

# SentiStrength

Irina Schmidt

Universität Koblenz-Landau  
Campus Koblenz  
Institute for Web Science and Technologies

## 1 Motivation

Jeden Tag wird in sozialen Netzwerken, Blogs und Diskussions-Foren eine hohe Anzahl an Nachrichten gepostet. Darin werden verstärkt Emotionen zum Ausdruck gebracht, um Freundschaften zu schließen oder aufrecht zu erhalten, um soziale Unterstützung zu zeigen oder um in Online-Debatten Stellung zu einem Thema zu beziehen. Besonders für die Forschung im Bereich der Sozialwissenschaften und Psychologie im Allgemeinen sind Analysen von Sprache und darin enthaltener Emotion ergiebig und wertvoll. Dazu zählen etwa Analysen zur Verknüpfung von Sprache und Psychologie [22], aber auch das Erkennen verschiedener empfindungsbezogener Muster. Beispielsweise lässt sich anhand der genutzten Sprache untersuchen, wie gut Menschen mit schmerzlichen Verlusten umgehen [11]. Weiterhin sind damit umfassende Analysen über geschlechterspezifische Unterschiede im Ausdruck von Empfindungen oder über grundsätzlich erfolgreiche Kommunikationsstrategien [9] interessant. Damit gewinnen Methoden und Anwendungen zur Analyse und Auswertung von Stimmungen und Meinungen in derartigen Texten verstärkt an Bedeutung [14]. So wurde beispielsweise mit Hilfe solcher Stimmungs- und Meinungsauswertungen in sozialen Medien bereits die Erkenntnisse gewonnen, dass negative Empfindungsdarstellungen in Online-Debatten im Schnitt zu längeren Diskussionen führen, oder, dass Frauen durchschnittlich mehr positive Empfindungen ausdrücken und ihnen entgegengebracht werden als bei Männern [20].

Es gibt bereits verschiedene Ansätze zur automatischen Analyse von Stimmungen und Meinungen in Texten allgemein. Sie fallen unter die Kategorie der Sentimentanalyse, die in Kapitel 2 allgemein beschrieben wird. Kapitel 3 durchleuchtet die menschliche Einschätzung des Sentiments, die die Grundlage zur Bewertung einer Sentimentanalyse darstellt und bedarf somit der gesonderten Betrachtung. In Kapitel 4 wird ein Ansatz zur Sentimentanalyse vorgestellt, der als Beispiel einer ganzen Gruppe von Ansätzen, des sogenannten Machine-Learning, dienen soll. Kapitel 5 beschreibt die speziellen Anforderungen an eine Sentimentanalyse in sozialen Medien und Kapitel 6 stellt daraufhin den SentiStrength-Ansatz vor, der versucht, diesen besonderen Eigenheiten gerecht zu werden. Anschließend werden im Kapitel 7 die Ergebnisse zusammengefasst, die bei dem Vergleich des SentiStrength-Ansatzes mit verschiedenen Ansätzen des Machine-Learning ermittelt wurden.

## 2 Sentimentanalyse

Die Sentimentanalyse, auch bekannt als „Opinion Mining“, beschreibt die Entnahme von als positiv oder negativ zu bewertenden Meinungen aus (unstrukturiertem) Text [10]. Abstrahiert betrachtet sind in jedem Verfahren zu einer solchen Analyse mindestens drei Phasen der Betrachtung – oftmals mehr, zur Steigerung der Korrektheit – erkennbar: Zunächst werden Sätze oder Satzteile extrahiert, die irgendeine Art von subjektiver Empfindung enthalten. Im Gegensatz zu subjektiven stehen hierbei die objektiven Informationen, die bei der Analyse durch die Anwendung dieses ersten Schrittes zumeist komplett aus der Betrachtung gezogen werden. Extrahierte Sätze oder Satzteile werden anschließend bezüglich ihres Sentiments untersucht. In diesem Teil unterscheiden sich die meisten Verfahren, da hier die eigentliche Bewertung der subjektiven Empfindung stattfindet. Der letzte Schritt besteht darin, das Sentiment des kompletten zu betrachtenden Textes zu bestimmen, etwa durch Zusammenfassen der Sentiments der einzelnen extrahierten Bereiche. Unter der Bestimmung des Sentiments kann man in diesem Kontext beispielsweise die Klassifizierung der betrachteten Textabschnitte, Sätze oder Satzteile in positiv, neutral oder negativ und die Bewertung der Stärke des Sentiments anhand verschiedener Skalen verstehen [15]. Es gibt auch Sentimentanalysen, die versuchen andere Stimmungen zu erkennen, allerdings gilt das nicht für die im Folgenden vorgestellten Verfahren.

Die Idee der Einteilung in positiv, negativ oder neutral und der Bewertung nach einer bestimmten Skala basiert auf den Erkenntnissen aus der Emotionspsychologie. Danach fallen unter Emotionen u.a. Empfindungen wie Freude, Trauer, Angst und Überraschung. Im Dimensionsmodell von Russel [13] lassen sich Emotionen grundsätzlich folgendermaßen einteilen: Man spricht von einer Wertigkeit (positiv oder negativ) und einer Stärke (niedrig oder hoch), die eine Emotion besitzen kann. Dabei zeigen andere Forschungen weiterhin, dass als positiv und negativ bewertete Empfindungen auch - relativ unabhängig voneinander - koexistieren können [5,7,24,23].

