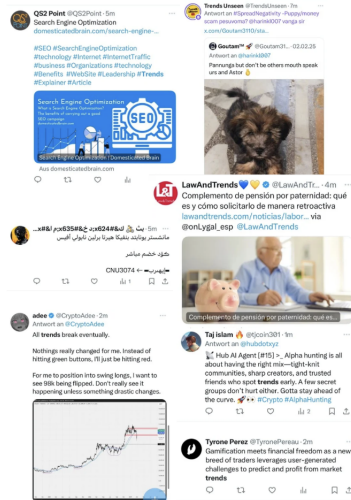


Stimmungsanalyse mit Twitter

Anne Huber, Andreas Franke, Felix Lindner, Burak Özkan, Milomir Soknic

Projektpraktikum Web Science,
Artificial Intelligence Group,
Universität Hagen, Deutschland

18. März 2025



- ▶ Twitter als Echtzeit-Plattform für Meinungen und Trends
- ▶ Große Datenmengen für maschinelles Lernen nutzbar
- ▶ Herausforderungen: Ironie, Sarkasmus, Emojis, Abkürzungen
- ▶ Einsatz in Politik, Marketing und Krisenmanagement

- ▶ Wie effektiv sind verschiedene maschinelle Lernverfahren bei der Stimmungsanalyse von Tweets?

1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

- ▶ Prüfung diverser Datensätze
- ▶ Entscheidung für „Sentiment140“
- ▶ Besonderheiten:
 - ▶ Artikel: „Twitter Sentiment Classification using Distant Supervision“
 - ▶ Bessere Datenqualität
 - ▶ Ausbalancierte Klassen
 - ▶ Emoticons als Sentiment-Indikatoren

Beispieltweet:

Original Tweet

Just got my dream job! So excited! :)



Datensatz

Just got my dream job! So excited!

Stimmung: Positiv

1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

Eigenschaften:

- ▶ Enthält **359** manuell gesammelte Tweets
- ▶ **177** negative und **182** positive Tweets
- ▶ Keine automatische Annotation durch Emoticons
- ▶ Dient zur unabhängigen Evaluierung von Modellen
- ▶ Testdatensatz enthält zusätzlich *Query-Terms*

Beispieltweet:

„no. it is too big. I'm quite happy with the Kindle2 “

Ohne Query Term: **Negativ**

Mit Query Term: **Positiv**

1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

Klassische Methoden

- ▶ **Logistische Regression**
- ▶ **Support Vector Machine (SVM)**
- ▶ **Naiver Bayes**
- ▶ Entscheidungsbäume
- ▶ Entscheidungswälder
- ▶ K-nächste Nachbarn

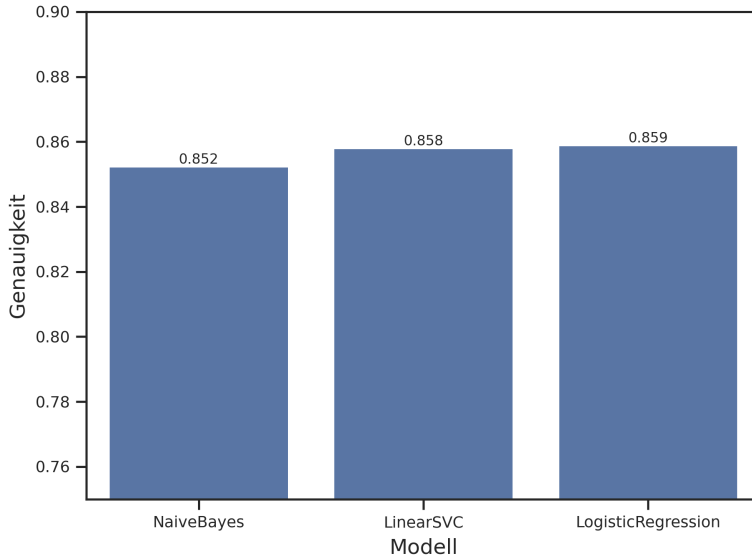
Parameter

- ▶ Vektorisierungsmethode
- ▶ Normalisierungsstrategie
- ▶ Strategie zur Entfernung von Stoppwörtern
- ▶ N-Gramm-Bereich
- ▶ Maximale Anzahl an Merkmalen

