

Bachelorarbeit

Deutscher Titel der Bachelorarbeit	
Englischer Titel der Bachelorarbeit	
Verfasser/in Familiennamen, Vorname(n)	
Matrikelnummer	
Studium	
Beurteiler/in Titel, Vorname(n), Familienname	

Hiermit versichere ich, dass

1. ich die vorliegende Bachelorarbeit selbständig und ohne Verwendung unerlaubter Hilfsmittel verfasst habe. Alle Inhalte, die direkt oder indirekt aus fremden Quellen entnommen sind, sind durch entsprechende Quellenangaben gekennzeichnet.
2. die vorliegende Arbeit bisher weder im In- noch im Ausland zur Beurteilung vorgelegt bzw. veröffentlicht worden ist.
3. diese Arbeit mit der beurteilten bzw. in elektronischer Form eingereichten Bachelorarbeit übereinstimmt.
4. (nur bei Gruppenarbeiten): die vorliegende Arbeit gemeinsam mit

entstanden ist. Die Teilleistungen der einzelnen Personen sind kenntlich gemacht, ebenso wie jene Passagen, die gemeinsam erarbeitet wurden.

Datum

NGerwin

Unterschrift

Bachelorarbeit

Automatische Verwertung von Mikro-Influencer-Feedback durch Sentimentanalyse

Eingereicht bei:

Dr. Ulrike Phieler

Institute for Interactive Marketing & Social Media

Wirtschaftsuniversität Wien

Eingereicht von:

Nikolaus Czernin

Friedrich Schmidtplatz 4/21

1080, Wien

Mobil: 0680 4450807

Mail: h11721138@s.wu.ac.at

Studium: Wirtschafts- & Sozialwissenschaften — Wirtschaftsinformatik

Fachsemester: W2021

Matrikelnummer: 11721138

Inhaltsverzeichnis

INHALTSVERZEICHNIS	2
ABBILDUNGSVERZEICHNIS	3
FORMELVERZEICHNIS	3
TABELLENVERZEICHNIS.....	3
1 EINLEITUNG	4
2 KAMPAGNE	8
2.1 FORMAT	8
2.2 BEISPIEL	8
2.3 EVALUIERUNG UND OPTIMIERUNG.....	9
2.4 DATEN.....	10
3 SENTIMENTANALYSE	11
3.1 THEORETISCHER HINTERGRUND	11
3.2 ANSÄTZE	15
3.2.1 <i>Maschinelles Lernen</i>	15
3.2.2 <i>Lexikon-basierte Ansätze</i>	15
3.2.3 <i>Wissensbasierter Ansatz</i>	17
3.3 ANSATZ DIESER ARBEIT	18
3.4 AUSWAHL DES SENTIMENT-LEXIKONS.....	19
4 DATENVERARBEITUNG.....	21
4.1 TOKENISIERUNG	21
4.2 PART-OF-SPEECH-TAGGING.....	21
4.3 STEMMING UND LEMMATISIEREN	22
4.4 GROß-KLEINSCHREIBUNG	22
4.5 STOPPWÖRTER	23
4.6 KONTEXTBEDINGTE EINFLÜSSE.....	23
4.7 SENTIMENTWERTE DER SÄTZE.....	25
4.8 NORMALISIERUNG DER SENTIMENTWERTE.....	26
5 ANALYSE	28
5.1 VERTEILUNG DER SENTIMENTE.....	28
5.2 ENTWICKLUNG DER SENTIMENTE	32
5.3 SENTIMENT-EXTREME	35
6 ZUSAMMENFASSUNG	38
ANHANG	40
LITERATURVERZEICHNIS.....	41
EIDESSTATTLICHE ERKLÄRUNG	48

Abbildungsverzeichnis

<i>Abbildung 1: Sentiment-Dichtefunktion: vor (schwarz, durchgehend) und nach (rot, strichliert) der Berücksichtigung der latenten Sentimente objektiver Worte</i>	26
<i>Abbildung 2: Sentiment-Verteilung nach Normalisierung mittels verschiedener Alpha-Werte</i>	27
<i>Abbildung 3: Sentiment-Verteilung der Kommentare</i>	29
<i>Abbildung 4: Sentiment-Verteilung der Blogeinträge</i>	29
<i>Abbildung 5: Sentiment-Verteilung der Kampagnen</i>	29
<i>Abbildung 6: Boxplot zur Anzahl der Kampagnen jedes Teilnehmers</i>	31
<i>Abbildung 7: Boxplot zum mittleren Sentiment jedes Teilnehmers</i>	31
<i>Abbildung 8: Sentiment-Entwicklung für Kampagne 17</i>	33
<i>Abbildung 9: Sentiment-Entwicklung für 50 Kampagnen</i>	33
<i>Abbildung 10: Sentiment-Entwicklung für Kampagnenmanager "nellisa"</i>	33
<i>Abbildung 11: Sentimente der Kampagne 198</i>	34
<i>Abbildung 12: durchschnittliche (türkis, durchgehend) und besonders negative (orange, strichliert) Sentimente der Kampagne 198</i>	35
<i>Abbildung 13: durchschnittliche (türkis, durchgehend) und besonders positive (pink, strichliert) Sentimente der Kampagne 198</i>	35
<i>Abbildung 14: Anzahl der Kommentare der Kampagne 198</i>	35
<i>Abbildung 15: Verteilung der Sentimente der Kampagne 198 Blogeintrag 7</i>	35