## 3 Menschliche Einschätzung des Sentiments

Ob eine Sentimentanalyse richtig arbeitet oder nicht, hängt davon ab, wie man dieses „richtig“ definiert. Eine Bewertung etwas Subjektivem wie Empfindungen ist selbst nicht-automatisiert keine triviale Angelegenheit. So individuell die Menschen sind, so verschieden ist ihre Einschätzung zu Emotionen. Sie hängt unmittelbar mit erlebten Erfahrungen, ganz persönlichem Empfinden [2] und oftmals auch dem Geschlecht des bewertenden Individuums zusammen [17]. Eine automatisierte Sentimentanalyse braucht aber einen Vergleichswert zur Bewertung ihrer Korrektheit. Dennoch scheint der Durchschnitt von Bewertungen durch Menschen eine sinnvolle Messlatte als Standard zum Bestimmen von Wertigkeit und Stärke einer Emotion zu sein. Dabei gilt es zu beachten, dass bei Bestimmung eines Richtwertes zur Analyse der Korrektheit einer Sentimentanalyse

die Interrater-Reliabilität der bewertenden Personen möglichst übereinstimmt. Dieses haben die Entwickler bei der Evaluierung ihrer Verfahren weitestgehend beachtet.

## 4 Beispiel eines Machine-Learning-Ansatzes

Ein Ansatz zur Sentimentanalyse ist das Machine-Learning. Turney stellt in seinem Paper einen einfachen (unüberwachten) Machine-Learning-Algorithmus vor, der folgend zusammengefasst nach [22] beschrieben wird. Beim Machine-Learning-Ansatz wird "versucht, dem System 'beizubringen', welche besonderen Merkmale typisch für positive, neutrale oder negative Sätze und Texte sind"[15]. Im Gegensatz dazu stehen wörterbuchbasierte Ansätze, die (feste) Wörterbücher nutzen, um das Sentiment zu bestimmen, wie etwa der SentiStrength-Ansatz, der im Kapitel 6 vorgestellt wird. Der Machine-Learning-Ansatz von Turney enthält typische Elemente des Machine-Learning allgemein und dient an dieser Stelle somit als Beispiel für diese Ansätze. Die Intention dieses Algorithmus ist es, das Sentiment in Produkt-Reviews zu bestimmen, um daraus abzuleiten, ob im jeweiligen Review das darin beschriebene Produkt empfohlen wird oder nicht.

Der Machine-Learning-Ansatz von Turney bekommt als Input ein Review und produziert als Output die Klassifikation dieses in „empfohlen“ oder „nicht empfohlen“. Dazwischen werden folgende Schritte ausgeführt:

Als erstes werden Ausdrücke, die ein Sentiment enthalten identifiziert und extrahiert. Dabei sind Adjektive und Adverbien grundsätzlich gute Indikatoren, dass in diesem Bereich des Satzes ein Sentiment ausgedrückt sein könnte [24,26,25]. Dabei würden einzelne extrahierte Adjektive und Adverbien nicht genügend Kontext liefern, um das Sentiment des Gesamtausdrucks eindeutig zu bestimmen. Zum Beispiel würde das Adjektiv „unvorhersehbar“ im Kontext eines Review über Fahrzeuge im Zusammenhang mit der Beschreibung einer Lenkung etwas Negatives suggerieren. Hingegen würde das gleiche Adjektiv im Kontext eines Reviews über Filme im Zusammenhang mit der Beschreibung eines Filmausgangs durchaus mit etwas Positivem verbunden sein. Deshalb wurde für das Extrahieren von Ausdrücken, die ein Sentiment enthalten, ein Muster entwickelt, nach dem zwei aufeinanderfolgende Wörter herausgenommen werden, wenn sie diesem Muster entsprechen. Eines der Wörter ist demnach immer ein Adjektiv oder Adverb und das andere Wort liefert den kontextuellen Hintergrund. Die Einteilung der Wörter in Wortgruppen erfolgt mit Hilfe eines Part-Of-Speech-Taggers [3].

Der nächste Schritt besteht darin, für diese extrahierten Ausdrücke das Sentiment zu bestimmen. Hier kommt der sogenannte PMI-IR-Algorithmus (Pointwise Mutual Information and Information Retrieval) zum Einsatz. Die Annahme hinter der Entwicklung des Algorithmus ist, dass Wörter, die ein positives Sentiment ausdrücken, tendenziell häufiger mit weiteren Wörtern, die ein positives Sentiment ausdrücken, vorkommen als mit Wörtern, die ein negatives Sentiment ausdrücken und umgekehrt [18]. Als Bezugswort zur Bestimmung eines positiven Sentiments wurde das Wort „excellent“ und zur Bestimmung eines negativen Sentiments das Wort „poor“ ausgewählt. Der

Algorithmus berechnet die Ähnlichkeit des extrahierten Ausdrucks mit den Bezugswörtern indem er Anfragen zu einer Suchmaschine sendet und die Anzahl der Übereinstimmungen als Wahrscheinlichkeit wertet, in der dieser Ausdruck im Zusammenhang mit dem jeweiligen Bezugswort auftaucht. Der Logarithmus aus der somit bestimmten Wahrscheinlichkeit des Auftretens wird als PMI des betrachteten Ausdrucks interpretiert. Die Differenz zwischen dem PMI des Ausdrucks und „excellent“ und dem PMI des Ausdrucks und „poor“ stellt gleichzeitig die Wertigkeit (gegeben durch das Vorzeichen) und die Stärke (gegeben durch den Zahlenwert) des Sentiments des betrachteten Ausdrucks dar.