Metrik: Genauigkeit

	Beispiel
Original-Tweet	<i>„@user I love this movie! http://example.com“</i>
Bereinigung	<i>„I love this movie“</i>
Tokenisierung	<i>[„I“, „love“, „this“, „movie“]</i>
Transformation	Lemmatization: <i>[„I“, „love“, „this“, „movie“]</i> Stemming: <i>[„I“, „lov“, „thi“, „movi“]</i>
Stoppwörter Behandlung	Ohne Stoppwörter: <i>[„love“, „movie“]</i>
Merkmalsextraktion	TF-IDF Beispiel: <i>(love: 0.75, movie: 0.85)</i>

Klassische Methode - Beste Genauigkeit pro Modell



1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

- ▶ 2018 von Google entwickelt
- ▶ *Bidirectional encoder representations from transformers* (BERT)
- ▶ Mit großem textuellen Korpus vortrainiert
- ▶ 110 Mio. Parameter
- ▶ Etabliert in der natürlichen Sprachverarbeitung (NLP)

DistilBERT-base-uncased

- ▶ Destilliertes BERT-Modell
- ▶ **Eigenschaften:**
 - ▶ Modelldestillation des bert-base-uncased Modells
 - ▶ 40% kleiner
 - ▶ 60% schnellere Inferenz
 - ▶ 97% der Fähigkeiten bleiben erhalten

Twitter-RoBERTa-base-sentiment

- ▶ Auf Twitter-Daten trainiertes RoBERTa-Modell
- ▶ **Eigenschaften:**
 - ▶ Mit 58 Mio. englischsprachigen Tweets weitertrainiert
 - ▶ *Finetuning* mit einem Stimmungsanalyse-Datensatz

1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

▶ Methode

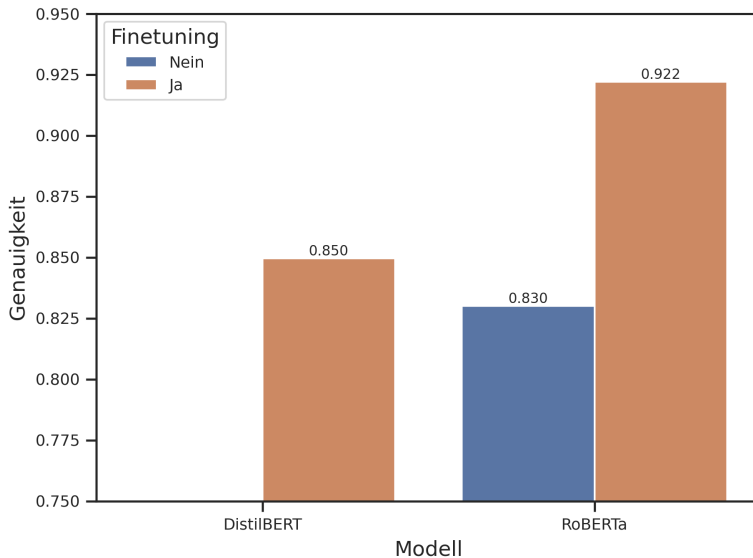
- ▶ *Finetuning* der Modelle mit Sentiment140
- ▶ Huggingface *transformers* Bibliothek

▶ Untersuchte Parameter

- ▶ Initiale Lernrate
- ▶ Größe des Trainingsdatensatzes

▶ Evaluationsmetrik

- ▶ Genauigkeit



1 Daten

- Datenauswahl
- Testdatensatz

2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

- ▶ **Reasoning-Modell**

- ▶ Basiert auf Transformer Architektur
- ▶ Trainiert mit Hilfe von *Reinforcement Learning*
- ▶ 671 Mrd. Parameter
- ▶ Destillierte Modelle verfügbar

- ▶ **Finetuning Ansatz**

- ▶ *DeepSeek-R1-Distill-Qwen-1.5B*
- ▶ Aufgrund von Hardwareanforderungen verworfen

Zero-Shot-Ansatz:

- ▶ Prompt enthält keine Beispiele
- ▶ Verwendete Modelle
 - ▶ DeepSeek-1.5B
 - ▶ DeepSeek-8B
 - ▶ DeepSeek-32B
 - ▶ DeepSeek-70B
- ▶ Inferenz mit und ohne Query-Terms

Prompt: Tweet sentiment? Sentiment Topic: {Query-Term}
Answer with positive or negative. Provide reasoning in JSON.
Tweet: „{tweet}“

