

Spezifikationsvortrag

Fabian Düker, Uli Steinbach

Universität Heidelberg, Institut für Computerlinguistik
Softwareprojekt, SoSe 2018

Prof Dr. Katja Markert

12.06.2018

Übersicht I

Übersicht

Autom. Erstellung eines Lexikons für die Erkennung von
Abusive Words

Anwendung auf Germeval Task I 2018

Binäre Klassifikation von 5000 Tweets

Problemstellung

- ▶ Problem: Hatespeech ist in ständiger Veränderung begriffen (Neologismen, Ambiguität, Kontext)
- ▶ Wiegand et al. 2016: Erstellung eines englischen Lexikons mit guten Ergebnissen auf cross-domain Evaluation
- ▶ SentiWS: Lexikon mit negativen Wörtern für das Deutsche

Lösungsansatz

- ▶ Erstellung Baselexikon aus SentiWS neg. Sentiment-Lexikon
- ▶ halbautomatische Erweiterung des Baselexikons mit deutschen Schimpfwörtern
- ▶ autom. Erweiterung mittels graphbasiertem Label-Propagation-Algorithmus
- ▶ Anwendung auf Germeval 2018 Datenset und Evaluation

Erstellung des Baselexikons

SentiWS

- ▶ Extraktion negativer Wörter aus SentiWS
- ▶ 686 Nomen
- ▶ 420 Verben
- ▶ 708 Adjektive
- ▶ Problem: Zu wenige explizite Schimpfwörter
- ▶ Lösung: Mehr Schimpfwörter hinzufügen

Erweiterung des Baselexikons

- ▶ Genius API: Erstellung eines Deutschrappkorpus
- ▶ Deutschrapp: zeitgemäße Verwendung von Schimpfwörtern (genrespezifisch, aber auch politisch, rassistisch, sexistisch)
- ▶ autom. Extraktion von Kandidaten mittels syntaktischer Pattern

Erweiterung des Baselexikons

- ▶ Automatischer Abgleich aller Nomen im Rapkorpus
- ▶ mit Schimpfwortliste aus dem Internet
- ▶ mit Pattern "du [NN]"
- ▶ etwa 280 potentielle Schimpfwörter
- ▶ manuelles Aussortieren von false positives (z.B. "Rapper")
- ▶ Auswahl der 200 häufigsten Schimpfwörter
- ▶ Erweiterung durch beleidigende Adjektive
- ▶ Suche nach Pattern "du [ADJ] Schimpfwort"
- ▶ etwa 280 potentiell beleidigende Adjektive

Erweiterung des Baselexikons

- ▶ Lemmatisierung mit IWNLP
- ▶ Lemmatisierung der nicht erkannten Adjektive von Hand
- ▶ Beseitigung von Duplikaten
- ▶ Finales Baselexikon:
 - ▶ 887 Nomen
 - ▶ 413 Verben
 - ▶ 824 Adjektive
 - ▶ - 2124 Wörter

Rapkorpus

- ▶ Texte von 30 willkürlich gewählten Rappern
- ▶ Bushido
- ▶ Chakuza
- ▶ K.I.Z.
- ▶ Kay One
- ▶ Kollegah & Farid Bang
- ▶ Prinz Pi
- ▶ Bass Sultan Hengzt
- ▶ Fler
- ▶ Azad
- ▶ Kool Savas
- ▶ ...

Rapkorpus

Auszug der extrahierten Schimpfwörter vor Handselektion

- ▶ *Rapper*
- ▶ *Kopf*
- ▶ **Arsch**
- ▶ **Bitch**
- ▶ **Schwanz**
- ▶ **Scheiße**
- ▶ *Gangster*
- ▶ *Block*
- ▶ *Baby*
- ▶ **Nutte**

Evaluation

- ▶ Baseline 1: Unigram und Bigram SVM:
- ▶ Baseline 2: Feature Selection (Mutual Information) SVM

Evaluation

Table: Baseline: tf-idf unigram SVM

[illegible]

Evaluation

Table: Baseline: tf-idf bigram SVM

[illegible]

Evaluation

Table: Baseline: Feature-Selection m. Mutual Information

[illegible]

Aufgabenverteilung

Zeitplan

Programmarchitektur

Datenstrukturen

Literatur