Zusammengefasst entsteht folgende Gleichung zur Bestimmung des Sentiments eines betrachteten Ausdrucks:

$$\begin{aligned} & \textit{Sentiment}(\textit{Ausdruck}) \\ &= \textit{PMI}(\textit{Ausdruck}, \textit{„excellent“}) - \textit{PMI}(\textit{Ausdruck}, \textit{„poor“}) \\ &= \log_2 \left[ \frac{p(\textit{Ausdruck AND „excellent“})}{p(\textit{Ausdruck}) \cdot p(\textit{„excellent“})} \right] - \log_2 \left[ \frac{p(\textit{Ausdruck AND „poor“})}{p(\textit{Ausdruck}) \cdot p(\textit{„poor“})} \right] \end{aligned}$$

Mit *AND* ersetzt durch *NEAR* (zeigte bessere Ergebnisse bei der Bestimmung des PMI nach [21]),  $p(X)$  interpretiert als  $\textit{hits}(X)$  (Treffer, die bei der Suche nach  $X$  ausgegeben wurden) und der Anwendung der Logarithmengesetze entsteht:

$$= \log_2 \left[ \frac{\textit{hits}(\textit{Ausdruck NEAR „excellent“}) \cdot \textit{hits}(\textit{„poor“})}{\textit{hits}(\textit{Ausdruck NEAR „poor“}) \cdot \textit{hits}(\textit{„excellent“})} \right]$$

Im letzten Schritt wird das gesamte Review anhand der Sentiments der darin vorkommenden Ausdrücke klassifiziert. Dazu wird der durchschnittliche Wert aus den Stärken der Sentiments im gesamten Review bestimmt. Hat der durchschnittliche Wert ein positives Vorzeichen, bedeutet dies, dass das Review als Empfehlung gewertet wird - bei negativem Vorzeichen als keine Empfehlung.

Der Algorithmus wurde ausgiebig getestet. Die Auswertungen bestätigten, dass das Verfahren im Durchschnitt eine Genauigkeit von 74% aufweist. Dabei wurden Reviews aus verschiedenen Themengebieten zur Evaluierung genutzt. Die Genauigkeit der Messung hängt demnach davon ab, aus welchem Themengebiet die Reviews entnommen wurden. Reviews über Fahrzeuge wurden zu durchschnittlich 84% genau bewertet, hingegen Reviews über Filme nur zu durchschnittlich 66%. Allerdings unterscheidet sich die Bestimmung des Sentiments in Reviews zu Filmen von Reviews zu anderen Produkten. Filme enthalten oftmals Szenenbeschreibungen, die Unbehaglichkeit suggerieren, aber den Film trotzdem nicht abwerten, eventuell sogar eher aufwerten.

## 5 Sentimentanalyse in sozialen Medien

Die meisten existierenden Algorithmen sind tendenziell für kommerzielle Zwecke entwickelt worden. Darin gilt es eher Meinungen zu einem Produkt zu erkennen

und nicht das Nutzerverhalten zu untersuchen. Damit ist die Intention auch nur das übergreifende Sentiment eines Textes zu bestimmen. Für die Einteilung von Produktbewertungen ist diese Herangehensweise durchaus sinnvoll, jedoch im Kontext einer Sentimentanalyse der Texte in sozialen Medien eher oberflächlich. Diese Texte enthalten oftmals beides: positiv und negativ zu bewertende Empfindungsäußerungen und diese in verschiedener Stärke.

Weiterhin basiert die Bestimmung des Sentiments in Texten mit Hilfe von Machine-Learning-Algorithmen. Diese sind auf so kurze Texte (im Mittel: 13 Wörter) [20], wie sie in der Kommunikation innerhalb der sozialen Medien vorkommen, nicht ausgelegt.

Andere Probleme entstehen bei näherer Betrachtung der verwendeten Sprachausprägungen innerhalb der sozialen Medien. Die Kommunikation erfolgt oftmals mit der Hilfe von abgekürzten Wörtern und Sätzen, unter Verwendung von Emoticons sowie wiederholten Satzzeichen und Buchstaben zur Betonung [1,4]. Der verstärkte Gebrauch von Slang und Dialekten stellt, neben der häufig fehlerhaften Orthografie und Grammatik, ebenfalls neue Herausforderungen an Algorithmen der Sentimentanalyse. Die meisten Analyseverfahren starten mit dem Part-Of-Speech-Tagging, das in der Regel von korrekter Rechtschreibung und Grammatik abhängig ist. Selbst unter Verwendung einer Rechtschreibkorrektur würden nur solche Verstöße erkannt und beseitigt werden, die vom Autor unbeabsichtigt aufgetreten sind [8,12].

