als Englisch

- Zunehmend Verfahren für multilinguale Analysen (Licht & Lind, [2023](#))





## Trend IV: Mehr als Automatisierung

---

- Semi-automatisierte Verfahren,
  - z.B. Active Learning
  - z.B. Seeded Sentimentanalyse (Watanabe, [2021](#))
- Automatisierte Verfahren mit geringen Mengen an Trainingsdaten, z.B. Few-Shot Learning (Rieger et al., [2024](#))
- Gesundes Misstrauen in menschliche Codierungen (Baden et al., [2024](#); Oschatz et al., [2024](#))

## 2. Evaluation



# Evaluation: The good & the bad

---

- In 1 Satz, was hat euch am Workshop **gefallen/geholfen**/wovon hätte es **mehr Inhalt** sein dürfen
- In 1 Satz, was hat euch am Workshop **nicht gefallen/nicht geholfen**/wovon hätte es **weniger** Inhalt sein dürfen

# Danke! Fragen?



**Dr. Valerie Hase**  
IfKW, LMU Munich



valeriehase



valerie-hase.com



**Luisa Kutlar**  
IfKW, LMU Munich



luisakutlar