

# Ansätze zur Erkennung von Kommunikationsmodi in Online-Diskussionen

Matthias Liebeck  
Heinrich-Heine-Universität Düsseldorf  
Institut für Informatik  
Universitätsstr. 1  
D-40225 Düsseldorf, Deutschland  
liebeck@cs.uni-duesseldorf.de

## Zusammenfassung

Bei der automatisierten Analyse von Textbeiträgen aus Online-Plattformen erfolgt oft eine Einteilung in positive und negative Aussagen. Bei der Analyse von Textbeiträgen eines kommunalen Online-Partizipationsverfahrens ist eine Aufteilung der geäußerten Meinungen in Kommunikationsmodi sinnvoll, um eine Filterung nach Argumenten und Emotionsäußerungen für nachfolgende Verarbeitungsschritte zu ermöglichen. In dieser Arbeit werden zwei Ansätze zur Erkennung von Kommunikationsmodi vorgestellt. Das erste Verfahren unterscheidet verschiedene Kommunikationsmodi anhand von Wortlisten. Die zweite Methode berücksichtigt Wortarten und extrahiert weitere sprachliche Eigenschaften. Zur Evaluation der Ansätze wird ein Datensatz aus Schlagzeilen von Nachrichtenartikeln der Internetseite ZEIT ONLINE und der Satire-Website Postillon erstellt. Die Ansätze werden zur Erkennung des Kommunikationsmodus Satire eingesetzt. Das beste Ergebnis mit einem durchschnittlichen  $F_1$  von 75,5 % wird durch den zweiten Ansatz mit einer Support Vector Machine erreicht.

## Kategorien

I.2.7 [Natural Language Processing]: Text Analysis;  
H.3.1 [Information stogare and retrieval]: Text Mining

## Schlüsselwörter

Natural Language Processing, Text Mining, Text Analysis, Sentiment Analysis, Opinion Mining, Emotion Recognition, Satire Detection, Postillon

## 1. EINLEITUNG

Im Internet gibt es viele verschiedene Plattformen, auf denen Meinungen, z. B. über Produkte, Filme oder politische Themen, als Textbeiträge geäußert werden können. Die Benutzer können untereinander Diskussionen führen, in denen sie idealerweise argumentativ ihre Meinungen darlegen. Die

se Textbeiträge können automatisiert analysiert werden, um Meinungsbilder über individuelle Themen zu erstellen.

### 1.1 Analyse von Online-Diskussionen

Von besonderem Interesse ist die automatisierte Analyse von Online-Partizipationsverfahren, bei denen Bürger die Möglichkeit nutzen, ihre Meinung zu lokalkommunalen Themen zu äußern. Bei einer oft erwünschten, hohen Teilnehmerzahl an Bürgern kann schnell das Problem auftreten, dass die beteiligten Bürger viele Textbeiträge erstellen und dadurch ein hoher Aufwand für eine manuelle Auswertung entsteht. Dieser nicht unerhebliche Arbeitsaufwand kann für Kommunen mit geringem Budget zu dem Problem führen, dass für die Analyse nicht genügend personelle Ressourcen zu Verfügung stehen und eine Analyse durch einen externen Dienstleister finanziell ebenfalls nicht möglich ist.

Eine weitere Schwierigkeit entsteht, wenn in einer Kommune erstmalig ein Online-Beteiligungsverfahren eingesetzt wird und die beteiligten Bürger vermehrt Inhalte äußern, die nicht zu dessen Thema passen. Werden in einem Verfahren beispielsweise gemeinsam Sparmaßnahmen diskutiert, so sind Beiträge, in denen Bürger kostenintensive Baumaßnahmen an der städtischen Infrastruktur vorschlagen, nicht konstruktiv und sollten herausgefiltert werden können.

Um diese Probleme zu reduzieren, sind mehrere automatisierte Schritte denkbar, die zu einer Arbeitsreduktion bei einer qualitativen Analyse führen. Durch diese Arbeitsreduktion kann eine Verwaltung umfassender mit den beteiligten Bürgern über eingereichte Verbesserungsvorschläge diskutieren. Zu diesen automatisierten Schritten gehören die thematische Gruppierung von Textbeiträgen und die themenspezifische Bestimmung einer Tonalität  $t \in \{\text{positiv, negativ, neutral}\}$ , um ein Stimmungsbild abschätzen zu können. Dadurch kann beispielsweise ermittelt werden, dass sich viele Bürger über eine Parkplatzsituation in einem Stadtteil beschweren und Anpflanzungen neuer Bäume in einem Park befürworten.