Diese Eigenheiten der Textausprägungen innerhalb der Kommunikation in sozialen Medien erschweren die Bestimmung ihres Sentiments. Zusammengefasst ist ein neuer darauf weitestgehend spezialisierter Ansatz der Sentimentanalyse notwendig, um diesen Anforderungen möglichst gerecht zu werden.

## 6 SentiStrength-Ansatz

SentiStrength ist ein wörterbuchbasierter Ansatz, der eigens dafür entwickelte Methoden anwendet, um gleichzeitig positive und negative Sentiments aus kurzen informalen und elektronischen Texten zu bestimmen. Dabei nutzt SentiStrength insbesondere eigene Methoden, um die tatsächlich vorkommende Grammatik und den Schreibstil des Cyberspace auszuwerten. [19]

Die Eingabe des Algorithmus ist somit ein kurzer informaler elektronischer Text. Die Ausgabe ist dann eine Bewertung der in diesem Text ausgedrückten Sentiments. Sowohl die Stärke jedes gefundenen positiven Sentiments als auch jede eines negativen Sentiments werden simultan und zusammen mit dem (je nach Auswahl) durchschnittlichen oder jeweils höchsten Wert beider Orientierungen ausgegeben.

Der Kern dieses Algorithmus besteht aus einer Nachschlagtabelle (im Folgenden als Wörterbuch bezeichnet), die Begriffe enthält, die mit dem Ausdruck eines Sentiments in Verbindung gebracht werden. Dieses Wörterbuch wurde zunächst manuell mit der Hilfe von Ratern erstellt und anschließend durch einen Trainingsalgorithmus optimiert. Zu jedem der darin enthaltenen Begriffe ist eine ganze Zahl aus dem Bereich  $[-5, 5] \setminus \{0\}$  zugeordnet. Der Zahlenwert gibt die Stärke und das Vorzeichen die Wertigkeit des Sentiments an. Dabei werden die Zahlen +1 und -1 jeweils mit der Wertigkeit neutral, die anderen positiven Zahlen mit der Wertigkeit positiv und dementsprechend die anderen negativen Zahlen mit der Wertigkeit negativ assoziiert. Ein Begriff, der beispielsweise mit der Zahl -5 belegt ist, wird im Allgemeinen also mit extrem negativem Sentiment in Verbindung gebracht. Neben diesem Wörterbuch sind weitere Wörterbücher und spezielle auf die Erkennung des Sentiments in sozialen Medien ausgelegte Interpretationsalgorithmen zur Auswertung integriert. [19]

Nachdem die erste Version von SentiStrength einige Defizite aufwies, wurde es überholt und eine zweite Version von SentiStrength kam heraus. SentiStrength in seiner ursprünglichen Version konnte in die Kategorie „überwachte wörterbuchbasierte Sentimentanalyse“ eingeteilt werden, da seine Wörterbücher nach der manuellen Erstellung mit einem Trainingsalgorithmus optimiert wurden. Er wird im folgenden Unterkapitel 6.1 beschrieben. Parallel zur Überholung der ersten überwachten Version von SentiStrength wurde zusätzlich eine unüberwachte Ausprägung (d.h. ohne weitere Optimierung) entwickelt, womit insgesamt SentiStrength in seiner zweiten Version entstand. Die wesentlichen Änderungen werden im Unterkapitel 6.2 aufgelistet. [18]

### 6.1 Überwachter SentiStrength-Ansatz (Version 1)

Der Algorithmus des SentiStrength-Ansatzes, folgend beschrieben nach [19], wird auf jeden Satz separat angewendet. Satzbegrenzungen werden entweder am Vorhandensein eines Zeilenumbruchs oder eines Satzzeichens erkannt, das kein Emoticon darstellt. Der Kern des Algorithmus ist das Wörterbuch mit 298 Begriffen, die mit einem positiven Sentiment, 465 Begriffen, die mit einem negativen Sentiment assoziiert werden und dem Wert ihres Sentiments, der gleichzeitig durch sein Vorzeichen die Wertigkeit und durch den Zahlenwert die Stärke angibt. Manche Begriffe besitzen am Ende ein Sternchen, das suggeriert, dass dieser Begriff einen ganzen Wortstamm abbildet und somit alle Wörter, die damit beginnen diesem zugeordnet werden können. Unter den verzeichneten Wörtern oder Wortgruppen befinden sich sowohl Standard- als auch Nichtstandardbegriffe, die jedoch allgemein gebräuchlich sind. Dabei drücken manche explizit eine Emotion aus und andere tragen indirekt zur Konstruktion eines positiven oder negativen Kontexts bei.

Dieses Wörterbuch wurde zunächst durch menschliche Einschätzung erstellt und danach durch automatische Modifizierung (Trainingsphase) überholt. Die Modifizierung geschah mit Hilfe eines Trainingsalgorithmus, der für jeden

Term berechnet, ob eine Erhöhung oder Verringerung seiner Stärke um jeweils eins die allgemeine Korrektheit der Klassifizierungen erhöhen würde. Jede Änderung von mindestens zwei wird notiert. Ein Festhalten jeder Änderung auch um nur eins würde den Algorithmus zu sehr verlangsamen, weshalb man sich für die untere Grenze von zwei entschied, trotz der Gefahr, dass damit eventuell manche nützliche, aber selten vorkommende Änderungen, vernachlässigt werden. Der Algorithmus betrachtet jeden Eintrag so lange, bis keine weiteren Änderungen die Korrektheit erhöhen würden.