Formelverzeichnis

<i>Formel 1. Normalisierung von Sentimenten</i>	27
---	----

Tabellenverzeichnis

<i>Tabelle 1: positive und negative Worte</i>	12
<i>Tabelle 2: Sentiment-Modifikatoren</i>	13
<i>Tabelle 3: idiomatische Aussage</i>	13
<i>Tabelle 4: Emoji-Sentimente</i>	14
<i>Tabelle 5: Negationen durch Valenzumkehrung</i>	16
<i>Tabelle 6: Negationen durch Intensität-Subtraktion</i>	16
<i>Tabelle 7: Sentiment-Modifikation durch Großbuchstaben</i>	23
<i>Tabelle 8: Nutzung von „nie“ und seine Auswirkungen auf Sentimente</i>	24
<i>Tabelle 9: Nutzung von „aber“ und seine Auswirkung auf Sentimente</i>	24
<i>Tabelle 10: Alle Kommentare von Teilnehmer „blackstar1487“</i>	32
<i>Tabelle 11: besonders negative Kommentare</i>	36
<i>Tabelle 12: besonders negative Kommentare der Kampagne 198</i>	37

1 Einleitung

Dank seiner Effektivität und der immer häufiger genutzten Möglichkeiten seiner gezielten Generierung nimmt Word-of-Mouth (WOM) längst einen wichtigen Platz im Marketing Mix ein (Fulgoni und Lipsman 2015). Bei sogenannten Seeding-Kampagnen, einer Art von Kampagne, die gezielt WOM generieren soll, erhalten ausgewählte, mehr oder minder einflussreiche Zielgruppenmitglieder, sogenannte Seed-Agents, Produkt-Exemplare, in der Hoffnung, dass möglichst viele ihrer Mitmenschen zum Kauf verleitet werden können (Haenlein und Libai 2017, 71). Durchschnittlich finden 90 Prozent aller solchen markenbezogenen Unterhaltungen offline statt, also in Person oder per Telefonat (Keller und Fay 2016). Da unmöglich alle Kaufentscheidungen auf Begründungen durch Offline-WOM untersucht werden können, ohne die Käufer explizit danach fragen zu müssen, sind Schätzungen notwendig. Bewertungen, Kommentare und Berichte von den Seed-Agents, mit denen bereits Kontakt besteht, bilden die Grundlage für solche Einschätzungen. Hierbei kann es sich ohne Weiteres um tausende Texte handeln. Zwischen den Marken und der Verarbeitung dieser Texte stehen also zwei Hindernisse: Der für den Computer nicht immer leicht verständlichen Form der Daten und der, für den Menschen zu großen, Menge der Daten. Um also Schlüsse aus diesen Daten ziehen zu können, sind automatische Textanalysemethoden gefragt.

Diese Arbeit beschäftigt sich mit den Anwendungsmöglichkeiten von Natural Language Processing (NLP) im Bereich von WOM-Marketing. Diese Methoden ermöglichen die Verarbeitung von unstrukturierten Texten, die durch Text-Mining (Salloum, Al-Emran und Shaalan 2017) von diversen Quellen bezogen werden können. Sie analysieren gruppierte Zeichenfolgen und erkennen Muster und Zusammenhänge. Inhaltliche Themen und implizite Informationen wie Emotionen der Autoren können so beispielsweise extrahiert werden, um Kundenmeinungen quantifizieren zu können. Durch NLP sollen die ansonsten ungenutzten Informationen, die sich hinter den Kommentaren verbergen, genutzt werden, um einerseits Mehrwert für die Marken selber durch qualitative Analysen und weitere Metriken für die Qualität der Marketingkampagnen zu schaffen.

Im Fokus stehen Sentimentanalysen. Es ist für Menschen meist nicht schwer einzuschätzen, ob ein Kommentar gut- oder böswillig, subjektiv oder objektiv, oder sogar ehrlich oder sarkastisch gemeint ist. Computer können jedoch diese impliziten

Informationen nicht intuitiv auslesen. Algorithmen, die diese quantifizieren, sind erforderlich. Mittels Sentimentanalysen lassen Datenanalysten Computer diese unerschwelligen Emotionen automatisch erkennen, um große Mengen an unstrukturierten Daten effizient verarbeiten zu können. Positive Sentimente der Zielgruppe sind Marken selbstverständlich wichtig (Gräve 2019), Zufriedenheit mit einer Marke erhöht neben der Kaufwahrscheinlichkeit auch die Wahrscheinlichkeit, jene Marke an Mitmenschen weiterzuempfehlen. Mithilfe von Sentimentanalysen können Dinge wie Kundenzufriedenheit näherungsweise quantifiziert und somit als Einflussvariable für Weiterempfehlungsquoten, einer Erfolgsmetrik für WOM-Marketing, genutzt werden. Das erlaubt Unternehmen, sonst nicht beobachtbare Veränderungen der Meinungen ihrer Kunden zu analysieren und ihre Entscheidungen auf Daten, anstatt Intuition, zu stützen. Die überwältigende Menge an Daten, die hierfür oft als Grundlage dient, erfordert ebenfalls automatisierte Prozesse. Händische Quantifizierung von Sentimenten in Kommentaren ist schnell nicht mehr effizient und verlässlich genug. Man bräuchte für die Aufbereitung der Daten entweder Angestellte, die dafür beträchtlich viel Arbeitszeit aufbringen müssten, was nur bei kleinen Kampagnen mit wenig Daten sinnvoll ist, oder das Budget, um die Verarbeitung via Crowdsourcing auslagern zu können.

