



Krypto-Stimmungen erkennen – Vergleich moderner Sprachmodelle

Ein Projekt zur Sentimentanalyse auf Reddit-Posts und -Kommentaren

Warum Reddit? Warum NLP?

- Reddit ist ein zentraler Ort für Krypto-Diskussionen Meinungen oft direkt & subjektiv
- Datenmenge ist zu groß für manuelle Auswertung
- Klassische Tools erfassen Sprache nur oberflächlich
- **Ziel:**
*Automatisierte Einordnung von Reddit-Beiträgen in **bullish**, **neutral** oder **bearish** mithilfe von NLP*



Zwei Datensätze – Kommentare & Posts separat behandelt

◆ Posts (Langtexte)

- Quelle: Eigene Reddit-Scrapes, gelabelt mit GPT-Unterstützung
- Trainingsdaten: 1.280 Texte, Validation: 320
- Klassenverteilung im Training:
 - neutral (1): 842
 - bullish (0): 391
 - bearish (2): 47
- Testset: 400 Texte, bullish nur 15

◆ Comments (Kurztexte)

- Quelle: IEEE-Datensatz mit 599.975 gelabelten Kommentaren
- Klassenverteilung im Training:
 - bullish (2): 305.079
 - neutral (1): 239.043
 - bearish (0): 55.853
- Testset: 999 Kommentare, gleichverteilt (je 333 pro Klasse)

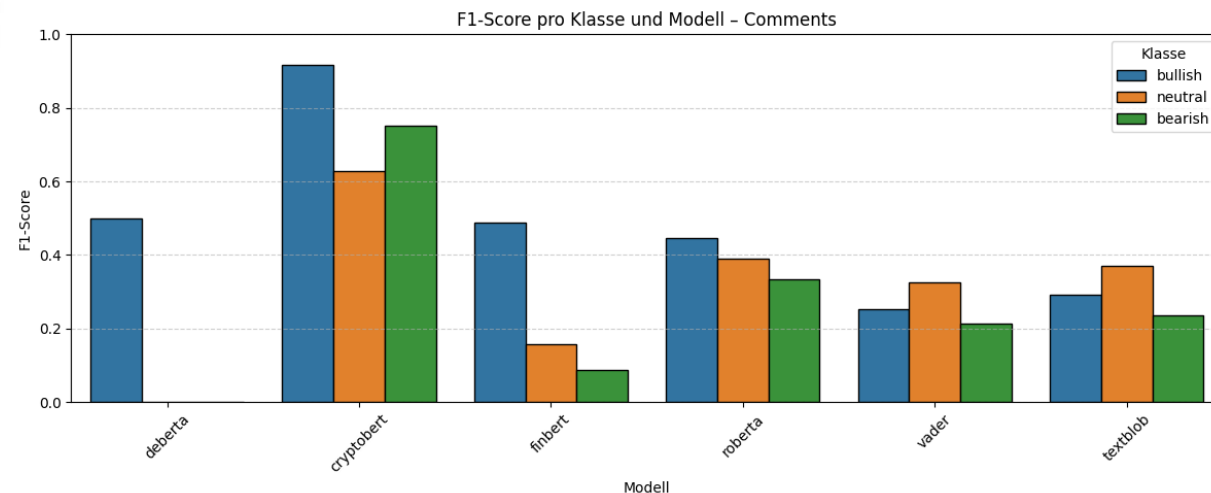
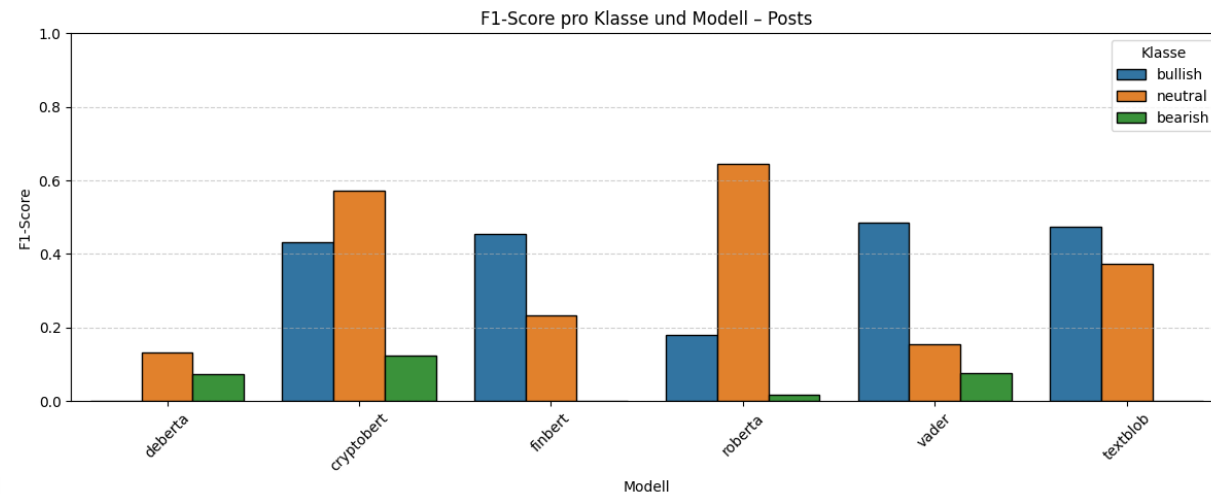
Mehrere Modelle im Vergleich