Ein eigener Rechtschreibkorrekturalgorithmus beseitigt zur Analyse doppelte Buchstaben, wenn damit aus einem Nichtstandard- ein Standardwort entstehen würde. Gewöhnliche Rechtschreibkorrekturalgorithmen würden Probleme mit dem in sozialen Medien üblichen Jargon bekommen.

Zusätzlich zu dem Kernwörterbuch existiert ein weiteres, das Begriffe beinhaltet, die eine Negation darstellen und bei Vorkommen den Wert des Sentiments des darauffolgenden Wortes in seiner Wertigkeit umkehren. Ein weiteres Wörterbuch mit Emoticons ermöglicht es, deren Einbindung in Texten mit einem Sentiment von entweder +2 oder -2 zu bewerten.

Weiterhin existiert ein Wörterbuch, das Begriffe enthält, die die Stärke eines Sentiments des darauffolgenden Wortes entweder um eins bis zwei erhöhen oder um eins erniedrigen. Dabei wird bei einem Vorhandensein von mindestens zwei zusätzlichen Buchstaben (die ein Standardwort von einem Nichtstandardwort unterscheiden) ebenfalls eine Erhöhung oder Verringerung der Stärke des Sentiments dieses Wortes um eins eingeleitet, es sei denn, es handelt sich um ein solches, dessen Wertigkeit als neutral bezeichnet ist. Ein „neutrales“ Wort wird durch die mehrmalige Wiederholung von Buchstaben mit einem Wert von +2 belegt. Eine ähnliche Wirkung hat das Vorkommen wiederholter Satzzeichen mit mindestens einem darin vorhandenen Ausrufezeichen. Die Stärke des Sentiments des unmittelbar vorangehenden Wortes oder Satzes wird dadurch um eins vergrößert oder reduziert.

Das Wort „miss“ ist das einzige, das sowohl mit einer positiven als auch negativen Wertigkeit belegt ist. Es wird zumeist in einem Kontext verwendet, in dem gleichzeitig Betrübnis und Liebe ausgedrückt werden (z.B. „I miss you.“). Eine weitere Besonderheit ist, dass Sentiments mit einer negativen Wertigkeit in Fragen ignoriert werden. Die Annahme dabei ist, dass Fragen keinerlei Negativität enthalten können. Nicht ignoriert werden in Fragen hingegen Sentiments mit einer positiven Wertigkeit, da Testfälle gezeigt haben, dass viele Fragen ein als zumindest mild positiv zu bewertendes Sentiment ausdrücken. Außerdem wurde aufgrund weiterer Recherche festgelegt, dass ein Satz mit Ausrufezeichen zumindest ein positives Sentiment mit der Stärke +2 besitzen muss, wenn dieses nicht neutral ist.

Einige Ergänzungen des Verfahrens wurden ausgegliedert (in der nächsten Version aber teilweise wieder eingegliedert), da sie in der damaligen Form nicht zum erhofften Anstieg der Korrektheit geführt hatten. Zu nennen wären etwa die Identifizierung von Idiomen und semantischen Doppeldeutigkeiten (z.B.: „you rock“ und „I like rock music“).

Nach ausgiebigem Testen des SentiStrength-Ansatzes wurde festgestellt, dass diese Version bei der Bestimmung von negativen Sentiments auffallend schlechtere Ergebnisse geliefert hatte als bei der Bestimmung von positiven Sentiments. Weiterhin wurde SentiStrength nur auf einem Vertreter eines sozialen Netzwerkes getestet. Über die Effizienz auf anderen sozialen Medien konnte keine Aussage getätigt werden. Auch erschien es nützlich, den Algorithmus dahingehend zu verbessern, dass er auch in einer unüberwachten Version genügend korrekt arbeitet, um ihn für die Sentimentanalyse in sozialen Netzwerken einsetzen zu können, für die keine Trainingsdaten zur Verfügung stehen. Somit wurde diese Version des SentiStrength-Ansatzes überarbeitet und eine verbesserte Version kam dabei heraus, deren wesentlichen Anpassungen im folgenden Kapitel erläutert werden. [18]

## 6.2 Überwacher und unüberwachter SentiStrength-Ansatz (Version 2)

Diese Version des SentiStrength-Ansatzes, folgend beschrieben nach [18], basiert auf der vorherigen und folgt dem gleichen Prinzip. Der Grund der Veränderung war, dass neben der Verbesserung der Korrektheit im Erkennen von negativem Sentiment auch die Adaption des Algorithmus an eine breitere Auswahl von Textarten notwendig erschien. Der Algorithmus wurde soweit verbessert, dass er in die Lage versetzt wurde, selbst in seiner unüberwachten Version ausreichend korrekt zu arbeiten. Die durchgeführten Modifikationen werden im Einzelnen folgend aufgeführt.