Ziel dieser Arbeit ist, das Potenzial von Sentimentanalysen bei der Evaluierung von Seeding-Kampagnen mittels einer lexikon-basierten Sentimentanalyse der Kommentare auf einer Plattform für solche Kampagnen aufzudecken. Die Aussagekraft der impliziten Informationen in Kommentaren in Kundenrezensionen und Social-Media-Texten wurde bereits erkannt und ihre Gewinnung und Nutzung erforscht. Jedoch bewegten sich diese Studien vorwiegend im Kontext von herkömmlichem Social Media, Analysen von Tweets oder Facebook-Kommentare (Vermeer, et al. 2019), oder Kundenrezensionen von Käufern (Zhang, Li und Chen 2012). Die Teilnehmer von Seeding-Kampagnen sind jedoch weder willkürliche Social-Media-Nutzer, noch schreiben sie aus freien Zügen Rezensionen auf einer Produktseite. Sie sind freiwillige, ausgewählte Produkttester, die aus eigenem Interesse an einer Kampagne teilgenommen haben. Dieser Motivationsunterschieds begründet eine Trennung der Anwendungsmöglichkeiten von Sentimentanalysen in diesen Kontexten, in denen Sentimentanalysen noch nicht ausreichend erforscht wurden.

Die generierte Menge an WOM hat nachweislich den größten positiven Einfluss auf Erfolgsmetriken wie Markenwert oder Verkaufszahlen, die aus WOM-Kampagnen resultieren (Chintagunta, Gopinath und Venkataraman 2010). Die ebenso nachweislich signifikanten Implikationen von Sentimenten für die Effektivität von WOM-Marketing sind Motivation für das Erstellen eines Sentimentanalysemodells für die Daten, die das Nebenprodukt einer solchen Kampagne sind. Mithilfe eines Analysemodells können den Kampagnenmanagern und den beauftragenden Marken zusätzlich Grundlagen zur Einschätzung der Kampagnenqualität dargelegt werden. Ebenso kann die Beziehung der Emotionalität und der Weiterempfehlungsrate erforscht werden. Korrelieren diese bei einem Produkt positiv, kann das unter Umständen zukünftige Marketingentscheidungen beeinflussen, emotional stärkere Reaktionen auf Produkte erfordern mehr Feinfühligkeit im Kundenumgang (Harvey 2021). Kommentare mit besonders starken Sentimenten sollten individuell analysiert werden, um besonders positive oder besonders negative Erfahrungen verstehen zu können. Darüber hinaus können Manager bei Kampagnen, die über mehrere Blogposts laufen, die Sentimente und Emotionalität über einen längeren Zeitraum analysieren. So könnte sich beispielsweise zeigen, ob das langfristige Nutzen eines Produktes oder eines Service zu anderen Kundenmeinungen führt. Begründungen für Spitzenwerte von Emotionen oder Veränderungen in der Emotionalität müssten nach der Sentimentanalyse entweder händisch oder automatisiert durch Topic-Modeling mittels Latenter Semantischer Analyse (Systemic Sciences 2021) oder eines TF-IDF-Algorithmus (Jayaswal 2020) extrahiert werden. Topic-Modeling würde jedoch den Rahmen dieser Arbeit sprengen. Die Ergebnisse dienen also einerseits als Einsichten in die Meinungen der Zielgruppe für die Brandmanager und andererseits als Metriken für die Effektivität der Kampagnen und ihrem geschätzten Beitrag zum Markenwert.

In den folgenden Kapiteln wird zunächst das Kampagnenformat und die damit verbundenen Daten für diese Analyse behandelt. Ein umfangreicher Teil der Recherche wird das Auswerten der Optionen für einen Ansatz des Sentimentanalysemodells sein. Darüber hinaus werden weitere Daten und Verarbeitungsmodule als Grundlage für das Modell erforderlich sein. Optionen werden ebenfalls anhand der Literatur ausgewertet und die Auswahl den vorhandenen Daten entsprechend modifiziert. Zur Datenverarbeitung wird der theoretische Hintergrund der gängigen Bearbeitungsschritte dargelegt und deren Anwendungen im Kontext dieser Arbeit dokumentiert, zwecks der Rechtfertigung der Datenmanipulation und der Verständlichkeit der Analyse. Nach der

Optimierung des Analysemodells werden die Ergebnisse stichprobenartig visualisiert und inhaltlich interpretiert. Hiermit soll beantwortet werden, inwiefern das Modell den Kampagnen- und Brandmanagern weiterhelfen kann. Es werden grundlegende Statistiken zu den erkannten Emotionen berechnet, Veränderungen von Emotionen über Kampagnenlaufzeiten analysiert, besonders aussagekräftige Texte erlesen, Korrelationen zwischen Emotionen und anderen Metriken visualisiert und Schwächen des Modells aufgedeckt und beurteilt.