◆ Transformer-Modelle

- **CryptoBERT** überzeugt bei Kommentaren – vor allem bei bullish (F1: ~0.91) und bearish (F1: ~0.75)
- **FinBERT** zeigt bei Posts Potenzial bei bullish, schwächelt jedoch bei neutral und bearish
- **RoBERTa & DeBERTa** liefern inkonsistente Ergebnisse – kein klarer Favorit

● Klassische Modelle (Text-basiert)

- **VADER & TextBlob** stabil bei neutral, aber sehr schwach bei bullish/bearish
- Beide Methoden ungeeignet für differenzierte Stimmungsanalyse auf Reddit

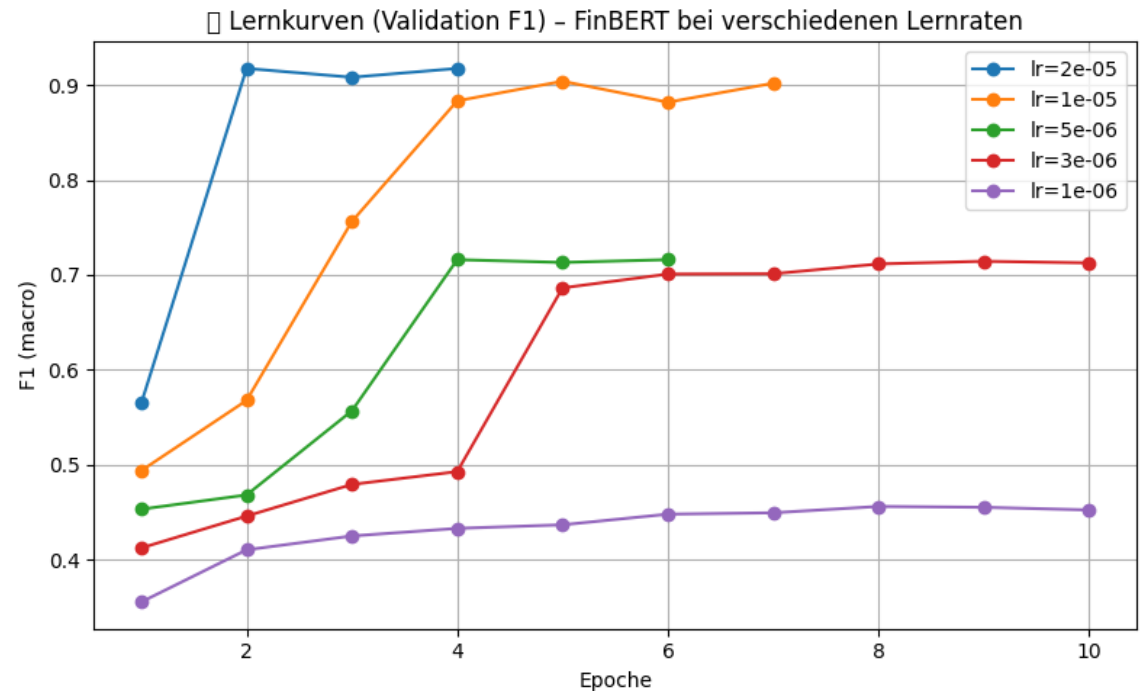
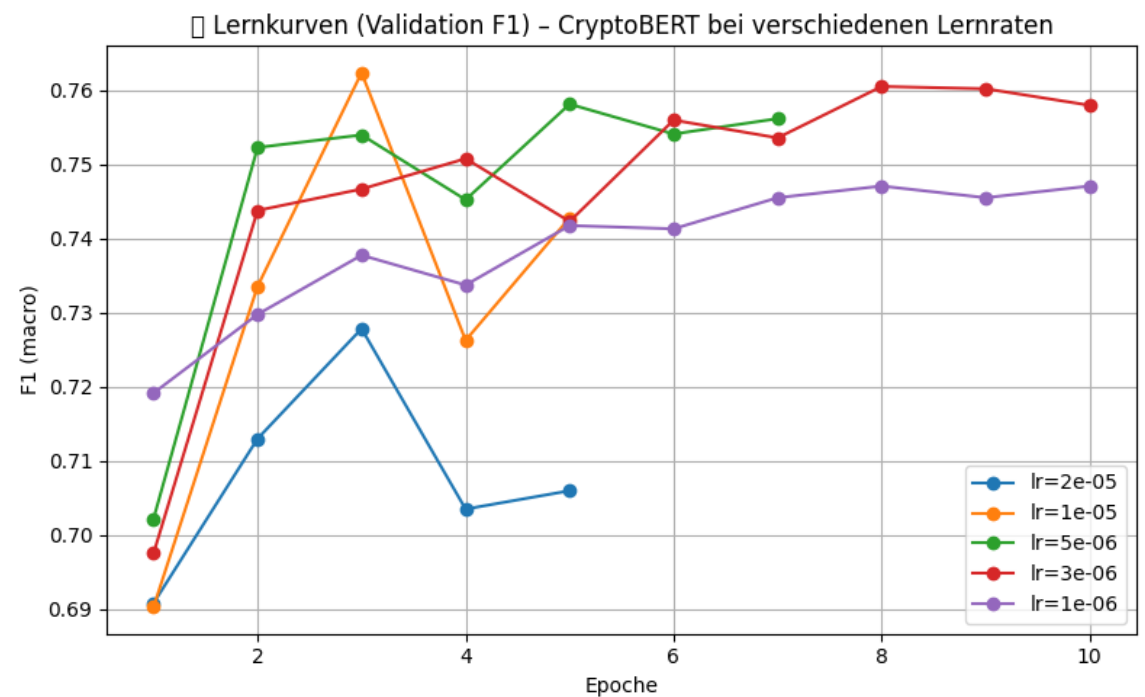