Zunächst wurden viel mehr Wörter mit Sternchen versehen. Dies ermöglicht zwar eine größere Vielfalt an Wortausprägungen zu identifizieren, bringt aber damit ein neues Problem mit sich: Je mehr Buchstaben eines Begriffes variabel gelassen werden, desto größer wird die Gefahr, dass man neben den verschiedenen Ausprägungen des Begriffes auch zusätzlich völlig fremde Begriffe, die aber dem Muster entsprechen als passend erkennt. Ein Beispiel wäre der Begriff „admir\*“, zu dessen Entsprechung auch die Erwähnung von „admiral“ zugeordnet werden würde. Aus diesem Grund wurden ebenfalls solche Ausnahmen explizit zum Kernwörterbuch hinzugefügt. Der Algorithmus sucht sich jeweils den längsten passenden Begriff aus dem Wörterbuch heraus, wenn mehrere vorhanden sind.

Sobald Begriffe, die mit einem negativen Sentiment belegt sind negiert werden, werden sie damit eher mit einem neutralen Sentiment als mit einem positiven assoziiert, weshalb hierzu ebenfalls eine Anpassung vorgenommen



wurde. Das Wörterbuch mit Idiomen wurde hinzugefügt, um u.a. Begriffe zu identifizieren, die in bestimmten Verbindungen mit anderen in ihrem Ausdrücken eines Sentiments unterschiedlich interpretiert werden müssen. Ein Beispiel dazu wäre etwa „I like something“ und „this is like something else“, was deutlich macht, dass diese Sätze beide den Begriff „like“ enthalten, aber nicht beide demselben Sentiment zugeordnet werden dürfen. Sobald ein Begriff in der Ausprägung vorkommt, wie er im Wörterbuch mit Idiomen verzeichnet ist, wird jede ursprüngliche Bewertung des Sentiments mit der aus diesem Wörterbuch überschrieben. Die Sonderregelung, dass das Vorkommen von Begriffen in Fragen ignoriert werden, die ein negatives Sentiment ausdrücken, wurde entfernt. Dazu geführt hat die neue Erkenntnis, dass es durchaus möglich ist, dass auch Fragen ein negatives Sentiment enthalten können.

Der Algorithmus wurde vor und nach der Modifizierung mit Hilfe des Trainingsalgorithmus angewendet. Dabei wurde festgestellt, dass die meisten durchgeführten Modifikationen relativ geringen Einfluss auf die Korrektheit der Ergebnisse haben. Diese Erkenntnis führt zu der Behauptung, dass auch der unüberwachte SentiStrength-Ansatz schon genügend gut arbeitet und außerdem noch den Vorteil mit sich bringt, dass man ihn in Domänen einsetzen kann, für die wenig oder gar keine Trainingsdaten zur Verfügung stehen.

## 7 Machine-Learning-Ansätze vs. SentiStrength-Ansatz

Die Machine-Learning-Ansätze zeichnen sich insbesondere dadurch aus, dass sie neben explizit vorkommenden Begriffen, die direkt mit einem Sentiment assoziiert werden auch implizite Informationen über ein vorhandenes Sentiment in Texten interpretieren können. Für viele Anwendungen, wie zur Forschung und zur Analyse von Produktbewertungen, sind solche Ansätze somit am besten geeignet. Sie benötigen zwar eine große Menge an Trainingsdaten und fordern eine saubere Rechtschreibung und korrekte Anwendung der Grammatik, können aber sobald sie auf eine bestimmte Domäne ausgerichtet wurden die besten Ergebnisse erzielen. Dies sind die größten Stärken der Machine-Learning-Ansätze im Vergleich zum SentiStrength-Ansatz. Allerdings werden diese Stärken zu ihren Schwächen, wenn man sie zur Bestimmung des Sentiments in Texten verwendet, die in sozialen Medien verfasst werden. SentiStrength arbeitet mit all seinen Wörterbüchern und Sonderregeln in diesem Umfeld durchschnittlich präziser [19].

Machine-Learning-Ansätze sind in der Regel überwacht. In den meisten Fällen müssen für die korrekte Arbeitsweise zunächst Testdaten für jede neue Domäne manuell angelegt werden, um die Machine-Learning-Algorithmen zu trainieren. Das ist für den Einsatz in sozialen Medien nicht optimal, da es darin einfach zu viele Themenfelder gibt und diese auch direkt nebeneinander und miteinander vernetzt vorkommen können. Wurde ein Algorithmus auf einem Gebiet trainiert, kann es passieren, dass er danach auf einem anderen

nicht mehr korrekt arbeitet. Der SentiStrength-Ansatz kann hier ohne großes Einbüßen der Korrektheit der Auswertung in seiner unüberwachten Version eingesetzt werden. [18]

