

SOCIAL MEDIA Trend Analyse

ML Operations • Sofie Pischl • 17 Mai 2025



Inhaltsverzeichnis



O1
Die Vision

02

Die Daten

03

Die Architektur



O4Das Ergebnis

05



Rückblick





01 Die Vision







Umfassende Social Media Trend Analyse

Welche Themen verlieren oder gewinnen an Beduetung?

Welche Themen sind grade plattformübergreifend aktuell?



Marketing

Influencer & Content Creator

Medien & Journalisten

Politische Akteure



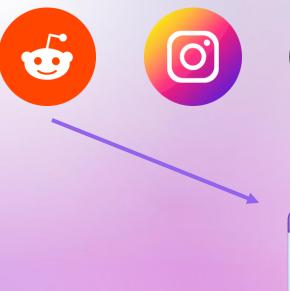


02 Die Daten





Die Daten

















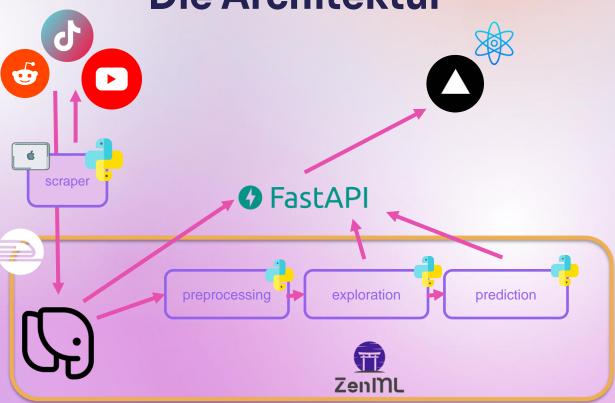
O3
Die
Architektur







Die Architektur









preprocessing

- Textbereinigung (cleaned_text): Kleinbuchstaben, URLs, HTML, Sonderzeichen, Zahlen entfernen
- Textfeatures: Textlänge, Wortanzahl, URL-, Mention- und Hashtag-Erkennung
- Hashtags extrahieren als neue Spalte
- Plattform-Features: One-Hot-Encoding (platform_*)
- Fehlende Werte: Strings → leer, Zahlen → O
- Zeitfeatures (bei timestamp): Wochentag, Uhrzeit, Wochenende
- Engagement normalisieren: Plattform- oder globalbasierte Z-Skalierung







exploration

- Plattformverteilung Beitragshäufigkeit und Prozentanteile je Plattform
- Engagement-Analyse Kennzahlen (Ø, Median, Std, Min, Max) pro Plattform
- Zeitliche Muster Beiträge nach Stunden, Wochentagen, Monaten gruppiert
- Sentimentverteilung O-Stimmung insgesamt und nach Plattform Einordnung in Kategorien: negativ, neutral, positiv
- Häufigste Wörter Top-N Wörter aus lemmatized_text
- Korrelationen Zusammenhang zwischen Engagement, Sentiment & Zeitmerkmalen
- **Ergebnisexport** Ergebnisse als JSON-Datei speicherbar (data/processed/)







predictions

- Feature-Engineering
 - TF-IDF-Vektorisierung + numerische Features (Zeit, Sentiment)
- Modelltraining
 - Random Forest Regressor + Test/Train-Split
- Evaluation
 - Metriken: MSE, RMSE, MAE, R², Varianz, Median-Fehler
 - K-Fold-Cross-Validation zur Stabilitätsprüfung
- Feature Importance
 - Analyse der wichtigsten Text- & numerischen Einflussgrößen
- Vorhersage & Analyse
 - Engagement-Werte für neue Daten prognostiziert

ARCHITEKTUR





04 Das Ergebnis



ARCHITEKTUR





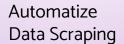
05 Rückblick

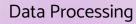




Meine Erfahrung







Was ist ML OPS überhaupt?

Deployment



Data Saving



Gelerntes





- Verschiedene einfache Orchestrierungstechniken für das Data Scraping
- Viele Herausforderungen beim Deployment kennen gelernt
- Herausforderung von ML OPS verstanden ©
- Erfahrung ganze Anewndung alleine erstellen



Optimierungspotenziale



- Effizienteres Data Processing Nur neue, noch nicht verarbeitete Daten erneut analysieren
 Konsistente und automatisierte Datensammlung
- Anbindung weiterer Plattformen
- Verbessertes Topic- & Sentiment-Modelling
- Themen & Inhalte durchsuchbar machen
- Nutzungsbasierte Gewichtung der Daten Daten aus häufig genutzten Plattformen stärker berücksichtigen
- Plattformspezifische AnalysenModellparameter im Frontend anpassbar machen



THANKS

Do you have any questions?



CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon** and infographics & images by **Freepik**