Aus Gründen der leichteren Lesbarkeit wurde beim Schreiben der Arbeit auf die gleichzeitige Anwendung von weiblicher und männlicher Sprachformen verzichtet, jedoch gelten sämtliche Personenbezeichnungen beiderlei Geschlechter.

Die Daten, anhand derer die das Analysemodell erstellt und getestet wird, wurden vom *Institute for Interactive Marketing & Social Media* der *Wirtschaftsuniversität Wien* im Zuge einer Kooperation mit *Trnd* (Trnd 2021) zur Verfügung gestellt. Die personen- und kampagnenbezogenen Daten werden nicht veröffentlicht. Das Analysemodell *GerVADER* (Hutto und Gilbert 2014) ist die Grundlage für das letztendlich benutzte Analysemodell und ist mit einer MIT-Lizenz (GitHub 2021) lizenziert, die Nutzung und Modifizierung der Software ist demnach gestattet.

2 Kampagne

2.1 Format

Haenlein und Libai (2017) zählen die drei wichtigsten WOM-Kampagnenarten auf. Erstens, Referral-Kampagnen, bei denen Influencer mit ihren eigenen Referral-Codes ermutigt werden, ihre Follower zum Kauf zu bewegen (ein bekanntes Beispiel sind Affiliate-Marketing-Kampagnen) Zweitens, gibt es Empfehlungs-Kampagnen, die auf Empfehlungsseiten wie *TripAdvisor* und Social Media aufbauen. Drittens, und für diese Arbeit am relevantesten, sind Seeding-Kampagnen. Hierbei sollen Influencer das zu bewerbende Produkt erhalten und anschließend ihren Bekannten und Followern von ihren (hoffentlich positiven) Erfahrungen mit dem Produkt erzählen. Die Nutzung von Mikro-Influencern als Seed-Agents wird gerechtfertigt durch die Vermutung, dass Seed-Agents mit wenigen Followern, unter 100 000 (Zarei, et al. 2020), en masse einen höheren ROI als weniger Seed-Agents mit jeweils mehr Followern generieren, da die Beiträge ersterer bei ihren Betrachtern zu höherem Engagement führen und somit mehr Überzeugungskraft haben (Zarei, et al. 2020).

Um auf eine effektive Reichweite zu kommen, ist jedoch eine Vielzahl an Mikro- und Nano-Influencern notwendig. Hierfür gibt es auf Influencer-Marketing spezialisierte Agenturen. Teil der angebotenen Dienste von Influencer-Marketing-Agenturen ist die Führung von Kampagnen, bei denen die Agentur die Selektion von und Korrespondenz mit den Influencern übernimmt. Auf deren Plattformen können sich Nutzer als Influencer registrieren und stehen in Folge für Kampagnen zur Verfügung. Somit wird Marken der Aufwand des Managements der Influencer abgenommen.

2.2 Beispiel

Ein Beispiel für eine Influencer-Kampagne wäre das folgende: Eine Marke bucht eine Influencer-Marketing-Agentur für eine Seeding-Kampagne. Diese startet auf ihrer Homepage eine Ausschreibung für 2000 Kampagnenteilnehmer. Auf der Plattform registrierte Mikro-Influencer können sich nach Belieben für die Teilnahme bewerben. Diese werden im Idealfall nach demografischen und weiteren zielgruppen-relevanten Kriterien gewählt, um eine annähernd repräsentative Auswahl an Teilnehmern zu bilden. Die Kampagnenmanager richtet sich in mehreren Blogbeiträgen an die teilnehmenden Influencer und stellt jedem Teilnehmer ein oder mehrere Probe-Exemplare des Produkts

zur Verfügung. Die Teilnehmer werden im Gegenzug aufgefordert, sich umfassend mit dem Produkt auseinanderzusetzen, ihren Freunden und Familien von dem Produkt zu erzählen, sowie mit Kommentaren zu dem Produkt zu antworten und eine Stern-Bewertung von 0 bis 5 abzugeben. Teilnehmer werden auch aufgefordert, während der Kampagnenlaufzeit berichten, wenn sie Bekannten das Produkt weiterempfohlen haben. Die Kampagnenmanager fordern außerdem Engagement der Teilnehmer durch beispielsweise Aufforderungen, Fotos mit dem Produkt zu posten oder auf bestimmte Fragen dazu einzugehen. Das Unternehmen hat somit über die Mundpropaganda und das Kundenengagement hinaus Ratings von einer großen Stichprobe seiner Zielgruppe als Metriken für das betriebene WOM-Marketing. Die erhaltenen Blog-Posts und Kommentare können und sollten ebenfalls als Mittel zur Messung des Erfolgs der Kampagne herbeigezogen werden. In ihnen verbergen sich wertvolle, implizite Daten, also Informationen, die erst durch NLP quantifiziert und somit messbar gemacht werden können, wie inhaltliche Bedeutungen oder Emotionen.

2.3 Evaluierung und Optimierung

Die folgenden drei Erfolgsmetriken von Seeding-Kampagnen sind die Quantität werden von Marketing-Managern häufig genutzt: Die Menge an Interaktionen, die durch die Kampagne angeregt werden, Steigerungen der Markenwert bzw. -bekanntheit und Steigerungen der Verkaufszahlen (Haenlein und Libai 2017, 72). Erstere wird online durch leicht messbare Likes und dergleichen bemessen, offline sind Unterhaltungen über die Marke relevant, jedoch schwer messbar. Chintagunta, Gopinath und Venkataraman (2010) weisen nach, dass die Sentimente der Zielgruppe einen positiven Einfluss auf alle drei Metriken hat.

