

Emotionen in deutschen Texten

ein quantitativer Ansatz mit GATE

Elisabeth Weigl

Seminar (mit Bachelorarbeit)

Gliederung

- 1 Einführung
 - Text Mining
 - Emotionen
 - GATE
- 2 Umsetzung
 - Andere Arbeiten
 - GATE Konfiguration
 - Textauswahl
- 3 Ergebnisse

Gliederung

1 Einführung

Text Mining

Emotionen

GATE

2 Umsetzung

Andere Arbeiten

GATE Konfiguration

Textauswahl

3 Ergebnisse

Begriffe

- Data Mining (*Datenschürfung*) - verschiedene Daten
- Text Mining - textuelle Daten
 - statistische Verfahren (Zipfsche Gesetze, Häufigkeitsklasse, signifikante Kookurrenten, ...)
 - musterbasierte Verfahren (reguläre Ausdrücke, syntaktische Muster)
- Opinion Mining (*Meinungsschürfung*)

Gliederung

1 Einführung

Text Mining

Emotionen

GATE

2 Umsetzung

Andere Arbeiten

GATE Konfiguration

Textauswahl

3 Ergebnisse

Emotionen und Emotionsstrukturen in Sachtexten

- Bewertung (Reaktion auf Ereignisse)
- Selbstbetroffenheit, Ich-Beteiligung
 - 8 I-Variablen
- qualitative Analyse (Präsenz und Intensität von I-Variablen bzw. Bewertungen)
- quantitative Analyse (Berechnung der emotionalen Intensität)

affektive Wörterbücher

- **Dresdner Angst Wörterbuch (DAW)**
Gottschalk-Gleser-Verfahren, sechs Angstformen
- **Affektives Diktionär Ulm (ADU)**
Substantive & Adjektive, keine Verben, 26.000
Vollformwörter

acht bzw. zwölf Kategorien:

- 1 Liebe
- 2 Begeisterung
- 3 Zufriedenheit und Erleichterung
- 4 Freude und Stolz
- 5 Zorn
- 6 Furcht
- 7 Depressivität und Schuld
- 8 Ängstlichkeit und Scham

Gliederung

1 Einführung

Text Mining
Emotionen
GATE

2 Umsetzung

Andere Arbeiten
GATE Konfiguration
Textauswahl

3 Ergebnisse

Einführung

Text Mining
Emotionen

GATE

Umsetzung

Andere Arbeiten
GATE Konfiguration
Textauswahl

Ergebnisse

Zusammen-
fassung

- “General Architecture for Text Engineering”
- Java-Toolkit für Text Mining
- Annotationstool
- Language Resource
- Processing Resource
 - OntoGazetteer - Ontologiekonzepte
 - Jape-Transducer - Manipulieren von Annotationen

Gliederung

1 Einführung

- Text Mining
- Emotionen
- GATE

2 Umsetzung

- Andere Arbeiten**
- GATE Konfiguration
- Textauswahl

3 Ergebnisse

Andere Arbeiten

- Einteilung nach positiv/negativ & Suche nach Emotionen
- verwendeten Textkorpora (Schlagzeilen, Blogeinträge, Kundenrezensionen, Kindergeschichten)
- manuelle Annotation der Trainingsdaten

Gliederung

1 Einführung

Text Mining
Emotionen
GATE

2 Umsetzung

Andere Arbeiten
GATE Konfiguration
Textauswahl

3 Ergebnisse

Konfiguration

- 1 aduw.dex -> mehrere .lst Dateien
- 2 lists.def erstellen

```
Erleichterung .lst :emotion :erleichterung :german
```

