

Vergleich dreier Textanalyse-Modelle (BoW, TF-IDF, SBERT) zur automatischen Sentiment-Erkennung in Schweizer Mundart-Chatnachrichten

Zwischenpräsentation

24.10.2025 | Raphael Weiss

Wie gut schneiden BoW, TF-IDF und SBERT bei der Sentimentanalyse von Mundart-Chatnachrichten ab?

Oder: Welches Modell versteht:

jo, mega guet... 🙄

Agenda

- Ziel & Datensatz
- Modelle
- Deskriptive Analyse
- Meine Erwartungen
- Modellevaluation
- Visualisierung (PCA+KMeans, 3D)
- Demo (Interaktive Eingabe)
- Key-Findings und mögliche Verbesserungen

Ziel & Datensatz

- **Ziel**

- Textklassifikation
(negativ/neutral/positiv)
für Schweizer Mundart
Chatnachrichten
- Vergleich:
 - BoW + LR (mit
Preprocessing)
 - TF-IDF + LR (mit Uni- und
Bigram und Preprocessing)
 - SBERT + LR (mit
paraphrase-multilingual-Mi
niLM-L12-v2)

- **Datensatz**

- CSV: mundart_chat.csv
(Spalten: text, label)
- 600 Zeilen
 - hammer 🥰, *positiv*
 - mega blööd 👎, *negativ*
 - maal luege, *neutral*
 - solala, *neutral*
 - richhitg nice 🎉, *positiv*
 - so huere pienlich 😞, *negativ*
- Train/Test-Split: 75/25

Modelle

- **BoW** («Bag of Words»)
 - Methode, die zählt, wie oft Wörter (oder N-Gamme) in einem Datensatz (hier: alle Chatnachrichten) vorkommen, unabhängig vom Kontext
- **TF-IDF** («Term Frequency Inverse Document Frequency»)
 - Baut auf BoW auf und gewichtet die Wörter (oder N-Gramme) zusätzlich nach ihrer Wichtigkeit im Datensatz (hier: alle Chatnachrichten)
- **SBERT** («Sentence Bidirectional Encoder Representations from Transformers»)
 - Vortrainiertes neuronales Sprachmodell, das darauf optimiert ist, ganze Sätze oder Texte als semantische Vektoren (Embeddings) darzustellen



**(multinomiale)
logistische
Regression als
Klassifikator**

Deskriptive Analyse

```
– Klassenverteilung –  
label  
neutral    199  
negativ    199  
positiv    199  
Name: count, dtype: int64  
  
– Länge (Zeichen) –  
mean      16.063652  
50%       15.000000  
min        2.000000  
max       37.000000  
Name: text, dtype: float64
```

Deskriptive Analyse

```
- Top-1g: negativ -  
ngram count  
🤔 49  
👉 44  
💔 44  
😡 41  
😭 37
```

```
- Top-1g: neutral -  
ngram count  
😐 71  
😏 65  
😄 60  
so 49  
😬 26
```

```
- Top-1g: positiv -  
ngram count  
😊 42  
✨ 39  
💕 35  
😍 34  
👍 32
```

```
- Top-2g: negativ -  
ngram count  
gar nöd 12  
nie wieder 12  
so en 11  
nöd guet 8  
mega blöd 6
```

```
- Top-2g: neutral -  
ngram count  
i o 13  
so i 12  
jo passt 11  
neutral gseh 11  
zur kenntnis 11
```

```
- Top-2g: positiv -  
ngram count  
top sache 11  
passt perfekt 10  
gfallt mir 7  
bin zfrieden 6  
gute idee 6
```

```
- PMI-2g: negativ -  
bigram count PMI  
so en 11 5.828665  
gar nöd 12 5.713188  
nöd guet 8 5.713188  
en quatsch 4 5.565631  
funktioniert nid 5 5.539159
```

```
- PMI-2g: neutral -  
bigram count PMI  
ist okay 5 6.011663  
kann maan 3 5.993047  
man machen 8 5.944549  
i o 13 5.886132  
kann man 9 5.855544
```

```
- PMI-2g: positiv -  
bigram count PMI  
cool gmacht 4 6.291554  
bin zfrieden 6 5.829449  
gfallt mir 7 5.732587  
passt perfekt 10 5.661504  
top sache 11 5.439112
```

*PMI: Pointwise Mutual Information
/ Wie stark Wörter gemeinsam
auftreten, verglichen mit Zufall*

Meine Erwartungen (gemäss Theorie😊)