Ein üblicher Ansatz zur Bestimmung von Tonalitäten ist der Einsatz eines Tonalitätslexikons, in dem für einzelne Wörter jeweils ein numerischer Tonalitätswert angegeben ist. In [10] wurde gezeigt, dass für das deutsche Tonalitätslexikon SentiWS [14] nur eine geringe Abdeckung für die untersuchten Kommentare aus einem kleineren Online-Partizipationsverfahren und einem Nachrichtenportal erreicht wurde und daher weitergehende Ansätze zur Bestimmung von positiven und negativen Aussagen nötig sind. In dieser Publikation werden daher zwei Ansätze vorgestellt, die eine differenziertere Analyse von Meinungsäußerungen

ermöglichen sollen, indem genauer auf Kommunikationsmodi und geäußerte Emotionen eingegangen wird.

## 1.2 Kommunikationsmodi

In Online-Diskussionen verwenden die Teilnehmer verschiedene Kommunikationsmodi. Die einzelnen Beteiligten können beispielsweise Aussagen tätigen (1), Argumente für oder gegen einen Standpunkt formulieren (2) oder Emotionsäußerungen zum Ausdruck bringen (3).

- (1) *Ich bin für den Bau eines Schwimmbads.*
- (2) *Es sollte kein Geld für die Oper ausgegeben werden, da unsere Schulen das Geld dringender benötigen.*
- (3) *Die steigende Kriminalitätsrate macht mir Angst.*

Die Emotionsäußerungen können wiederum in verschiedene Emotionen differenziert werden. Bei der Untersuchung von Textbeiträgen aus Online-Partizipationsverfahren sind zunächst die Emotionen  $E := \{\text{Freude, Hoffnung, Empörung, Enttäuschung, Angst}\}$  für ein Stimmungsbild der Bürgermeinungen interessant. Eine fundiertere Einteilung wird in zukünftigen Arbeiten durch Experten erfolgen. Die folgenden Beispiele aus einem fiktiven Online-Partizipationsverfahren veranschaulichen die unterschiedenen Emotionen.

- (4) Freude: *Das wäre wirklich schön.*
- (5) Hoffnung: *Ich hoffe, dass an der Hauptstraße neue Bäume gepflanzt werden können.*
- (6) Empörung: *Das gehört doch verboten!*
- (7) Enttäuschung: *Die zur letzten Wahlperiode versprochene Änderung konnte meine Erwartungen nicht erfüllen.*
- (8) Angst: *Ich befürchte, dass meine Buslinie durch diese Änderung eingestellt wird.*

Bei der automatisierten Erkennung von Emotionen variiert die Einteilung der Emotionen je nach Textmaterial. [17] unterscheidet in die sechs Emotionen *anger, disgust, fear, joy, sadness* und *surprise*, wohingegen [8] nur die vier Emotionen *anger, fear, joy* und *sadness* betrachtet.

Der Grund für die Untersuchung von Kommunikationsmodi ist die Arbeitshypothese, dass durch eine automatische Erkennung der Emotionen  $E$  ein detaillierteres, themenspezifisches Stimmungsbild angegeben werden kann, als es ein Mittelwert über numerische Tonalitätsangaben ermöglicht. Dazu muss ein Klassifikator  $K$  erstellt werden, der jedem Satz eines Textbeitrags individuell eine Emotion oder die Klasse *neutral* zuordnet. Erschwerend bei der Klassifikation ist die subjektive und kontextabhängige Wahrnehmung von Emotionen.

Der Rest dieser Arbeit ist wie folgt aufgebaut: Im nächsten Kapitel werden verwandte Arbeiten vorgestellt. Nachdem in Kapitel 3 zwei Ansätze zur Erkennung von Kommunikationsmodi präsentiert werden, erfolgt in Kapitel 4 eine Evaluation beider Ansätze am Beispiel des Kommunikationsmodus Satire. Anschließend wird in Kapitel 5 ein Fazit gezogen und Ideen für zukünftige Arbeiten angegeben.

## 2. VERWANDTE ARBEITEN

Der Bereich Sentiment Analysis beschäftigt sich mit der automatisierten Bestimmung von Tonalitäten in Textdokumenten. Übliche Anwendungsgebiete sind die Analyse von Produktrezensionen [7] und Filmrezensionen [12], die in positive und negative Äußerungen kategorisiert werden. Die automatisierte Extraktion von Tonalitäten, bei der einer Aussage eine Tonalität  $t \in \{\text{positiv, negativ, neutral}\}$  zuge-

ordnet wird, hat sich für die Analyse von Zeitungsartikeln im Rahmen einer Medienresonanzanalyse [16] bewährt.

Die Erkennung von Emotionen in Texten ist bereits mehrfach [1, 8, 17] untersucht worden. Das Ziel von [1] ist die Erkennung von Emotionen in Märchentexten. Die Autoren fokussieren sich dabei auf die binäre Klassifikation von englischsprachigen Sätzen und untersuchen, ob in einem Satz Emotionen auftreten. Dafür annotieren sie einen Datensatz von 1580 Sätzen aus 22 Geschichten der Gebrüder Grimm, auf dem ein linearer Klassifikator trainiert und evaluiert wird. Zur vektoriellen Repräsentation der einzelnen Sätze verwendet [1] mehrere sprachliche Eigenschaften wie die Verteilung von POS-Tags, Satzlängen, Interpunktionszeichen und mehrere Listen von Wörtern, die auf Emotionen hindeuten.