- 3 jape-Dateien:
 - counter.jape (Wortanzahl)
 - emotion.jape (unterschiedliche Emotionsworte)
 - allemotions.jape (alle Emotionsworte)

emotion.jape

```
Phase: Emotion
Input: Split Lookup
Options: control = appelt

Rule: negationRule
Priority: 30
( {Lookup.majorType == verneinung} {Lookup.majorType == emotion} )
|
( {Lookup.majorType == emotion} {Lookup.majorType == verneinung} )
:negation -->
{}

//
// ANGST
//
Rule: angstRule
(
({Lookup.majorType == emotion, Lookup.minorType == aengstlich})

):emotionlabel
-->
:emotionlabel.Aengstlichkeit = {kind="aengstlich", rule = "angstRule"}
```

Gliederung

1 Einführung

Text Mining
Emotionen
GATE

2 Umsetzung

Andere Arbeiten
GATE Konfiguration
Textauswahl

3 Ergebnisse

Textauswahl

- 11 emotionale Texte (z.B. Vier-Pfoten Artikel, Artikel von Abtreibungsgegnern, etc.)
- 11 nicht-emotionale Texte (Wikipediaartikel, MSN-Datenschutzbestimmungen, etc.)

Einführung

Text Mining

Emotionen

GATE

Umsetzung

Andere Arbeiten

GATE Konfiguration

Textauswahl

Ergebnisse

Zusammen- fassung

Text	Gesamt	alle E in %	meiste E 12	meiste E 8
Vier-Pfoten 1	287	6,272	Depressivität	Ängstlichkeit
Vier-Pfoten 2	427	3,044	Begeisterung	Begeisterung
Jugend-FPÖ	915	2,404	Furcht	Freude
Grüne-Jugend	599	2,17	Depr.,Freude	Depr.,Freude
GNE	966	2,277	Ängstlichkeit	Ängstlichkeit
Yahoo Answer	132	5,30	Depr.,Freude	Depr.,Freude
Widerstand.info	368	2,446	Furcht	Furcht,Freude
Liebe	355	1,972	Liebe	Liebe
Das Kind	933	1,179	Depressivität	Depressivität
Microsoft	250	0,8	Liebe,Scham	Liebe,Ängstl.
Welt.de	787	2,541	Freude	Freude

nicht-emotionale Texte

Einführung

Text Mining

Emotionen

GATE

Umsetzung

Andere Arbeiten

GATE Konfiguration

Textauswahl

Ergebnisse

Zusammen- fassung

Blog-Bedienung	825	1,09	Begeisterung	Begeisterung
Linux	4682	1,153	Freude	Freude
Heise-Artikel	332	0,602	Liebe,Scham	Liebe,Ängstl.
Zeit-Artikel 1	612	1,633	4a,7a,8a	Freude
Donauau	457	1,97	Freude	Freude
Presse-Artikel	483	1,035	4a,4b,7a,8a	Freude
Zeit-Artikel 2	616	0,649	2,4a,6,7a	2,4,6,7
Schnapsherstellung	357	0,28	Freude	Freude
Zeit-Artikel 3	412	0,243	Ängstlichkeit	Ängstlichkeit
ORF	602	0,831	Freude,Furcht	Freude,Furcht
MSN	2788	0,681	Begeisterung	Begeisterung

(nicht-)erreichte Ziele

Einführung

Text Mining
Emotionen
GATE

Umsetzung

Andere Arbeiten
GATE Konfiguration
Textauswahl

Ergebnisse

Zusammen- fassung

- Einteilung nach Gefühlskategorien
 - zu wenig Annotationen
 - ADU für Psychoanalyse
- Einteilung in emotional/nicht-emotional
 - keine Normalverteilung
 - Tendenz
 - Wertebereich zw. 1,179 - 1,97

Probleme der Textanalyse

- gesprochene vs. geschriebene Sprache
- Tippfehler
- Sarkasmus, Zynismus
- Interpretation (“Freiheit”- Politik vs. Software)
- Unterstellung von Emotionalität

Einführung

Text Mining
Emotionen
GATE

Umsetzung

Andere Arbeiten
GATE Konfiguration
Textauswahl

Ergebnisse

Zusammen-
fassung

Fragen, Wünsche, Beschwerden, Anregungen, ... ?