# 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 Schimpwörtern
- 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 Deutschrapkorpus
- Deutschrap: 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

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

### Table: Baseline: tf-idf unigram SVM

Fold	#train	pos	neg	tok/tw	#test	pos	neg	tok/tw	F1
1	4500	1524	2976	146.94	500	160	340	144.124	0.7026952348344236
2	4500	1560	2940	146.73	500	124	376	146.05	0.7430423747345847
3	4500	1525	2975	146.73	500	159	341	146.0	0.6822048536263124
4	4500	1505	2995	146.57	500	179	321	147.512	0.6637390819218496
5	4500	1498	3002	145.94	500	186	314	153.154	0.6466493970138341
6	4500	1504	2996	147.0	500	180	320	143.568	0.6675527979162865
7	4500	1504	2996	146.61	500	180	320	147.106	0.6563892890626787
8	4500	1519	2981	146.16	500	165	335	151.13	0.6681861534976389
9	4500	1504	2996	147.08	500	180	320	142.86	0.650127611518916
10	4500	1513	2987	146.83	500	171	329	145.104	0.6737827241885149
	Total Accuracy: 0.68 (+/- 0.05)								

Table: Baseline: tf-idf bigram SVM

Fold	#train	pos	neg	tok/tw	#test	pos	neg	tok/tw	F1
1	4500	1524	2976	146.94	500	160	340	144.124	0.6713317445366677
2	4500	1560	2940	146.73	500	124	376	146.05	0.7143231005141465
3	4500	1525	2975	146.73	500	159	341	146.0	0.6607936275023581
4	4500	1505	2995	146.57	500	179	321	147.512	0.6407256780069586
5	4500	1498	3002	145.94	500	186	314	153.154	0.6004567364407795
6	4500	1504	2996	147.0	500	180	320	143.568	0.636733440679629
7	4500	1504	2996	146.61	500	180	320	147.106	0.6211176439430055
8	4500	1519	2981	146.16	500	165	335	151.13	0.6254687762688433
9	4500	1504	2996	147.08	500	180	320	142.86	0.6184350158030838
10	4500	1513	2987	146.83	500	171	329	145.104	0.6386443278943279
							To	tal Accura	acy: 0.64 (+/- 0.06)

#### Table: Baseline: Feature-Selection m. Mutual Information

Fold	#train	pos	neg	tok/tw	#test	pos	neg	tok/tw	F1
1	4500	1524	2976	146.94	500	160	340	144.124	0.7506333288282201
2	4500	1560	2940	146.73	500	124	376	146.05	0.8159589743589742
3	4500	1525	2975	146.73	500	159	341	146.0	0.7661802897341697
4	4500	1505	2995	146.57	500	179	321	147.512	0.761125609472951
5	4500	1498	3002	145.94	500	186	314	153.154	0.717136006424597
6	4500	1504	2996	147.0	500	180	320	143.568	0.7715594209711858
7	4500	1504	2996	146.61	500	180	320	147.106	0.7808015386464718
8	4500	1519	2981	146.16	500	165	335	151.13	0.7812137931034483
9	4500	1504	2996	147.08	500	180	320	142.86	0.7738412698412699
10	4500	1513	2987	146.83	500	171	329	145.104	0.7651150793650793
	Total Accuracy: 0.77 (+/- 0.05)								

# Aufgabenverteilung

# Zeitplan

# Programmarchitektur

### Datenstrukturen

### Literatur