Eine automatisierte Erkennung von sechs verschiedenen Emotionen erfolgt in [17]. Die Autoren untersuchen mehrere Techniken auf einem Datensatz aus 1000 Schlagzeilen von Nachrichtenartikeln. Für ein Baseline-Verfahren erstellen die Autoren sechs Wortlisten aus *WordNet-Affect* [18], einer um Emotionen annotierte Erweiterung von *WordNet* [11]. In dem Baseline-Verfahren erfolgt die Klassifikation eines Satzes  $s$  zu Emotionen durch das Auftreten der einzelnen Wörter aus  $s$  in den Wortlisten. In einem fortgeschrittenen Verfahren, das aus einer Kombination aus *Latent Semantic Analysis* [9] und Synonymen aus *WordNet* und *WordNet-Affect* besteht, erreicht [17] einen durchschnittlichen  $F_1$  Wert von 17,57 % als bestes Ergebnis für die Erkennung der sechs Emotionen.

In [3] wird ebenfalls das binäre Klassifikationsproblem der Satire-Erkennung in Zeitungsartikeln behandelt. Dabei berücksichtigen die Autoren zusätzlich den Inhalt der Nachrichtenartikel. Als Datensatz untersucht [3] dabei insgesamt 4000 englischsprachige Zeitungsartikel, von denen 233 Satire beinhalten. Als Basismodell wählen die Autoren ein Bag-of-Words-Modell mit binärer Gewichtung. Eine deutliche Steigerung der Ergebnisse kann mit *Bi-Normal Separation (BNS)* [6] als Gewichtung, dem Nachschlagen von Wörtern in einem Lexikon und einer Google-Suche nach auftretenden Personen und Organisationsnamen erreicht werden. Zur Klassifikation setzt [3] eine lineare Support Vector Machine ein.

## 3. ANSÄTZE

Im Folgenden werden zwei Ansätze vorgestellt, die für die Erkennung von Kommunikationsmodi in Textbeiträgen aus Online-Partizipationsverfahren naheliegend sind. Beide Ansätze versuchen, die menschliche Erkennung von Emotionen nachzuahmen, indem sie auf die in den Textbeiträgen vorhandenen Wörter, in Form von Signalwörtern und bestimmten sprachlichen Konstruktionen, achten.

Dazu müssen die zu untersuchenden Texte satzweise analysiert werden. Die Eingabetexte werden zunächst durch eine Natural Language Processing Pipeline aufbereitet. Die Anzahl an Verarbeitungsschritten der Pipeline ist von der konkreten Aufgabenstellung abhängig. Für die vorgestellten Ansätze werden insgesamt vier Schritte in der NLP-Pipeline durchgeführt: Mittels eines Tokenizers wird ein Eingabetext in einzelne Wörter zerlegt. Durch einen Sentence Splitter werden die Wörter in Sätze gruppiert. Anschließend werden für jedes Wort ein Part-of-Speech Tag (POS-Tag) bzw. eine Wortart bestimmt und eine Lemmatisierung durchgeführt, durch die für jedes Wort zusätzlich eine Grundform (z. B.

Schwimmbäder → Schwimmbad) angegeben wird.

Der erste Ansatz untersucht, inwiefern bestimmte Schlüsselwörter auf einzelne Kommunikationsmodi oder Emotionen hinweisen. Der zweite Ansatz arbeitet unter der Hypothese, dass eine Korrelation bestimmter sprachlicher Eigenschaften zu einzelnen Kommunikationsmodi auftritt. Dabei werden die verwendeten Wortarten untersucht. Motiviert wird dieser Ansatz dadurch, dass eine positive Emotion beispielsweise mit einer überdurchschnittlichen Anzahl an Adjektiven korrelieren könnte.

In beiden Ansätzen wird jedem Satz mittels eines Klassifikators ein Kommunikationsmodus bzw. eine Emotion oder die Klasse *neutral* zugeordnet. Um beide Ansätze evaluieren zu können, muss ein annotierter Datensatz bzw. ein Korpus als Trainingsmenge verwendet werden, in dem auf Satzebene alle Sätze mit entsprechenden Kommunikationsmodi annotiert sind. Der Korpus wird in eine Trainings- und eine Testmenge aufgeteilt, anhand derer ein Klassifikator trainiert bzw. bewertet werden kann.

