

Implementierung und Evaluation von Word Embeddings für die Emotionsklassifikation in deutschsprachigen historischen Dramen

Thema:

Implementierung und Evaluation von Word Embeddings für die Emotionsklassifikation in deutschsprachigen historischen Dramen

Art:

MA

BetreuerIn:

Thomas Schmidt

BearbeiterIn:

Alina El-Keilany

Status:

in Bearbeitung

Stichworte:

Word Embeddings, Emotionsklassifikation, NLP, Digital Humanities, CLS, Word2Vec, fastText

angelegt:

2024-04-01

Antrittsvortrag:

2024-04-29

Hintergrund

Die Repräsentation von Wörtern durch Word Embeddings (wie etwa Word2Vec [1], GloVe [2] oder fastText [3]) stellt eine wichtige Technik für das Natural Language Processing (NLP) dar. Auch für die Digital Humanities (DH) sind Word Embeddings von Bedeutung. So können sie etwa als Forschungsgegenstände behandelt werden, um anhand ihrer zum Beispiel lexikalischen Bedeutungswandel oder gesellschaftliche Veränderungen nachzuvollziehen (so z.B. in [4] und [5]). Oder sie werden als Features für NLP-Aufgaben verwendet, wie etwa bei der Named Entity Recognition (NER), der Emotionsklassifikation und der Sentiment Analyse, welche allesamt im Bereich der quantitativen Kultur- und Literaturanalyse für die Generierung von analysierbaren Daten relevant sind.

Ein Forschungsprojekt aus dem Gebiet der Computational Literary Studies (CLS) ist das *Emotions in Drama* Projekt (EmoDrama; [6]), welches (u.a.) Emotionsannotationen für deutsche Dramen von 1650 bis 1815 bietet. Diese gelabelten Daten ermöglichen die Anwendung maschineller Lernverfahren für die computergestützte Emotionserkennung.

Wie auch in anderen Fächern ist in den Digital Humanities der Stand der Technik für die Emotionserkennung inzwischen die Nutzung von dynamischen (Satz-)Embeddings anstelle von statischen Word Embeddings, und die Verwendung von Transformer-Architekturen und Large Language Models (LLM) wie etwa BERT hat sich etabliert. Diese Modelle müssen erfahrungsgemäß aufwendig task- und domänenspezifisch angepasst werden. Statische Embeddings hingegen werden oft nur noch für Baseline-Experimente genutzt, und weder optimiert noch im Hinblick auf die

Textgrundlage ausgewählt. Es ergibt sich eine Forschungslücke im Anwendungsfeld der Klassifikation von Emotionen in historischen deutschen Texten. Denn die Optimierungsansätze, die für LLMs längst gebräuchlich sind, wurden hier noch nicht konsequent auch für statische Embeddings angewendet.

Zielsetzung der Arbeit

Im Rahmen dieser Arbeit soll die Verwendung von statischen Word Embeddings als Features für die Emotionsklassifikation von historischen deutschen Dramen untersucht werden. Das Forschungsziel ist dabei die Evaluation der Embeddings: Dafür sollen die verschiedenen Arten der Embeddings (z.B. Word2Vec, fastText, GloVe), die möglichen Trainingsmethoden und Trainingsparameter, sowie unterschiedliche Textgrundlagen für das Training implementiert und getestet werden und die Embeddings anhand verschiedener Klassifikationstasks am EmoDrama-Datensatz verglichen werden. Aus den gewonnenen Ergebnissen sollen Schlüsse sowohl für die Anwendung statischer Embeddings im allgemeinen Kontext der Emotionsklassifikation, als auch im Speziellen für den Bereich der CLS und historischer deutscher Literatur gezogen werden.

Als Deliverables entstehen zudem verschiedene historische Word Embedding Modelle, welche für andere Anwendungsfälle der Digital Humanities genutzt werden können.

Konkrete Aufgaben

- Literaturrecherche
- Sichtung, Auswahl und Akquise möglicher Word Embeddings, Korpora und Klassifikatorarchitekturen
- Implementierung des Klassifikators und Vorbereitung des EmoDrama-Datensatzes
- Erstellung eines Evaluationsframeworks
- Implementierung verschiedener vortrainierten Word Embeddings und anschließende Analyse der Performance und der Probleme bei der Verwendung moderner Embeddings für die Klassifikationsaufgabe
- Akquise und Aufbereitung verschiedener literarischer Korpora historischer deutscher Sprache (z.B. German Drama Corpus, Projekt Gutenberg DE, Deutsches Textarchiv), sowie die Erstellung eigener Spezialkorpora daraus
- Training und Optimierung verschiedener statischer Word Embeddings mit unterschiedlichen Attributen und Parametern
- Nutzung der eigenen Word Embeddings für die Klassifikationsaufgabe
- Analyse, Exploration und Diskussion der Ergebnisse
- Formulierung von Best Practices und Empfehlungen zu Word Embeddings in Kontext der Emotionserkennung in historischen Texten

Technische Umsetzung: Der praktische Teil der Arbeit wird mit Python umgesetzt. Für die Klassifikation wird mit tensorflow und der Keras-Bibliothek gearbeitet. Für das Training eigener Embeddings werden die Gensim-Bibliothek [7] und das fastText-Modul verwendet

Erwartete Vorkenntnisse

- Natural Language Processing (NLP)
- Grundlagen der Korpuslinguistik
- Grundlagen des Deep Learnings (DL)
- Python: tensorflow/ Keras, pandas, numpy, NLTK, scikit-learn, GENSIM, fastText

Weiterführende Quellen

[1] Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. In *Proceedings of Workshop at ICLR 2013*.

[2] Pennington, J., Socher, R. & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. In A. Moschitti, B. Pang & W. Daelemans (Hrsg.), *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (S. 1532-1543). Association for Computational Linguistics.

[3] Bojanowski, P., Grave, E., Joulin, A. & Mikolov, T. (2017). Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 5, 135-146.

[4] Hamilton, W. L., Leskovec, J. & Jurafsky, D. (2016). Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics.

[5] Garg, N., Schiebinger, L., Jurafsky, D. & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 115(16), 3635-3644.

[6] Dennerlein, K., Schmidt, T. & Wolff, C. (2023). EmoDrama. Ein Korpus mit Emotionsinformationen in Dramen von 1650-1815. *Zeitschrift für digitale Geisteswissenschaften*, 8.

[7] Řehůřek, R. & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (S. 45-50). ELRA.

Řehůřek, R. Gensim: Topic modelling for humans. <https://radimrehurek.com/gensim>

From:

<https://wiki.mi.ur.de/> - **MI Wiki**

Permanent link:

https://wiki.mi.ur.de/arbeiten/word_embeddings_emotionsklassifikation

Last update: **23.04.2024 11:13**