Die Auswahl der richtigen Teilnehmer für Seeding-Kampagnen kann seinen Erfolg stark beeinflussen. Eine gängige Vorgehensweise ist, Seed-Agents mit großem Einfluss auf möglichst viele Zielgruppen-Mitglieder zu wählen, sogenannte *Opinion Leader* (Dodds 2007, 4). Gewisse Studien argumentieren hingegen, für die Effektivität von Seeding-Kampagnen sei die Natur der Zielgruppe und ihre Anfälligkeit auf sozialen Einfluss wichtiger als die Breite des Einflusses der Seed-Agents (Dodds 2007, 1). Auch Haenlein und Libai (2017) betonen, dass Opinion Leader nicht unbedingt überragende Followerzahlen brauchen, um für effektives WOM sorgen zu können — nicht nur seien Mikro-Influencer weitaus billiger als oft prominente Mega-Influencer, sie hätten auch

weitaus mehr Unterhaltungen als Durchschnittsverbraucher und ihren Empfehlungen gingen verglichen mit denen von Mega-Influenzern mehr Leute nach. Die Aktivierung vielverdienender Seed-Agents, bestenfalls sogenannter *Revenue Leader*, birgt ebenfalls Vorteile. Nicht nur verfügen diese mit hoher Wahrscheinlichkeit über mehr hochwertige Kontakte, sie haben auch oft überdurchschnittlichen Einfluss auf Kontakte aufgrund ihrer von Kontakten vermuteten Expertise (Haenlein und Libai 2017, 80).

2.4 Daten

Für diese Arbeit werden Textdaten mehrerer Seeding-Kampagnen verwendet. *Trnd* (Trnd 2021) ist eine WOM-Plattform. Benutzer können sich darauf registrieren und für diverse Seeding-Kampagnen kostenlos bewerben. Jede Kampagne wird vor der Bewerbung umfassend auf *Trnd* beschrieben. Teilnehmern einer Kampagne werden ein oder mehrere, bezahlte Produkte zugeschickt, beziehungsweise die Kaufpreise ersetzt. In den Kampagnen führen die Kampagnenmanager von *Trnd* einen Blog, in dem Nachrichten, Fragen und Anweisungen an die Teilnehmer formuliert werden. Die Teilnehmer geben zu diesen Blogeinträgen Kommentare ab. Schließlich bewerten sie die Kampagnen mit einer Bewertung von 1 bis 5 und geben ihre Wahrscheinlichkeit, das Produkt weiterzuempfehlen, ab. Diese Daten wurden von *trnd.com* gescraped und in Tabellenform gespeichert. Von 573 gefundenen Kampagnen sind für 227 Resultate in Form von Bewertungen und Weiterempfehlungsquoten verfügbar, wovon für 208 Kampagnen Blogkommentare aus durchschnittlich 10 Blogeinträgen geladen wurden, wovon wiederum bei 106 Kampagnen die Teilnehmeranzahlen und Kampagnenmanager bekannt sind. In den Kampagnen haben im Schnitt jeweils 1800 Personen als Seed-Agents teilgenommen, nur ein Dutzend Kampagnen verzeichneten unter 100 Teilnehmern.

3 Sentimentanalyse

3.1 Theoretischer Hintergrund

NLP befasst sich mit dem Verarbeiten von unstrukturierten Daten in Form von Texten und den dadurch extrahierbaren Informationen. Strukturierte Daten und semi-strukturierte sind bereits geparkt, also in Form valider Datenformate, die für den Computer lesbar und durch minimale Umstrukturierung in relationalen Datenbankensystemen gespeichert werden können (Vishwakarma 2021). Unstrukturierte Daten hingegen, welche, laut Schätzungen, 80-90% der gesamten digitalen Datenmenge ausmachen (Davis 2019), sind ohne die Hilfe der richtigen Werkzeuge nicht analysierbar. Beispiele sind Fotos, Videos, Audio und natürlich die natürliche Sprache (Marr 2019), gesprochen oder transkribiert. Ihre Semantik, also die Bedeutung der ausgedrückten Worte, ist vom Computer nicht ersichtlich. Das Ziel von NLP ist, eine maschinell lesbare Struktur aus dem komplexen, unstrukturierten Gewirr aus Informationen zu abstrahieren (Liu 2020, 16). Somit können Fließtexte als Datensätze für statistische Berechnungen und automatisierte Verarbeitungen genutzt werden.

Sentimentanalysen, auch *Meinungsanalyse* (Siegel, et al. 2017, 67) oder *Opinion Mining* (Wolfgruber 2015, 18) genannt, „extrahier[en] subjektive Informationen aus Texten und klassifiziert Wörter, Phrasen, Texte und Dokumente nach den jeweils enthaltenen Meinungen und Stimmungen [der Autoren]“ (Wolfgruber 2015, 18). Sie sind somit ein semantisches Analyseproblem, jedoch sind sie weitaus konzentrierter als grundsätzliche Informationsgewinnungsprobleme, bei denen nicht nur Sentimente extrahiert werden. Texte müssen also nicht vollständig verstanden werden, lediglich Teilaspekte (Liu 2020, 14).