### 3.1 Wortlisten

Der erste Ansatz basiert auf der Annahme, dass das Auftreten bestimmter Wörter mit einem Kommunikationsmodus bzw. einer Emotion korreliert. Das Ziel des Ansatzes ist die Verwendung von Wortlisten, die eine Klassifikation eines Satzes, basierend auf den in ihm enthaltenen Wörtern, ermöglichen. Für den nachfolgenden Satz können die Wörter *Angst* und *verliere* auf die Emotion *Angst* hinweisen:

(9) *Ich habe Angst, dass ich verliere.*

Um diese Erkennung zu automatisieren, können für jede Emotion charakteristische Wörter aus einer Trainingsmenge extrahiert werden. Dazu werden in der Trainingsmenge auftretende Wörter untersucht und in disjunkte Wortlisten eingeteilt. Bei der Konstruktion dieser Wortlisten könnten für Beispiel (9) die Wörter *Angst* und *verliere* als charakteristisch identifiziert werden, falls sie auch in anderen Sätzen der Trainingsmenge auftreten, die ebenfalls mit der Emotion *Angst* annotiert sind. Die größte Schwierigkeit bei diesem Ansatz ist die passende Auswahl der Wörter für die Wortlisten. So muss darauf geachtet werden, keine Wörter zu verwenden, die in allen Klassen häufig vorkommen. Daher bietet es sich an, nur Wörter zu betrachten, die überwiegend in einer Klasse (relative Häufigkeit größer als ein Schwellwert  $\tau$ ) und damit nur selten in anderen Klassen vorkommen.

Wird allein auf den Schwellwert  $\tau$  geachtet, so entsteht das Problem, dass auch Wörter in die Wortlisten aufgenommen werden, die insgesamt nur selten auftreten. Tritt beispielsweise das Wort *Glück* nur einmal in der Trainingsmenge auf, und zwar in einem mit der Emotion *Empörung* annotierten Satz, so würde das Wort *Glück* in die Wortliste für die Emotion *Empörung* aufgenommen werden, anstatt in die Wortliste der Emotion *Freude*. Um diese Problematik zu vermeiden, bietet sich ein Parameter *supp* an, der angibt, wie häufig ein Wort insgesamt in der Trainingsmenge auftreten muss, bevor es in eine Wortliste eingefügt werden darf.

Durch diese beiden Parameter werden häufig in allen Emotionen auftretende Wörter, wie Artikel und Pronomen, herausgefiltert. Die konkrete Wahl der beiden Parameter kann durch ein Experiment geschätzt werden.

Um einen neuen Satz einer Emotion zuzuordnen, kann jedes Wort  $w$  des Satzes in den Wortlisten nachgeschlagen werden. Dem Satz wird diejenige Emotion zugeordnet, für die am meisten Wörter in der jeweiligen Wortliste gefunden

werden.

In einer Modifikation dieses Ansatzes werden nicht die Wörter, sondern deren, durch eine Lemmatisierung bestimmte, Grundformen in Wortlisten geführt bzw. nachgeschlagen.

### 3.2 Sprachliche Eigenschaften

Der zweite Ansatz basiert auf der Hypothese, dass die verschiedenen Kommunikationsmodi bzw. Emotionen im Text charakteristische sprachliche Merkmale besitzen. Um diese Vermutung auf einem Datensatz zu überprüfen, ist eine Methode  $M$  notwendig, die sprachliche Eigenschaften eines Satzes in eine vektorielle Darstellung überführt. Dazu werden aus allen Sätzen einer Trainingsmenge sprachliche Eigenschaften durch  $M$  extrahiert. Für einen zu klassifizierenden Satz einer Testmenge werden ebenfalls sprachliche Eigenschaften mittels  $M$  extrahiert, die anschließend mit einem Klassifikationsverfahren und einer Distanzfunktion zu einem Kommunikationsmodus zugeordnet werden können.

Eine einfache Annahme ist, dass bestimmte Kommunikationsmodi bzw. Emotionen mit einer bestimmten Verteilung von POS-Tags korrelieren. Als erste vektorielle Modellierung eines Satzes  $s$  wird daher für jeden POS-Tag  $p$  eines Tagsets die Häufigkeit von  $p$  in  $s$  angegeben.

Diese vektorielle Darstellung kann um weitere sprachliche Eigenschaften ergänzt werden, die eventuell charakteristisch für eine Emotion sein können, beispielsweise welche Wortart am Satzanfang und am Satzende steht oder welches Interpunktionszeichen (Punkt, Fragezeichen oder Ausrufezeichen) einen Satz beendet. Ferner kann das Auftreten von Negationen oder von verschachtelten Nebensätzen berücksichtigt werden.

## 4. KOMMUNIKATIONSMODUS SATIRE

Da zum aktuellen Zeitpunkt noch keine ausreichende Datenmenge an Diskussionsbeiträgen aus Online-Partizipationsverfahren vorliegt, werden die in Kapitel 3 beschriebenen Ansätze zur Erkennung von Kommunikationsmodi konkret auf den Kommunikationsmodus Satire angewendet, indem die Erkennung von Satire in Nachrichtenartikeln evaluiert wird. In zukünftigen Arbeiten werden diese Techniken ebenfalls für die Erkennung von Emotionsäußerungen in Online-Partizipationsverfahren erprobt und evaluiert.

