

Entwicklung eines Chatbots

Präsentation

Von

Alexander Beutler, Benedikt Göller und Tim
Wildberger

Agenda

- Was war unser Ziel?
- Unser Chatbot
- Unser Entwicklungsprozess
 - Datensatz
 - Datentransformation
 - Clustering
 - Klassifikator
 - Konzept (Demo)
- Lessons Learned
- Fragen?

Was war unser Ziel?

- Einen Chatbot für Smalltalk



Unser Chatbot



Datensatz von [kaggle.com](https://www.kaggle.com)



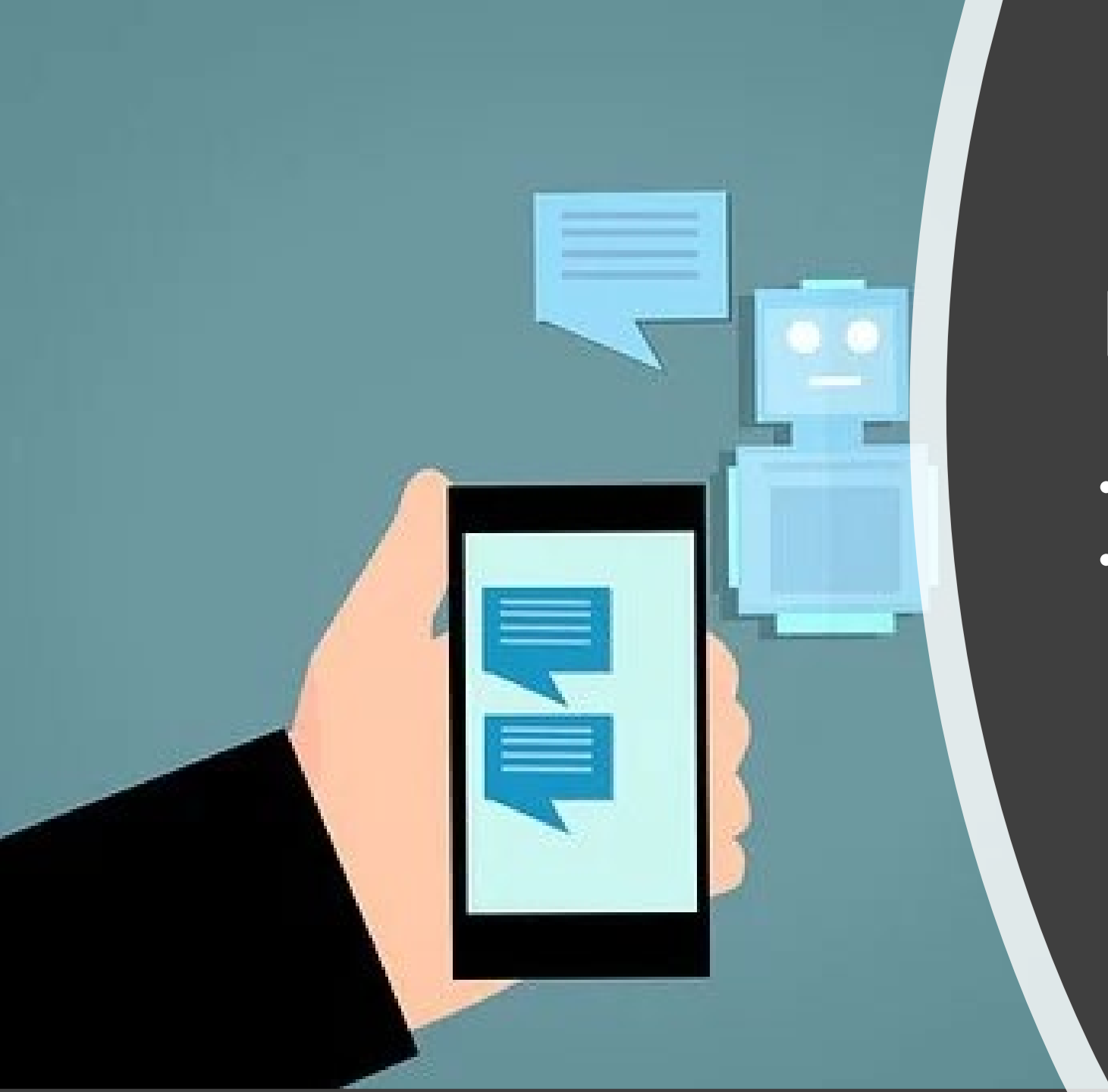
Framework zum lernen: **scikit learn**



Jupyter Notebook als IDE



Der Entwicklungsprozess



Datensatz

- Gespräch zwischen 2 Bots
- Deckt weite Themenbereiche ab

Datenstransformation

- Lemmatisierung
 - Alle Wörter auf Wordstamm runterbrechen
- Stop-Wörter entfernen?
 - Nein, da die Qualität des Datensatzes verringert wird um Alltagsgespräche zu führen