Die Bestimmung von indirekt vorhandenem Sentiment führt in Texten aus sozialen Medien oftmals zur Fehlinterpretation des darin ausgedrückten Sentiments. Es kann passieren, dass etwa das Vorhandensein von Begriffen aus bestimmten Themenbereichen, die zeitweise vermehrt den Kontext zu Texten mit vorwiegend negativem Sentiment lieferten, dazu führt, dass ein Text mit einem positivem oder neutralem Sentiment negativer interpretiert wird, als es sein sollte. Zum Beispiel würde ein Text mit dem Thema Irak eventuell mit einem negativen Sentiment bewertet, weil der Algorithmus etwa auf der Grundlage von Nachrichtenmeldungen trainiert wurde oder aufgrund der Debatten aus den USA. Nun verfassen aber auch Menschen im Irak in sozialen Medien über ihr Land Texte mit neutral oder positiv zu bewertendem Sentiment. Es ist dann schlichtweg falsch, diese in irgendeiner Weise mit einem negativen Sentiment in Verbindung zu bringen, wenn nur der Kontext Irak implizit mit etwas Negativem assoziiert wird. Ein ähnliches Problem tritt auf, wenn man die Entwicklung von Trends in sozialen Medien beobachten möchte. Diese Probleme haben ihren Ursprung darin, dass der Machine-Learning-Ansatz von der Domäne abhängig ist, auf der er trainiert wurde [19]. SentiStrength ist domänenunabhängig und robust [18], zumindest was das Bestimmen des Sentiments in Texten aus Sozialen Medien betrifft, wofür er allerdings auch ausgelegt wurde.

Noch hinzukommt, dass der Machine-Learning-Ansatz bei der Bewertung nicht unterscheidet, in welcher Zeit der zu bewertende Text erstellt wurde. Wenn man implizite Informationen in Texten, die in sozialen Medien verfasst sind, betrachtet, spielt der Zeitraum indem er verfasst wurde eine nicht unwichtige Rolle [19]. Zum Beispiel wenn eine Debatte über das World Trade Center vor dem 11. September 2001 verfasst wurde, ist ihr Sentiment sicherlich ein anderes als danach. Da der SentiStrength-Ansatz solche impliziten Informationen vernachlässigt kann dieses Problem nicht entstehen.

## 8 Fazit und Ausblick

Die Sentimentanalyse wird in allen möglichen Bereichen zunehmend interessanter. Bisher sind Anwendungen als nützlich angedacht wie etwa das Erkennen der Popularität von Filmen aus mehreren Online-Reviews und die Diagnose, welche Autoteile in besonderem Maß gemocht oder nicht gemocht werden. Dies geschieht, indem man auf bestimmten Webseiten oder Foren eine Sentimentanalyse durchführt. Eine weitere interessante Anwendung der Sentimentanalyse ist es, Suchmaschinen und andere Anwendungen, die eine Suche beinhalten, dahingehend zu optimieren, dass sie in die Lage sind, Suchergebnisse sortiert nach Treffern des gewünschten vorkommenden Sentiments auszugeben [6]. In sozialen Medien könnte die Sentimentanalyse die Überwachung der verfassten

Inhalte unterstützen um eventuellen Missbrauch leichter und schneller zu erkennen [16] oder um für die Allgemeinheit potentiell gefährliche User frühzeitig aufzufinden. Auch denkbar sind Anwendungen, die eine möglichst kurze Zusammenfassung eines Textes automatisch generieren können, indem sie den Satz mit dem größten Sentiment des vorherrschenden Sentiments des gesamten Textes extrahieren. Ebenfalls denkbar ist ein Tool, das bei der Erstellung von Texten überprüft, ob sie genügend neutral verfasst sind oder doch ein positives oder negatives Sentiment in der Argumentation vorhanden ist. Diese kann dann etwa für akademische Zwecke eingesetzt werden [22].

Mit dem SentiStrength-Ansatz ist eine Bestimmung des Sentiments selbst für sehr kurze Texte, die in sozialen Medien verfasst wurden, möglich [19]. Dabei ist sein Algorithmus robust und kann dort selbst ohne die Verfügbarkeit von Trainingsdaten in dessen unüberwachter Version mit genügender Treffsicherheit eingesetzt werden [18]. Mit den Machine-Learning-Ansätzen ist es weiterhin möglich, in die Sentimentanalyse auch den Kontext mit einzubeziehen. Für andere angedachte Anwendungen wäre es sinnvoll, dass neben der Unterscheidung zwischen Emotion und Information, auch eine breitere Analyse mit Erkennung von Hintergründen, warum ein Sentiment an dieser zu analysierenden Stelle vorkommt, betrieben wird. Eine intelligente Miteinbeziehung von Kontext und Differenzierung zwischen Intention, Argumentation und Spekulation in den betrachteten Texten [27] wäre anzustreben. Im Moment ist dazu (noch) kein Ansatz in der Lage. Alle Sentimentanalyse-Algorithmen haben außerdem gemeinsam, dass Sarkasmus und Ironie bei der Ermittlung des Sentiments in Texten Probleme bereiten. Die Detektion solcher Konstrukte ist wohl eines der anzustrebenden Ziele in zukünftiger Weiterentwicklung von allen Ansätzen zur Sentimentanalyse [18].