Zur Evaluation wird ein Datensatz aus Nachrichtenartikeln zusammengestellt. Basierend auf den Überschriften der Nachrichtenartikel soll das binäre Klassifikationsproblem gelöst werden, ob ein Nachrichtenartikel von der Satire-Webseite *Postillon*<sup>1</sup> stammt oder auf ZEIT ONLINE<sup>2</sup> veröffentlicht wurde.

### 4.1 Datensatz

Der zur Satire-Erkennung verwendete Datensatz setzt sich aus den beiden Nachrichtenquellen *Postillon* und ZEIT ONLINE zusammen. Die Schlagzeilen der Artikel beider Webseiten wurden jeweils über einen JSON-Webservice heruntergeladen. Für den *Postillon* werden 3650 Artikel aus dem Zeitraum Oktober 2008 bis März 2015 betrachtet. Die beiden nachfolgenden Schlagzeilen sind Beispiele für Satire-Artikel aus dem *Postillon*:

<sup>1</sup><http://www.der-postillon.com>

<sup>2</sup><http://www.zeit.de>

- (10) *Sensation! Autobahn-Fahrer entdeckt weitere Fahrbahn rechts neben der Mittelspur*  
 (11) *Kölner Dom von Unbekannten über Nacht um 360 Grad gedreht*

Eine genauere Betrachtung der Satire-Artikel hat ergeben, dass bestimmte Artikelformate in regelmäßigen Abständen vorkommen, wie z. B. *Sonntagsfragen* oder *Newsticker*. Diese wurden für die weitere Betrachtung entfernt, um das Klassifikationsproblem zu erschweren. Da einige Artikel mehrfach veröffentlicht wurden, wird von allen Artikeln mit demselben Namen jeweils nur die chronologisch erste Veröffentlichung verwendet. Durch diese Filterungsschritte reduziert sich die Anzahl der zur Verfügung stehenden Postillon-Artikel auf 2260.

Die zweite Klasse des Datensatzes setzt sich aus klassischen Zeitungsartikeln zusammen, die auf ZEIT ONLINE veröffentlicht wurden. Aus den Kategorien Wirtschaft, Gesellschaft, Sport, Wissen und Digital wurden jeweils die 2000 aktuellsten Artikel vor dem Stichtag 1.4.2015 mittels *ZeitOnlineAPISharp*<sup>3</sup> heruntergeladen. Für die weitere Verarbeitung wurden aus diesen 10000 Artikeln insgesamt 2260 Artikel zufällig ausgewählt, um einen balancierten Datensatz betrachten zu können. Die beiden folgenden Schlagzeilen stammen aus Artikeln von ZEIT ONLINE:

- (12) *Lehrerverband warnt vor Risiken für Bildung*  
 (13) *Energiekonzern verzichtet auf Atomenergie, Kohle und Gas*

Zur nachfolgenden Evaluation werden die Artikel in eine Trainings- und eine Testmenge aufgeteilt. Zum Training werden pro Klasse 1000 Artikel zufällig ausgewählt, sodass das Training auf insgesamt 2000 Artikeln stattfindet. Die Evaluation erfolgt auf der Grundlage der verbleibenden 2520 Artikel.

## 4.2 Evaluation

Der zusammengestellte Datensatz wird zunächst durch eine NLP-Pipeline aufbereitet: Für jede Schlagzeile erfolgt durch OpenNLP<sup>4</sup> eine Zerlegung in einzelne Wörter, eine Trennung in Sätze und eine Bestimmung von Wortarten für jedes einzelne Wort. Die ermittelten Wortarten stammen aus dem *Stuttgart-Tübingen-Tagset (STTS)* [15], welches aus 54 verschiedenen Wortarten besteht. Die Lemmatisierung von Wörtern erfolgt durch Mate Tools [2].

### 4.2.1 Wortlisten

Für die Satire-Erkennung über Wortlisten müssen zwei disjunkte Wortlisten  $W_P$  und  $W_Z$  erstellt werden, in denen jeweils Wörter vertreten sind, die überwiegend nur in den Schlagzeilen des Postillons bzw. in den Schlagzeilen von ZEIT ONLINE auftreten. Zur Erstellung dieser Wortlisten werden zunächst die einzelnen Wörter als Datengrundlage verwendet. Im Postillon treten 4025 (3583 Lemmata) und in ZEIT ONLINE 3050 (2696 Lemmata) verschiedene Wörter auf. Für die Erstellung der Wortlisten wird zunächst pro Wort bestimmt, wie häufig es in der Trainingsmenge auftritt. Wörter, die weniger als  $\text{supp}$  mal auftreten, werden ignoriert. Anschließend wird für jedes Wort  $w$  berechnet, wie groß die relative Häufigkeit  $p$  von  $w$  in den Schlagzeilen des Postillons bzw. in den Schlagzeilen aus ZEIT ONLINE ist. Falls  $p \geq \tau$  ist, so wird  $w$  in die entsprechende Wortliste eingefügt. In einer zweiten Variante werden anstelle der