Finetuning: Wie wurden die Modelle verbessert

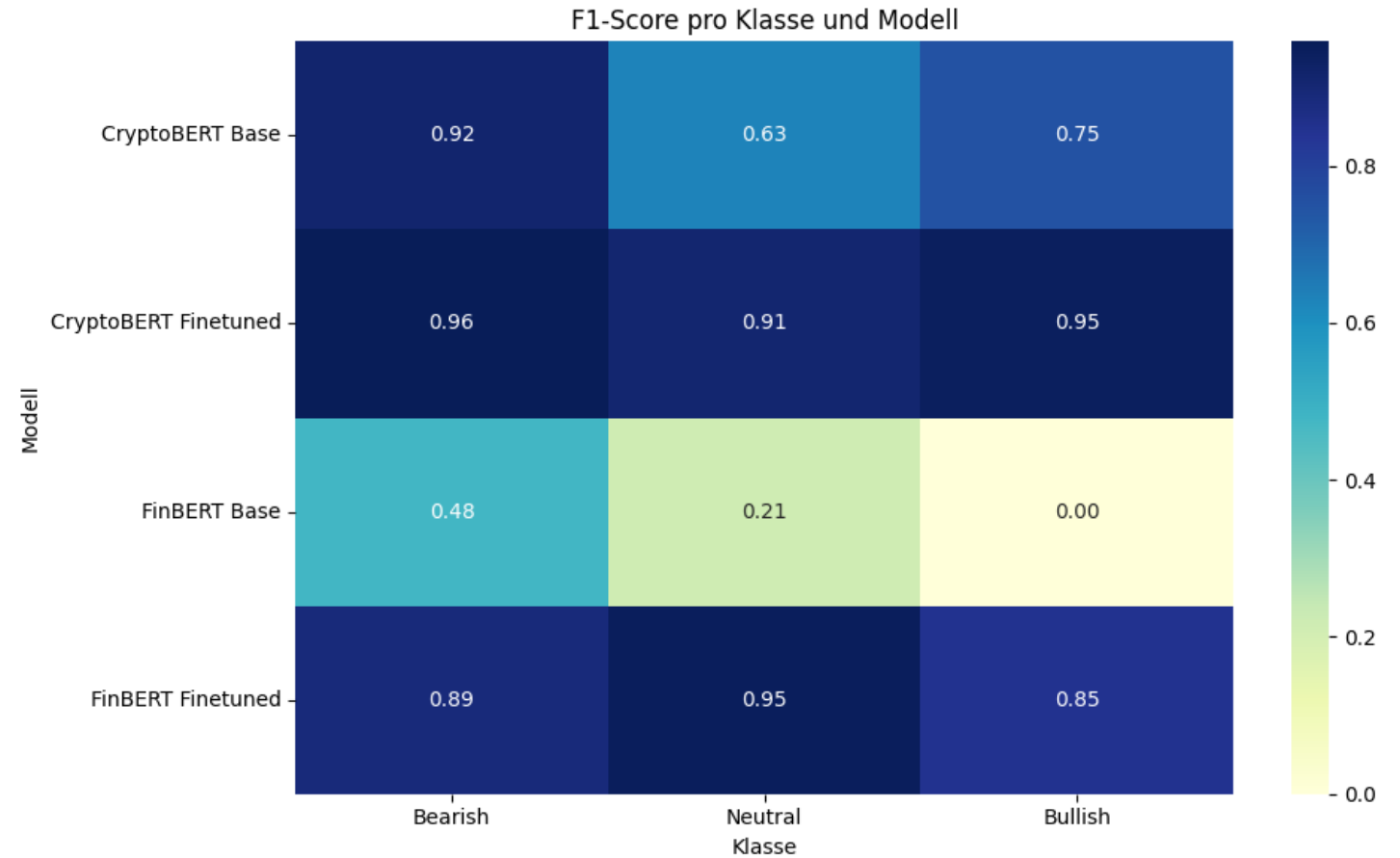
- Modelle separat trainiert: **CryptoBERT** (Comments), **FinBERT** (Posts)
- Verschiedene **Lernraten** getestet (z. B. $1e-5$, $2e-5$, $5e-6$)
- **Early Stopping** & beste F1-Score auf Validation-Set → Auswahlkriterium
- Verwendung einer **gewichteten Verlustfunktion**, um Klassen-Ungleichgewicht auszugleichen
- Ergebnis: **zwei spezialisierte, feinjustierte Modelle** für unterschiedliche Texttypen