Meinungen und Sentimente teilen eine essenzielle Eigenschaft: Subjektivität. Verschiedene Erfahrungen und wiederum verschiedene Wahrnehmungen und Verarbeitungen dieser Erfahrungen führen zu internen Zuständen, die nicht eindeutig prüfbar sind. Statt Sentimente von Meinungen abzugrenzen, nutzt Lui (2020) Meinungen als Überbegriff für Einstellungen, Einschätzungen und eben Sentimenten, die sie als interne Emotion beschreibt, die die Meinung bildet. Sentimentanalysen können diese, je nach Analysemodell, in verschiedenen Komplexitäten erkennen. Simple Varianten erkennen nur Sentimente, ohne Subjekt und Objekt festzulegen. Komplexere Modelle

können Entitäten wie Subjekt, Objekt, Objektaspekte und Zeitpunkt in einem Text unterscheiden (Liu 2020, 18).

Da Sentimente, wie oben beschrieben, mehrheitlich subjektive Zustände sind, sollten Sentimentanalysemodelle in der Lage sein, subjektive Aussagen von objektiven abgrenzen zu können. Sie untersuchen Texte auf Worte und Phrasen, anhand derer das Sentiment des umfassenden Textes erkannt werden können. Diese Sentiment-Träger (Wolfgruber 2015, 24) haben jeweils eine gewisse Valenz (auch Polarität genannt). Diese drückt aus, ob die grundlegende Bedeutung eines Wortes oder einer Phrase positiv, negativ oder, falls weder noch, neutral (Wolfgruber 2015, 22). Neutrale Worte können zwar semantisch bedeutungsvoll sein, tragen aber nichts zur Emotion der Aussage bei. Das Sentiment der Sentiment-Träger hat eine gewisse Intensität, es gibt selbstverständlich verschieden stark emotionale Ausdrücke. Besonders gut oder schlecht gelaunte Menschen verwenden emotional geladene Worte, um ihren Gefühlszustand verständlich zu machen. Quantifiziert man die Intensität von Sentiment-Trägern, so kann die Valenz als Vorzeichen verwendet werden (+ für positive Sentimente und – für negative) (Tabelle 1). Synonyme von Sentiment-Trägern haben die gleiche Valenz und Intensität, Antonyme haben die jeweils entgegengesetzte Valenz (Wolfgruber 2015, 22).

der Film ist	neutral	0.0000
der Film ist gut	positive	0.4767
der Film ist schlecht	negative	-0.4588
der Film ist toll	positive	0.4939
der Film ist furchtbar	negative	-0.5106
der Film ist ok	positive	0.3818
der Film ist mittelmäßig	neutral	0.0000

Tabelle 1: positive und negative Worte

Der Kontext eines Ausdruckes kann seine Sentiment-Intensität verändern, „ohne dabei die Valenz zu verändern“ (Wolfgruber 2015, 34). Partikel wie „sehr“ oder „wenig“ dienen hierbei als Modifikatoren der Intensität (Tabelle 2).

das Hotel ist ...	neutral	0.0000
das Hotel ist schön	positive	0.5106
das Hotel ist sehr schön	positive	0.5563
das Hotel ist hässlich	negative	-0.5423
das Hotel ist sehr hässlich	negative	-0.5849
das Hotel ist schmutzig	negative	-0.4404
das Hotel ist etwas schmutzig	negative	-0.3832

Tabelle 2: Sentiment-Modifikatoren

Auch die Valenz eines Ausdruckes kann kontextbedingt verändert werden. Negationen sowie sarkastische und ironische Aussagen können die Valenz des Sentiments ebenso beeinflussen, in manchen Methoden sogar gänzlich umkehren (Wolfgruber 2015, 23). Negationen können durch lexikalische Regeln erkannt werden. Sarkasmus zu erkennen ist jedoch noch weitaus schwieriger, da es oft keine sichtbaren Unterschiede zwischen in geschriebenen, sarkastischen und ehrlichen Aussagen gibt. Zum Beispiel werden bei „Wortironien“ einzelne Worte absichtlich unangemessen verwendet (Wolfgruber 2015, 27, zitiert nach Hartung, 1998, 70), um den Zuhörer oder Leser darauf hinzuweisen, dass das polare Gegenteil der Aussage gemeint ist. Idiomatiche Aussagen bilden eine weitere Komplikation: Hierbei werden in Redewendungen implizite Semantiken durch außergewöhnliche Kombinationen von Worten verschleiert. Diese Redewendungen können sich regional unterscheiden. Nicht nur erschweren dem Modell unbekannte Redewendungen die semantische Analyse eines Textes, bei Sentimentanalysen können Modellen emotional geladene Aussagen entgehen (Tabelle 3).