<sup>3</sup><https://github.com/Liebeck/ZeitOnlineAPISharp>

<sup>4</sup><https://opennlp.apache.org/>

**Tabelle 1: Satire-Erkennung mit Wortlisten**

$\tau$	Wort		Lemma	
	+ Default		+ Default	
0,55	59,24 %	56,84 %	62,88 %	60,21 %
0,6	<b>61,51 %</b>	<b>57,74 %</b>	<b>64,21 %</b>	<b>60,38 %</b>
0,65	59,23 %	55,73 %	57,59 %	54,03 %
0,7	58,87 %	52,11 %	53,75 %	47,70 %
0,75	52,36 %	46,11 %	50,90 %	45,24 %
0,8	49,02 %	41,52 %	48,13 %	40,97 %
0,85	43,41 %	33,58 %	43,81 %	33,58 %
0,9	42,42 %	30,76 %	42,96 %	30,06 %
0,95	41,39 %	28,77 %	42,04 %	29,30 %
1	41,39 %	28,77 %	42,04 %	29,30 %

Wörter die durch die Lemmatisierung bestimmten Lemmata der Wörter untersucht.

Da bisher noch keine Erfahrungswerte für die Parameterwahl vorliegen, werden die Auswirkungen verschiedener Parameter experimentell bestimmt, indem eine Gittersuche mit  $\tau \in \{0.55, 0.6, \dots, 1\}$  und  $\text{supp} \in \{3, 4, \dots, 10\}$  durchgeführt wird.

Die Ergebnisse der Satire-Erkennung durch Wortlisten sind in Tabelle 1 als durchschnittlicher  $F_1$  Wert über beide Klassen angegeben, wobei für jeden Wert von  $\tau$  das jeweils beste Ergebnis angegeben ist, welches durchgehend mit  $\text{supp} = 3$  erreicht wird. Bei dem Mehrheitsentscheid des Ansatzes ist es möglich, dass ein Unentschieden vorliegt. Dies bedeutet, dass keines der Wörter einer zu klassifizierenden Schlagzeile in  $W_P$  oder  $W_Z$  auftritt oder dass ein Gleichstand vorliegt. Für jede Parameterbelegung von  $\tau$  und  $\text{supp}$  werden zwei Evaluationen mit einem unterschiedlichen Standardwert  $d \in \{\text{Postillon}, \text{ZEIT ONLINE}\}$  zur Auflösung eines Gleichstands durchgeführt, von denen in Tabelle 1 jeweils das schlechtere der beiden Ergebnisse aufgeführt ist. Um die Auswirkungen eines Standardwerts beurteilen zu können, sind in Tabelle 1 ebenfalls die Ergebnisse einer Klassifikation aufgeführt, bei der ein Unentschieden bei einem Mehrheitsentscheid als falsche Klassifikation behandelt wird.

Für die Satire-Erkennung erreicht der erste Ansatz mit den auftretenden Wörtern als Datengrundlage als bestes Ergebnis den Wert 61,51 % für  $\tau = 0,6$  und  $\text{supp} = 3$ . Durch eine Lemmatisierung kann das Ergebnis auf 64,21 % gesteigert werden. Bei einem fixierten Wert für  $\tau$  und einem steigenden Wert für  $\text{supp}$  werden die Klassifikationsergebnisse schlechter, da die Größen der Wortlisten entsprechend abnehmen. Für den untersuchten Datensatz verschlechtern sich die Ergebnisse bei einem steigenden  $\tau$  aus demselben Grund.

### 4.2.2 Sprachliche Features

In einem ersten vektoriiellen Modell wird jede Schlagzeile durch die absoluten Häufigkeiten der auftretenden POS-Tags aus dem STTS-Tagset repräsentiert. Zum Vergleich wird ein zweites Modell untersucht, in dem die Auswirkungen einer Reduktion der 54 POS-Tags auf die 12 POS-Tags des UTS-Tagsets [13] beobachtet werden. Beide Modelle werden mit den drei Klassifikationsverfahren k-Nearest Neighbors (kNN), Support Vector Machine (SVM) und out-of-place measure [4] evaluiert.

Als SVM-Implementierung wird LIBSVM [5] verwendet. Eingesetzt wird eine soft-margin SVM mit einem RBF-Kernel  $K(x, y) = \exp(-\gamma \|x - y\|^2)$ . Die für das Training der

**Tabelle 2: Satire-Erkennung mit sprachlichen Eigenschaften**

Modell	SVM	kNN		Out-of-place
		+ Default		
STTS	73,65 %	71,77 %	71,04 %	<b>67,13 %</b>
STTS, Variante B	<b>75,50 %</b>	73,13 %	72,94 %	—
UTS	71,51 %	74,75 %	74,48 %	60,74 %
UTS, Variante B	73,55 %	<b>75,25 %</b>	<b>75,05 %</b>	—

SVM benötigten Werte für den Strafterm  $C$  und für  $\gamma$  werden pro Modell jeweils über eine Gittersuche mittels einer Kreuzvalidierung über die Trainingsmenge bestimmt. Als Konvergenzkriterium der SVM wird  $\epsilon = 10^{-3}$  gesetzt.