Feintuning-Qualität durch Lernkurven

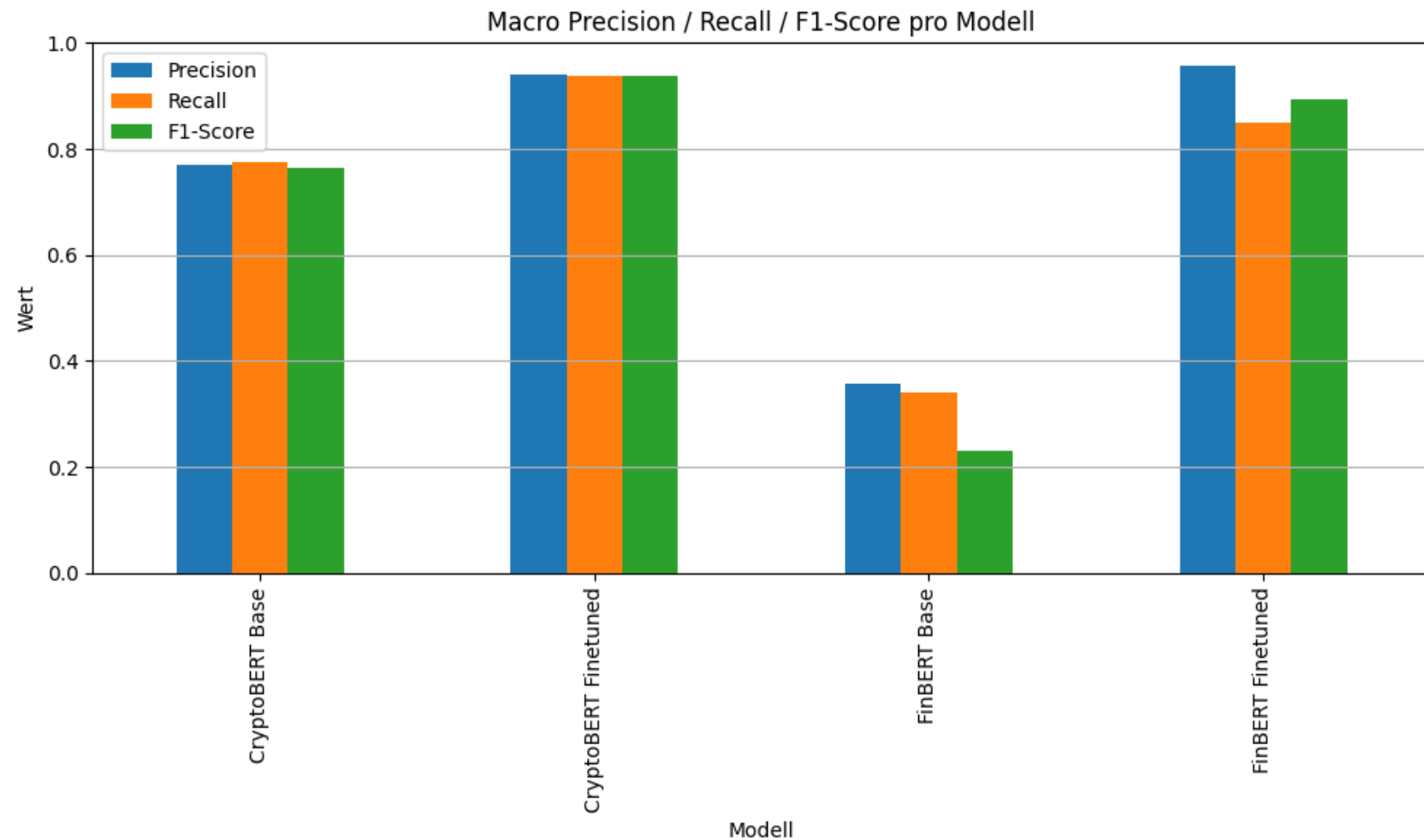
- CryptoBERT: **lr=1e-05** zeigt besten Peak – stabil & performant
- FinBERT: **lr=2e-05** mit rasantem Lernfortschritt und Top-F1
- Niedrige Lernraten lernen stabiler, aber weniger effektiv
- Entscheidung basierte auf bestmöglichem **Validation F1 (macro)**



Deutlich bessere Ergebnisse durch Finetuning



Ergebnisse pro Modell im Überblick



FinBERT (Posts)

- **Base:** F1-Score (macro): **0.23**, Accuracy: **0.35**
- **Finetuned:** F1-Score (macro): **0.89**, Accuracy: **0.93**

CryptoBERT (Comments)

- **Base:** F1-Score (macro): **0.77**, Accuracy: **0.77**
- **Finetuned:** F1-Score (macro): **0.94**, Accuracy: **0.94**

A large orange circle occupies the left side of the slide, partially cut off by the edge.

Sentimentanalyse, die funktioniert

- **Automatisierte Analyse** von Krypto-Diskussionen auf Reddit
- **Schnelle Erkennung** von Marktstimmungen ohne manuelle Auswertung