Das Essen war unter jeder Sau! neutral 0.0

Tabelle 3: idiomatiche Aussage

Um Sentimentanalysen noch weiter zu verkomplizieren, kommen domänenabhängige Modifizierungen von Sentiment-Trägern ins Spiel. Wolfgruber (2017, 36) nennt das Adjektiv „alt“ als Beispiel: Domänenunabhängig ist es neutral und beschreibt lediglich das relative Alter, in der Domäne von Hotelbewertungen kann „ein altes Hotel“ jedoch im positiven Sinne für Tradition und Urigkeit stehen. In der Domäne von Essen ist „altes Essen“ wiederum nichts Wünschenswertes und demnach negativ. Solche Veränderungen beachtende Sentimentanalysen müssen also einen gewissen Grad an Domänenwissen haben und Sentiment-Träger den richtigen Aspekten zuweisen können, um keine ungewollten Fehler zu machen.

Sentimentanalysemodelle sollten auch Interjektionen, also wortähnliche, gefühlsausdrückende Laute, erkennen und mit richtigen Sentimenten annotieren können (Wolfgruber 2015, 32). Ebenso können Emoticons und Emojis Hinweise auf Sentimente des Autors sein (Tabelle 4). Emoticons sind bildhafte Kombinationen aus ASCII-Zeichen, die auf jeder Tastatur zu finden sind. Emojis hingegen sind Grafiken, die durch Unicode-Kodierungen geladen werden. Dieser Unicode-Standard muss von Geräten unterstützt werden, damit Emojis angezeigt werden können, und nicht das Ausweich-Emoji ☐ auch „Tofu“ genannt (Khemka 2020).

😊	positive	0.4588
😍	positive	0.6124
😞	negative	-0.4588
😡	negative	-0.7184
😐	neutral	0.0000
😶	neutral	0.0000
:)	positive	0.4588
:(negative	-0.4404
>:(negative	-0.5719
<3	positive	0.4404

Tabelle 4: Emoji-Sentimente

Laut Liu (2020, 9) werden Sentimentanalysen vorwiegend auf Dokument-, Satz- oder Aspektebene durchgeführt. Sentimentanalysen auf Dokumentebene stellen fest, ob ein Dokument insgesamt ein gewisses Sentiment trägt. Ein häufiges Beispiel für Sentimentanalysen auf Dokumentebene sind Produktrezensionen, wo die Sentimente aller Sentiment-Träger im Dokument zu einer Meinung zusammengefasst werden. Die Algorithmen von Sentimentanalysen auf Satz- und Dokumentebene sind sehr ähnlich, es gibt neben dem Trennen in einzelne Sätze keine fundamentalen Unterschiede in der Verarbeitung (Yousef, Medhat und Mohamed 2014, 2). Auf Satzebene können zudem neutrale Sätze erkannt werden, welche laut Liu (2020, 9) faktisch und objektiv sind. Hung und Lin (2013) hingegen argumentieren, dass neutrale Sätze im Kontext von emotionalen Sätzen ebenfalls emotionale Tendenzen haben. Darauf wird im Kapitel zur Datenverarbeitung näher eingegangen. Auf Aspektebene annotieren Sentimentanalysen Meinungsausdrücke als Meinungsträger, Meinungsobjekt und gemeintem Aspekt des Objektes, sowie alle anderen relevanten Aspekte.

3.2 Ansätze

Medhat et al. (2014, 6) geben einen Überblick über die drei Ansätze von Sentimentanalysemodellen. Diese sind maschinelle Lernansätze, lexikon-basierte Ansätze und wissensbasierte Ansätze.

3.2.1 Maschinelles Lernen

Beim maschinellen Lernen für Klassifikationsmodelle werden ML-Algorithmen angewandt, um Muster von bestehenden Datenpunkten zu lernen und auf nicht klassifizierte Datenpunkte anzuwenden, um diese korrekt klassifizieren zu können. ML-Algorithmen sind statistische Methoden, somit „wird [im Bereich von NLP] der Sinn des Textes nicht erschlossen, was gerade bei komplexeren Aussagen von großem Nachteil sein kann.“ (Wolfgruber 2015, 48) Maschinelle Lernverfahren werden unter anderen in überwachte und unüberwachte Lernmethoden unterteilt (Khan, et al. 2016, 98). Unterteilungen können von Studie zu Studie variieren. Bei überwachtem maschinellern Lernen „werden fehlende Informationen auf Grund von Beobachtungen [existierender Informationen] vorherzusagen“ (Wolfgruber 2015, 48). Somit werden grammatikalische Regeln und Wortbedeutungen automatisch berücksichtigt, wenn auch nicht verstanden (Khan, et al. 2016). Beispiele für Algorithmen sind Entscheidungsbäume, Support Vector Machines und Deep Learning. Um ihre eigenen Parameter selbst korrigieren zu können, um ein möglichst genaues Modell zu bilden, müssen die Trainingsdaten alle mit den korrekten Klassen markiert sein (Wójcik 2019). Unüberwachte Lernmodelle werden nicht auf möglichst korrekte Klassifizierungen trainiert, sondern vergleichen Aspekte von Datenpunkten, um sie in homogene Gruppen zu teilen und Strukturen sichtbar zu machen. Die bekanntesten Algorithmen sind Clustering-Algorithmen (Wolfgruber 2015, 50). Durch solche Ansätze sind Sentimentanalysen auf Dokument-, Satz- und bei sogar Aspektebene möglich.