Für den kNN-Algorithmus werden verschiedene Werte für  $k \in \{1, \dots, 12\}$  erprobt. Bei der Bestimmung der nächsten Nachbarn werden jeweils die  $k$  nächsten Nachbarn per euklidischer Distanz ermittelt. Sollten mehrere Kandidaten für die Auswahl des  $k$ -nächsten Nachbarn  $p$  vorhanden sein, so wird die Liste der nächsten Nachbarn um alle Nachbarn erweitert, die zum Anfrageobjekt  $o$  denselben Abstand haben wie  $o$  zu  $p$ . Die Klassenzugehörigkeit erfolgt über einen Mehrheitsentscheid der Klassen aller gefundenen nächsten Nachbarn. Tritt dabei ein Gleichstand auf, so wird ebenfalls ein Standardwert verwendet.

Das out-of-place measure wird gewöhnlich für die Spracherkennung durch N-Gramme eingesetzt. Für die Evaluation wird es für den Vergleich von Verteilungen von POS-Tags verwendet, indem für beide Nachrichtenquellen sogenannte Kategorienprofile bestimmt werden. Ein Kategorienprofil besteht dabei jeweils aus einer nach absoluten Häufigkeiten absteigend sortierten Liste von POS-Tags der jeweiligen Trainingsmenge. Zur Klassifikation einer Schlagzeile wird ein Anfrageprofil mittels derselben Methode berechnet. Die Indexpositionen des Anfrageprofils werden mit den Indexpositionen der Kategorienprofile verglichen. Einer Schlagzeile wird dann diejenige Nachrichtenquelle zugeordnet, zu deren Kategorienprofil der kleinste Abstand zum Anfrageprofil besteht.

Die Modellierung wird in einer Variante B ergänzt, in der weitere sprachliche Eigenschaften als binäre Dimension hinzugefügt werden. Dabei wird berücksichtigt, ob in der Überschrift das erste Wort ein Nomen ist, ob das erste Wort ein Verb ist, ob das letzte Wort ein Verb ist, ob in der Überschrift ein Anführungszeichen vorhanden ist und ob ein Fragezeichen, ein Ausrufezeichen oder ein Komma (jeweils binär) auftritt.

Die Ergebnisse der drei Klassifikationsverfahren sind in Tabelle 2 mit durchschnittlichen  $F_1$  Werten über beide Klassen dargestellt. Das insgesamt beste Ergebnis von 75,5 % erreicht eine SVM mit STTS POS-Tags und Variante B. Das beste Ergebnis des kNN-Algorithmus ist minimal schlechter mit 75,25 %. Die Klassenzugehörigkeit beim kNN-Algorithmus kann in fast allen Fällen per Mehrheitsentscheid bestimmt werden. Interessant zu beobachten sind die Auswirkungen der Reduktion des STTS-Tagsets auf das UTS-Tagset, die je nach Klassifikationsverfahren unterschiedlich sind. Für den kNN-Algorithmus konnte eine Verbesserung der Ergebnisse erzielt werden. Bei den anderen Verfahren verschlechtert sich das Ergebnis im Vergleich zu den STTS-Tags.

## 5. FAZIT UND AUSBLICK

In dieser Arbeit wurden zwei Ansätze zur Erkennung von Kommunikationsmodi präsentiert. Für die nahe Zukunft ist geplant, beide Ansätze auf Textbeiträge eines Online-Partizipationsverfahrens anzuwenden. Dazu wird ein Codebuch entwickelt werden, mit dem der Datensatz in Bezug auf Kommunikationsmodi annotiert wird. Beide Ansätze werden dann auf das Multiklassenproblem der Erkennung von Emotionsäußerungen transferiert und evaluiert.

Bei der Untersuchung des Kommunikationsmodus Satire wurde gezeigt, dass die beiden Ansätze gute Ergebnisse von bereits 75,5 % erreicht haben. Bei der Evaluation ist aufgefallen, dass das Festlegen eines Standardwerts bei einem Gleichstand für den mit Wortlisten arbeitenden Ansatz eine größere Auswirkung auf die Klassifikationsergebnisse hat, als beim kNN-Algorithmus für die sprachlichen Eigenschaften. Für das binäre Klassifikationsproblem der Satire-Erkennung sind ein Reihe von weitergehenden Untersuchungen möglich. Es könnte untersucht werden, welche Auswirkungen durch die Filterung von Stoppwörtern, durch den Vergleich unterschiedlicher Distanzfunktionen für den kNN-Algorithmus und durch die Verwendung weiterer sprachlicher Eigenschaften entstehen können und ob dadurch die Ergebnisse gegenüber Variante B noch gesteigert werden können.