Clustering



Umwandlung des Textes in Numerische Daten

Word2Vec-Transformation



Ward Clustering

Cluster beinhalten Kontext
sind relativ gleichmäßig verteilt (keine
riesen Cluster)



Klassifikator

auf Basis des Modells
trainieren eines neuen
Modells zur
Klassifikation neuer
Eingaben

K-nearest Neighbour

Stellenweise
Ähnlichkeiten zur
Clustering-Methode

Höchste accuracy

Schnellste Laufzeit

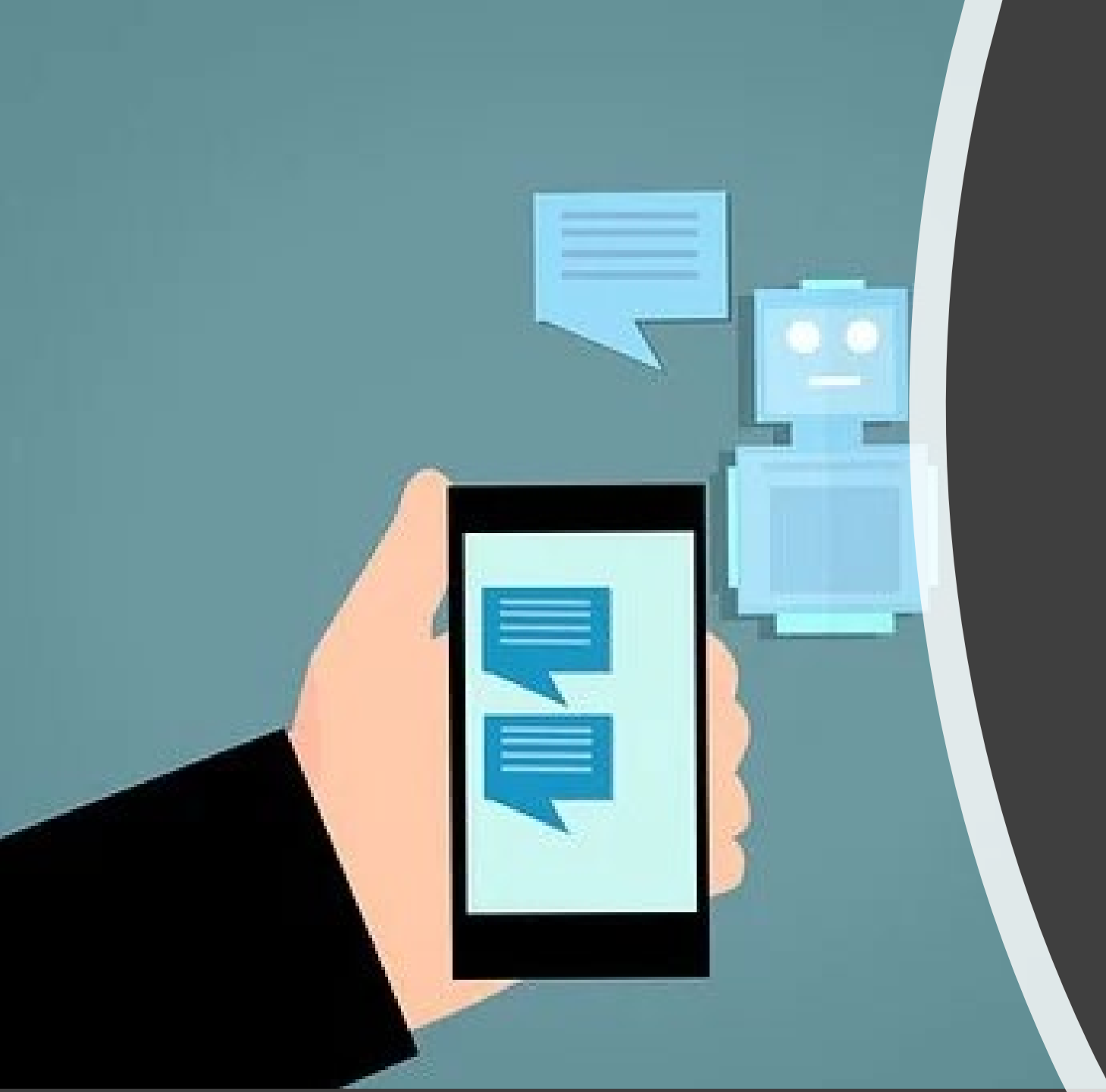




Konzept

Lessons Learned

- Sehr komplexer Ansatz/Zielstellung durch breite Themen
- Freie Arbeitsweise
 - Fokus auf Verständnis anstatt auf Perfektion
- Datensatz war qualitativ verbesserungswürdig
- Themen waren zu breit für die geringe Größe des Datensatzes



Fragen?