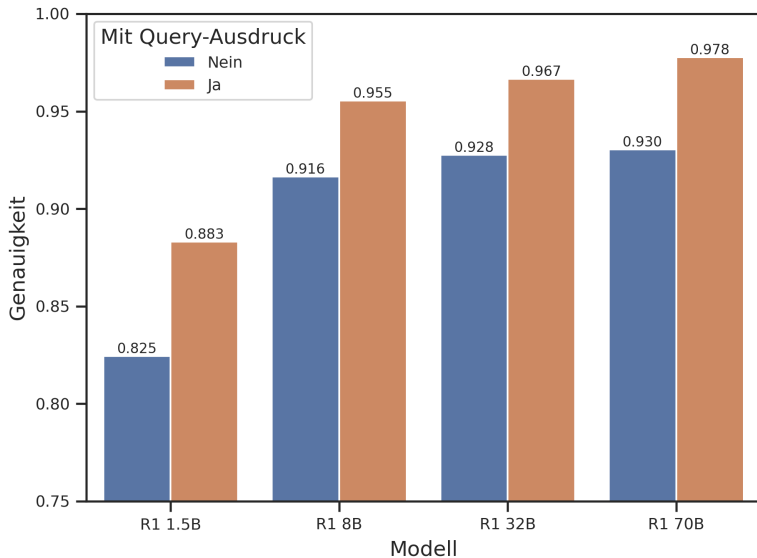
Tweet: „no. it is too big. I'm quite happy with the Kindle2 “

Reasoning ohne Query Term

„The tweet expresses dissatisfaction with something being 'too big,' indicating a negative sentiment.“

Reasoning mit Query Term:

„The tweet mentions being 'quite happy' with the Kindle2, which indicates a positive sentiment.“



1 Daten

- Datenauswahl
- Testdatensatz

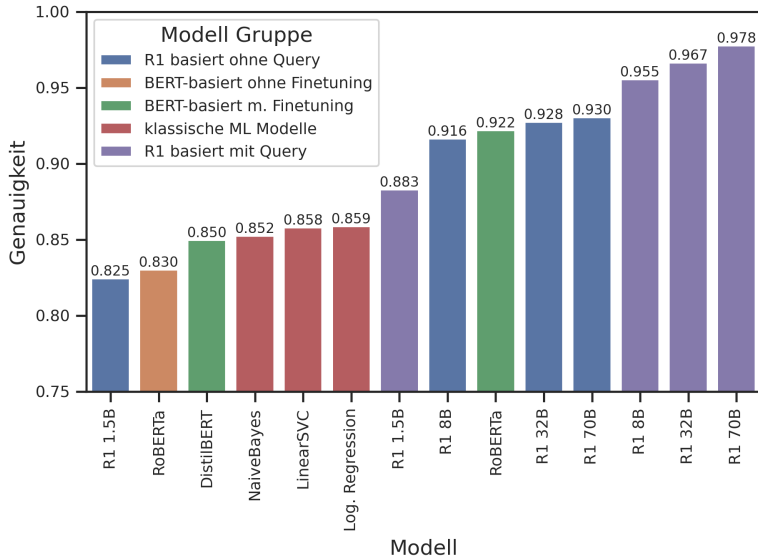
2 Klassische Methoden

3 Deep Learning

- Deep Learning - BERT-Modelle
- Deep Learning - *Finetuning* BERT-Modelle
- DeepSeek

4 Zusammenfassung

Zusammenfassung - Überblick Ergebnisse alle Modelle



► Erkenntnisse:

- LLMs führen zu höheren Genauigkeiten als klassische Verfahren
- *Finetuning* erhöht Genauigkeit der LLMs
- Query-Terms erhöhen die Genauigkeit für LLMs erheblich
- LLMs haben viel größere Hardwareanforderungen und längere Inferenzzeiten

► Ausblick:

- *Noisy-Label*
- *Aspect-Based-Sentiment-Analysis* (ABSA)
- Große *Deep Learning* Modelle

► Erkenntnisse:

- LLMs führen zu höheren Genauigkeiten als klassische Verfahren
- *Finetuning* erhöht Genauigkeit der LLMs
- Query-Terms erhöhen die Genauigkeit für LLMs erheblich
- LLMs haben viel größere Hardwareanforderungen und längere Inferenzzeiten

► Ausblick:

- *Noisy-Label*
- *Aspect-Based-Sentiment-Analysis* (ABSA)
- Große *Deep Learning* Modelle

Vielen Dank für Ihre Aufmerksamkeit!

Fragen?