Implementierung und systematische Evaluation von Methoden lexikonbasierter Sentiment Analyse für das Deutsche

Thema:
: Implementierung und systematische Evaluation von Methoden lexikonbasierter Sentiment Analyse für das Deutsche

Art:
: MA

BetreuerIn:
: Thomas Schmidt

BearbeiterIn:
: Jakob Fehle

ErstgutachterIn:
: Christian Wolff

Status:
: abgeschlossen

Stichworte:
: Sentiment Analysis, Text Mining, Corpus Analytics, Natural Language Processing, Digital Humanities

angelegt:
: 2020-05-18

Antrittsvortrag:
: 2020-06-15

Abschlussvortrag:
: 2021-01-11

Hintergrund

Sentiment Analysis ist eines der populärsten Forschungsgebiete im Bereich des Text Mining und Natural Language Processing. Wenngleich nicht mehr state-of-the-art, erfreuen sich lexikonbasierte Methoden der Sentiment Analysis noch immer großer Beliebtheit in sehr speziellen Domänen und „underresourced“ Sprachen (alles außer Englisch und Chinesisch). Grund hierfür ist die simple Generalisierbarkeit und die Nachvollziehbarkeit der Ergebnisse. Für das Deutsche gibt es einige Sentiment-Lexika und Vorverarbeitungsschritte, die jedoch noch nicht systematisch und erschöpfend evaluiert wurden.

Zielsetzung der Arbeit

In der vorliegenden Arbeit sollen die gängigsten Techniken und Lexika im Deutschen auf einer größeren Menge an Beispielkorpora unterschiedlicher Domänen angewandt und evaluiert werden. Ein Abgleich mit einfachen und aktuellen Machine Learning-Methoden soll die Ergebnisse in einen Gesamtkontext stellen. Es sollen damit Empfehlungen und best Practices für den Einsatz lexikonbasierter Methoden in deutscher Sprache gemacht werden.

Konkrete Aufgaben

- Einarbeitung in die Related Work
- Akquise gängiger deutscher Korpora für Sentiment Analysis
- Implementierung verschiedener lexikonbasierter Techniken
- Implementierung verschiedener Machine Learning-Techniken
- Evaluation der Ergebnisse
- Interpretation und Empfehlungen

Erwartete Vorkenntnisse

- Vorerfahrung im Bereich NLP
- Programmiererfahrung in Python und/oder R

Weiterführende Quellen

Liu, B. (2016). Sentiment Analysis. Mining Opinions, Sentiments and Emotions. New York: Cambridge University Press.

Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. Computer Science Review, 27, 16-32.

Schmidt, T., Burghardt, M., & Wolff, C. (2018). Herausforderungen für Sentiment Analysis-Verfahren bei literarischen Texten. INF-DH-2018.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), 267-307.

Vinodhini, G., & Chandrasekaran, R. M. (2012). Sentiment analysis and opinion mining: a survey. International Journal of Advanced Research in Computer Science and Soft-ware Engineering, 2(6), 282-292.

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.

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

Permanent link:
https://wiki.mi.ur.de/arbeiten/lexikonbasierte_sentia

Last update: **29.01.2021 11:51**