Die Entwicklung eines hybriden Ansatzes, der einen wörterbuchbasierten Ansatz und einen Machine-Learning-Ansatz intelligent miteinander kombiniert, scheint langfristig für die Bandbreite an verschiedenen Anwendungen am sinnvollsten zu sein [22,15]. Trotz aller Verbesserungen ist die Validierung eines Algorithmus darauf angewiesen, dass eine ordentliche Referenz vorhanden ist, welche Aussagen in welcher Verbindung mit einem Sentiment zu werten sind. Doch, wie bereits festgestellt, ist es mit der Einschätzung des Sentiments schon von der menschlichen Seite her nicht immer eindeutig. Deshalb wird man wahrscheinlich trotz einer ausgefeilten Technik im Ermitteln eines Sentiments in Texten immer darauf angewiesen sein, die Inhalte selbst und nach eigenem subjektivem Einschätzen nochmals zu reflektieren. Anwendungen, die eine Sentimentanalyse anbieten, werden wohl die Selektion und Klassifikation nach Meinungen und Empfindungen erleichtern, aber die menschliche Einschätzung nicht ersetzen können. Am Schluss entscheidet dann doch jeder Mensch für sich allein, in welchem Maße und ob überhaupt der zu betrachtende Text positiv, negativ oder neutral zu bewerten ist.

## Literatur

1. Baron, N.: Language of the internet. In: The Stanford handbook for language engineers, Stanford, CA (2003) 59–127
2. Barrett, L.F.: Valence is a basic building block of emotional life. *Journal of Research in Personality* **40** (2006) 35–55
3. Brill, E.: Some advances in transformation-based part of speech tagging. In: In Proceedings of the twelfth national conference on artificial intelligence. (1994) 722–727
4. Crystal, D.: Language and the internet (2nd ed.). Technical report, Cambridge University Press, Cambridge, United Kingdom (2006)
5. Diener, E., Emmons, R.A.: The independence of positive and negative affect. *Journal of Personality and Social Psychology* **47** (1984) 1105–1117
6. Hearst, M.: Direction-based text interpretation as an information access refinement. In Jacobs, P., ed.: *Text-Based Intelligent Systems*. Lawrence Erlbaum Associates (1992) 257–274
7. Huppert, F.A., Whittington, J.E.: Evidence for the independence of positive and negative well-being: Implications for quality of life assessment. *British Journal of Health Psychology* **8** (2003) 107–122
8. Kukich, K.: Techniques for automatically correcting words in text. *ACM Computing Surveys* **24** (1992) 377–439
9. Naveed, N., Gottron, T., Kunegis, J., Alhadi, A.C.: Bad News Travel Fast: A Content-based Analysis of Interestingness on Twitter. (2011)
10. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval* **2** (2008) 1–135
11. Pennebaker, J.W., Mehl, M.R., Niederhoffer, K.G.: Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology* **54** (2003) 547–577
12. Pollock, J.J., Zamora, A.: Automatic spelling correction in scientific and scholarly text. *Communications of The ACM* **27** (1984) 358–368
13. Russell, J.A.: Affective space is bipolar. *Journal of Personality and Social Psychology* **37** (1979) 345–356
14. Schönhalz, D.: Opinion mining und sentiment-analyse im web 2.0. Webseite (März 2010) <http://www.social-media-magazin.de/index.php/inhalt/opinion-mining-und-sentiment-analyse-im-web-20.html>.
15. Schönhalz, D.: Lernende maschinen. Webseite (März 2011) <http://www.social-media-magazin.de/index.php/heft-03-2011/lernende-maschinen.html>.
16. Spertus, E.: Smokey: Automatic recognition of hostile messages. In: Proceedings of the 14th National Conference on Artificial Intelligence and 9th Innovative Applications of Artificial Intelligence Conference (AAAI-97/IAAI-97), Menlo Park, AAAI Press (1997) 1058–1065
17. Stoppard, J.M., Gruchy, C.D.G.: Gender, Context, and Expression of Positive Emotion. *Personality and Social Psychology Bulletin* **19** (1993) 143–150
18. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology* **63**(1) (2012) 163–173
19. Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., Kappas, A.: Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology* **61**(12) (2010) 2544–2558

20. Thelwall, M., Wilkinson, D., Uppal, S.: Data mining emotion in social network communication: Gender differences in myspace. *J. Am. Soc. Inf. Sci. Technol.* **61**(1) (January 2010) 190–199
21. Turney, P.: Mining the web for synonyms: Pmi-ir versus lsa on toefl (2001)
22. Turney, P.D.: Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In: *ACL'02: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. (2002) 417–424
23. Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of Personality and Social Psychology* **54**(6) (1988) 1063–1070
24. Watson, D.: Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology* **54** (1988) 1020–1030
25. Wiebe, J., Bruce, R., Bell, M., Martin, M., Wilson, T.: A corpus study of evaluative and speculative language. In: *Proceedings of the 2nd ACL SIGdial Workshop on Discourse and Dialogue (SIGdial-2001)*, Aalborg, Denmark (2001) 186–195
26. Wiebe, J.M.: Learning subjective adjectives from corpora. In: *In AAAI*. (2000) 735–740
27. Wilson, T.: *Fine-grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of private states*. PhD thesis, Intelligent Systems Program, University of Pittsburgh (2007)

## Urheberschaftserklärung

Ich versichere hiermit, die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen verwendet zu haben. Alle wörtlichen und sinngemäßen Entlehnungen sind unter genauer Angabe der Quelle kenntlich gemacht. Die Arbeit wurde weder in dieser noch in ähnlicher Form als Prüfungsleistung für eine andere Prüfung eingereicht.

---

Vorname

---

Nachname

---

Ort, Datum

---

Unterschrift