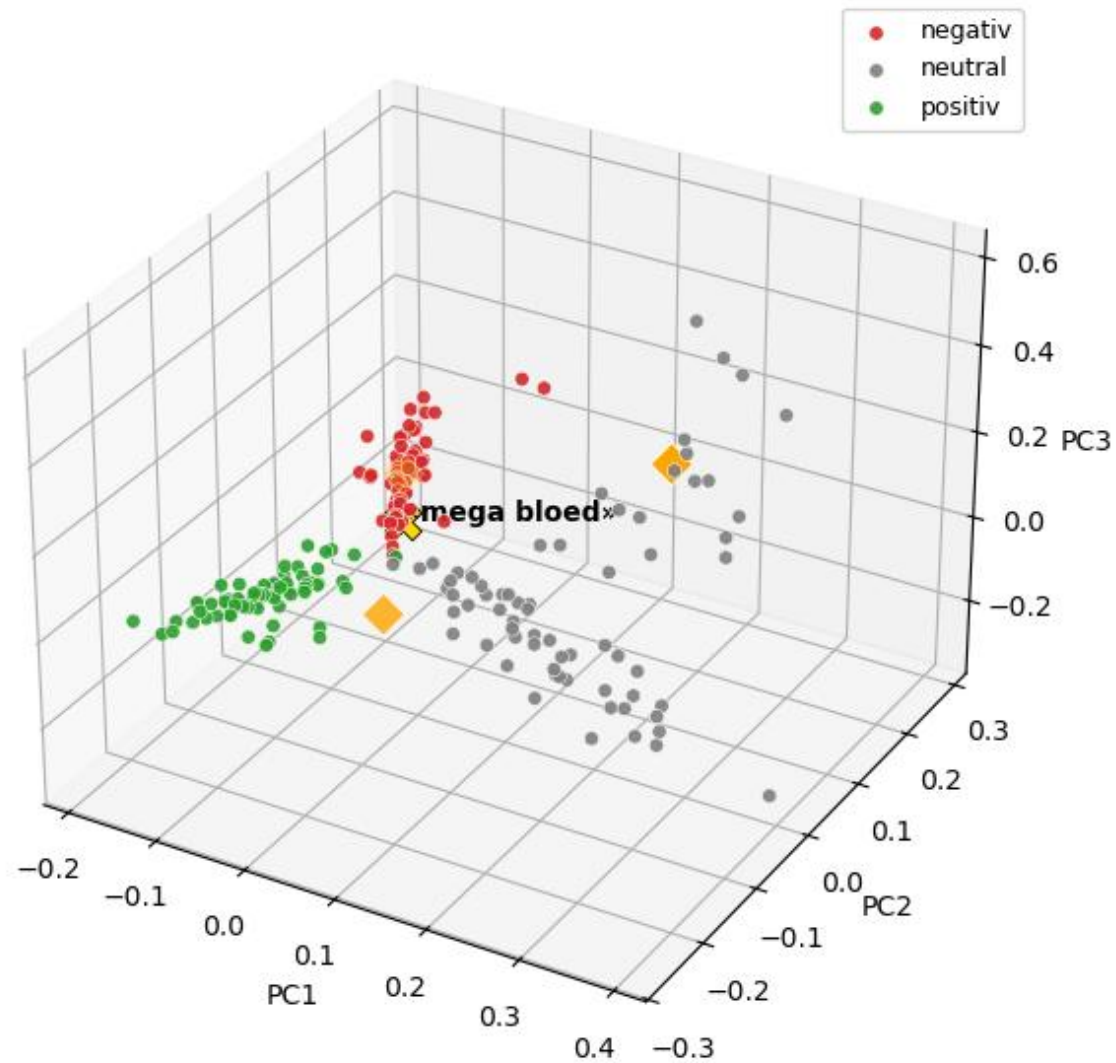
	BoW	TF-IDF (mit Bigram)	SBERT
Klare Keywords	✓ stark	✓ stark	✓ stark
Viele Füllwörter	⚠ schwächer	✓ stark	✓ stark
Synonyme/Paraphrasen	⚠ teils, da hier mit Dialekt-Mapping	⚠ teils, da hier mit Dialekt-Mapping	✓ stark (auch ohne Dialekt-Mapping)
Negationen („nicht gut“)	✗ schwach, da hier kein Bigram	⚠ teils, da hier Bigram	✓ stark
Rechtschreibfehler / Elongation	✗ schwach, braucht Vorkommen	⚠ teils (dank Bigram)	✓ robust
Emojis als Signal	⚠ teils, braucht Vorkommen	⚠ etwas besser (dank Gewichtung)	✓ robust
kurze Fragmente (1–3 Wörter)	✓ stark	✓ stark	⚠ teils schwächer, da keine Satzstruktur