Die größte Schwierigkeit des ersten Ansatzes ist die Auswahl charakteristischer Wörter. Bei einem Transfer des Ansatzes auf andere Datenquellen werden die Auswirkungen von manuell vorgegebenen Wortlisten untersucht. Insbesondere ist geplant, die Auswirkungen einer Erweiterung der Listen durch Synonyme aus einem Thesaurus zu beobachten. Bei der Generierung der Wortlisten konnten die Ergebnisse durch eine Lemmatisierung verbessert werden. In zukünftigen Arbeiten wird untersucht werden, inwiefern Wiktionary<sup>5</sup> zur Grundformreduktion eingesetzt werden kann. Außerdem wird der Ansatz für den Umgang mit Negationen erweitert werden.

Für den zweiten Ansatz sind weitere sprachliche Eigenschaften in einer vektoriiellen Darstellung vorstellbar, wie beispielsweise das Auftreten von POS-Tag-Bigrammen. Um bei einer Klassifikation gute Ergebnisse erzielen zu können, ist eine Filterung nach sprachlichen Eigenschaften notwendig, die besonders gut mit den einzelnen Klassen korrelieren. Darüber hinaus wird in zukünftigen Arbeiten untersucht werden, inwiefern beide Ansätze kombinierbar sind. Anstelle eines Mehrheitsentscheids des ersten Ansatzes kann die vektorielle Darstellung für jeden Eintrag einer Wortliste um binäre Dimensionen erweitert werden, die jeweils angeben, ob in dem zu klassifizierenden Satz das entsprechende Wort einer Wortliste auftritt. In nachfolgenden Arbeiten wird untersucht werden, welchen Einfluss eine Verkleinerung des STTS-Tagsets auf das UTS-Tagset auf anderen Datensätzen hat.

<sup>5</sup><https://de.wiktionary.org/>

## 6. LITERATUR

- [1] C. O. Alm, D. Roth, and R. Sproat. Emotions from Text: Machine Learning for Text-based Emotion Prediction. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, HLT '05*, pages 579–586. Association for Computational Linguistics, 2005.
- [2] A. Björkelund, B. Bohnet, L. Hafdell, and P. Nugues. A High-Performance Syntactic and Semantic Dependency Parser. In *Proceedings of the 23rd International Conference on Computational Linguistics: Demonstrations, COLING '10*, pages 33–36. Association for Computational Linguistics, 2010.
- [3] C. Burfoot and T. Baldwin. Automatic Satire Detection: Are You Having a Laugh? In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, ACLShort '09*, pages 161–164. Association for Computational Linguistics, 2009.
- [4] W. B. Cavnar and J. M. Trenkle. N-Gram-Based Text Categorization. In *Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval*, pages 161–175, 1994.
- [5] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011.
- [6] G. Forman. BNS Feature Scaling: An Improved Representation over TF-IDF for SVM Text Classification. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM '08*, pages 263–270. ACM, 2008.
- [7] M. Hu and B. Liu. Mining and Summarizing Customer Reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04*, pages 168–177. ACM, 2004.
- [8] S. M. Kim, A. Valitutti, and R. A. Calvo. Evaluation of Unsupervised Emotion Models to Textual Affect Recognition. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, CAAGET '10*, pages 62–70. Association for Computational Linguistics, 2010.
- [9] T. Landauer, P. Foltz, and D. Laham. An introduction to latent semantic analysis. *Discourse processes*, 25:259–284, 1998.
- [10] M. Liebeck. Aspekte einer automatischen Meinungsbildungsanalyse von Online-Diskussionen. In *Proceedings BTW 2015 - Workshops und Studierendenprogramm*, pages 203–212, 2015.
- [11] G. A. Miller. WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11):39–41, 1995.
- [12] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs Up?: Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP '02*, pages 79–86. Association for Computational Linguistics, 2002.
- [13] S. Petrov, D. Das, and R. McDonald. A Universal Part-of-Speech Tagset. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. European Language Resources Association, 2012.
- [14] R. Remus, U. Quasthoff, and G. Heyer. SentiWS – a Publicly Available German-language Resource for Sentiment Analysis. In *Proceedings of the 7th International Language Resources and Evaluation (LREC'10)*, pages 1168–1171, 2010.
- [15] A. Schiller, S. Teufel, C. Stöckert, and C. Thielen. Guidelines für das Tagging deutscher Textcorpora mit STTS (kleines und großes Tagset). Technical report, Universität Stuttgart, Universität Tübingen, 1999.
- [16] T. Scholz and S. Conrad. Opinion Mining in Newspaper Articles by Entropy-Based Word Connections. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1828–1839. Association for Computational Linguistics, 2013.
- [17] C. Strapparava and R. Mihalcea. Learning to Identify Emotions in Text. In *Proceedings of the 2008 ACM Symposium on Applied Computing, SAC '08*, pages 1556–1560. ACM, 2008.
- [18] C. Strapparava and A. Valitutti. WordNet-Affect: An affective extension of WordNet. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, pages 1083–1086. European Language Resources Association, 2004.