Modevaluation

Modell	Accuracy	Kommentar
BoW + LogisticRegression	0.986	Nahezu perfekt – lernt klare Wörter, aber wohl overfit, viele Wiederholungen im Datensatz
TF-IDF + LogisticRegression	0.993	Ebenfalls sehr stark, leicht robuster durch Gewichtung.
SBERT + LogisticRegression	0.920	Semantisch robust, aber niedrigere Scores (realistischer).

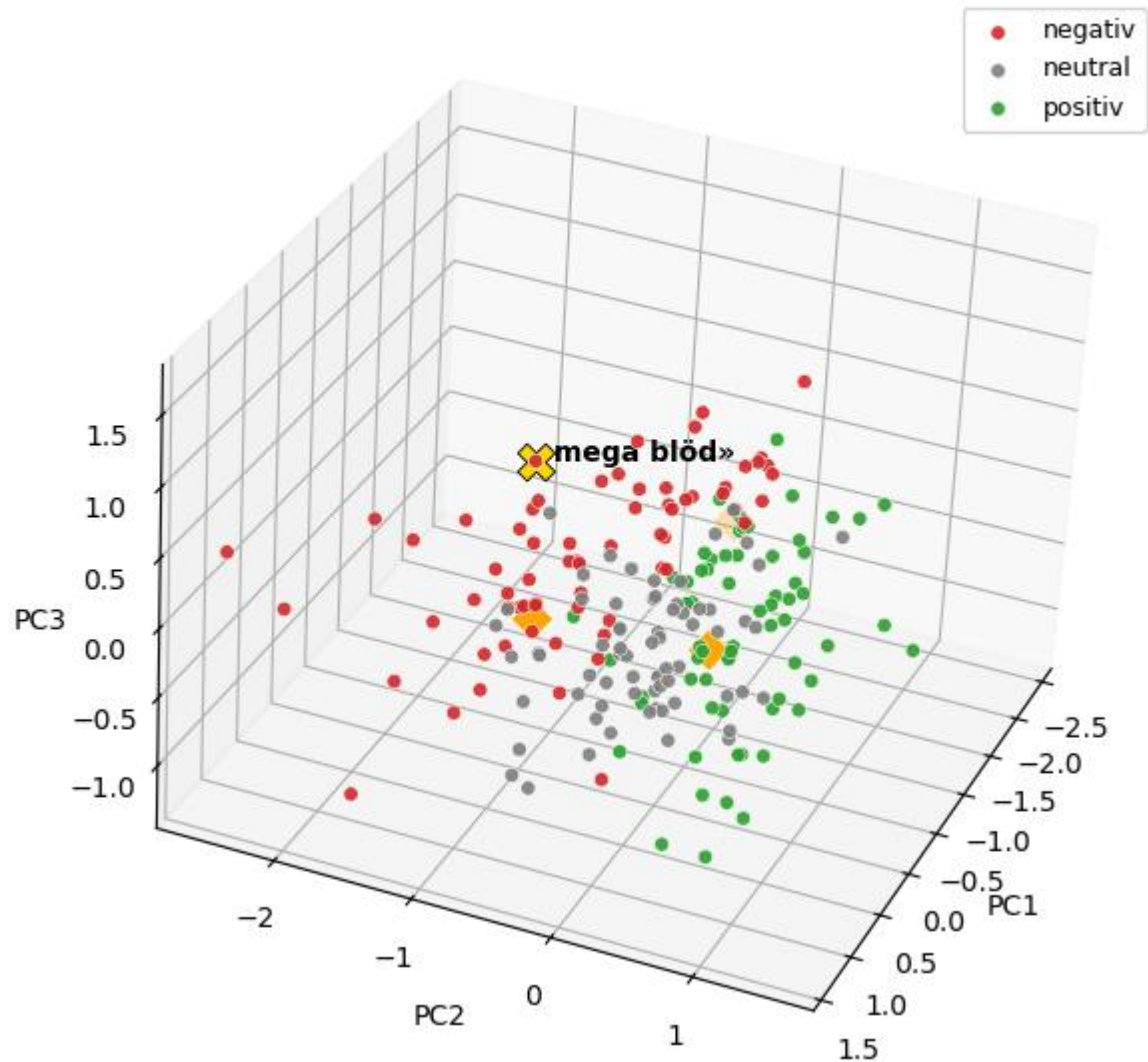
Visualisierung

PCA+KMeans (3D) - TF-IDF



Visualisierung

PCA+KMeans (3D) - SBERT



Demo

```
> Textanalütixs isch mega cool
```

```
– Ergebnisse –
```

BoW	-> positiv		negativ: 0.23		neutral: 0.34		positiv: 0.43
TF-IDF	-> positiv		negativ: 0.25		neutral: 0.35		positiv: 0.40
SBERT	-> positiv		negativ: 0.04		neutral: 0.25		positiv: 0.71

Demo

```
> mega blöd
```

```
– Ergebnisse –
```

BoW	-> negativ	negativ: 0.64	neutral: 0.12	positiv: 0.24
TF-IDF	-> negativ	negativ: 0.61	neutral: 0.17	positiv: 0.22
SBERT	-> negativ	negativ: 0.95	neutral: 0.02	positiv: 0.03

```
> mega blööd
```

```
– Ergebnisse –
```

BoW	-> negativ	negativ: 0.48	neutral: 0.16	positiv: 0.36
TF-IDF	-> negativ	negativ: 0.44	neutral: 0.24	positiv: 0.32
SBERT	-> positiv	negativ: 0.23	neutral: 0.08	positiv: 0.70

```
> so mega blöd
```

```
– Ergebnisse –
```

BoW	-> neutral	negativ: 0.44	neutral: 0.47	positiv: 0.10
TF-IDF	-> negativ	negativ: 0.51	neutral: 0.32	positiv: 0.17
SBERT	-> negativ	negativ: 0.80	neutral: 0.16	positiv: 0.03

```
> das isch so mega blöd
```

```
– Ergebnisse –
```

BoW	-> neutral	negativ: 0.19	neutral: 0.77	positiv: 0.04
TF-IDF	-> neutral	negativ: 0.40	neutral: 0.46	positiv: 0.15
SBERT	-> negativ	negativ: 0.74	neutral: 0.21	positiv: 0.05

```
> das isch so übel
```

```
– Ergebnisse –
```

BoW	-> neutral	negativ: 0.07	neutral: 0.88	positiv: 0.05
TF-IDF	-> neutral	negativ: 0.10	neutral: 0.81	positiv: 0.08
SBERT	-> negativ	negativ: 0.85	neutral: 0.14	positiv: 0.01

Key-Findings

- BoW/TF-IDF: überraschend stark bei kurzen Chatnachrichten
- SBERT: robuster bei neuen Formulierungen / Wörtern (semantischer Transfer), aber Mühe bei kurzen Sätzen (fehlender Kontext)
- Accuracy (99 %) ist nicht gleichbedeutend mit guter Generalisierung
- Schwierigkeiten mit Mundart Chatnachrichten:
 - *nicht standardisiert: „nöd“, „nid“, „ned“*
 - *mischt Dialekt, Hochdeutsch, Slang, Englisch*
 - *drückt viel über Tonfall, Kontext und Ironie aus, insbesondere mit Emoji. „jo, mega guet... 🙄“*

Key-Findings

```
> jo, mega guet... 🙄
```

```
– Ergebnisse –
```

BoW -> positiv	negativ: 0.26	neutral: 0.22	positiv: 0.52
TF-IDF-> positiv	negativ: 0.22	neutral: 0.24	positiv: 0.54
SBERT -> negativ	negativ: 0.54	neutral: 0.44	positiv: 0.01

Verbesserungen

	Hauptwirkung auf	Nutzen / Wirkung
Mehr (realistische) Daten & Kontext!	SBERT	Verbessert semantische Repräsentationen.
→ Vocabulary Growth Curve zur Validierung der Sprachvielfalt mit zunehmenden Daten → Zipf's Law zur Validierung der natürlichen Wortfrequenzverteilung		
Stoppwort-Optimierung	BoW / TF-IDF	Reduziert Rauschen, stärkt relevante Signale.
Bessere Normalisierung & Lemmatization, Dialekt-Mapping (auch für SBERT!)	alle	Erkennt Varianten & Schreibweisen besser.
Char-n-grams oder Subwords	BoW / TF-IDF	Erhöht Robustheit bei Tippfehlern & Dialektformen
N-Gram-Optimierung (1–3)	BoW / TF-IDF	Erkennt längere Phrasen & Negationen („nicht mega guet“)