3.2.2 Lexikon-basierte Ansätze

Eine weitere Form von Sentimentanalysen nutzt Sentiment-Lexika, welche für alle möglichen Sentiment-Träger zumindest die jeweilige Valenz und in vielen Fällen auch Sentiment-Intensität und Wortart annotieren (Liu 2020, 59). Auf Satz- oder Dokumentebene können die Sentimente summiert werden, eine positive oder negative Summe deuten dann auf einen insgesamt positiven oder negativen Satz hin, während eine Summe gleich 0 einen neutralen Satz prägt.

Wie granular solche Ansätze die Einflüsse jedes Wortes auf das gesamte Sentiment auswirken, kann variieren und sich somit auf die Genauigkeit des Modells auswirken. In der simpelsten Form werden Sentiment-Träger nur mit Valenzen, also ohne Intensitäten, annotiert. In Hu und Liu (2004) kehren Negationen diese Valenzen vollständig um. Aus „gut“, mit einer Valenz von 1, wird somit „nicht gut“, mit einer Valenz mit -1 (Liu 2020, 59). Die verschiedenen, emotionalen Intensitäten von Sentiment-Trägern motivieren genauere Annotierungen. Diese können zum Beispiel kategorisch, von 5 (positiv) bis -5 (negativ) reichen (Taboada et al. 2011), oder kontinuierlich annotiert sein, von 1 (positiv) bis -1 (negativ) (Loria 2020). Hier bei Negationen Valenzen ganz umzukehren (Tabelle 5), ist jedoch laut Liu (2020, 60) unklug. Wenn „perfekt“ im *GerVADER*-Lexikon (Tymann, et al. 2020) ein Sentiment von rund 0,64 hat, würde eine Umkehrung zu einem Sentiment von -0,64 führen, was nicht sinnvoll wäre – „nicht perfekt“ umschreibt nicht unbedingt für ein negatives Sentiment.

das Essen war ...	neutral	0.0000
das Essen war gut	positive	0.4767
das Essen war nicht gut	negative	-0.4767
das Essen war schlecht	negative	-0.4588
das Essen war nicht schlecht	positive	0.4588

Tabelle 5: Negationen durch Valenzumkehrung

Liu (2020, 60) schlägt darum vor, die Sentimente von negierten Worten um einen fixen Betrag in Richtung der entgegengesetzten Valenz zu modifizieren (Tabelle 6), anstatt sie mit -1 zu multiplizieren.

das Essen war ...	neutral	0.0000
das Essen war perfekt	positive	0.6369
das Essen war nicht perfekt	negative	0.2369
das Essen war schrecklich	negative	-0.5106
das Essen war nicht schrecklich	positive	-0.1106

Tabelle 6: Negationen durch Intensität-Subtraktion

Das vereinfacht auch die Integration von weiteren Regeln, wie zum Beispiel Modifikatoren, die nur die Intensität beeinflussen, ohne die Valenz zu verändern. Diese können die Intensität verstärken oder abschwächen. Bei kontinuierlichen Sentiment-Annotierungen ist somit eine granulare Unterscheidung in der Auswirkung jedes Wortes möglich. Solche Modifikatoren können neben Adjektiven und Adverbien auch durch

andere semantische Muster gebildet werden, wie der Nutzung von Großbuchstaben, Rufzeichen und Fragezeichen (K. Tyman, et al. 2021). Auch die Nutzung von „aber“ kann die Veränderung von Intensitäten anregen.

Sentiment-Lexika zu erstellen ist eine komplexe und zeitaufwändige Aufgabe, die äußerste Sorgfalt erfordert, um genaue Analysen zu ermöglichen. Ein strikt manuelles Annotieren, etwas durch Crowdsourcing oder von Experten (Fehle, Schmidt und Wolff 2021, 1), Auflisten und Annotieren aller Sentiment-Träger birgt obendrein ein großes Bias-Risiko, da Beteiligte untereinander und zu verschiedenen Zeitpunkten unterschiedliche Sentimente schätzen könnten (Hung und Lin 2013, 48). Sentiment-Schätzer sind sich schätzungsweise in 60–80 Prozent der Fälle über das Sentiment eines Textes einig (Ribeiro, et al. 2016, 13). Darum dienen manuelle Arbeitsschritte meist als Komplemente zu zuvor maschinell erstellten Lexika-Erstellungsprozessen (Wolfgruber 2015, 40). Hung und Lin (2013) schlagen vor, mit einigen Seed-Worten und ihren Valenzen zu beginnen, die Liste iterativ durch Synonyme und Antonyme aus Onlinelexika wie *Wordnet* zu erweitern und schließlich händisch zu kontrollieren und zu erweitern. Das Lexikon *SentiWordNet* wurde beispielsweise so erstellt (Esuli und Sebastiani 2010). *SO-CAL* (Taboada, et al. 2011), kurz für *Semantic Orientation CALculator*, ist ein Sentimentanalysemodell mit einem eigenen Lexikon. Es verwendet eine eigene Subjektivitätsmetrik, die Semantische Orientierung, um die Sentimente gewisser Worte zu errechnen (Taboada, et al. 2011, 268).

3.2.3 Wissensbasierter Ansatz

Wolfgruber (2015) nutzt einen wissensbasierten Ansatz, um Sentimente in Hotelbewertungen auf Aspektenebene zu analysieren. Das erfordert das maschinelle Analysieren und Erkennen lokaler Grammatiken. „[Diese] sind nicht als Grammatik [...] einer gesamten Sprache zu verstehen, sondern [...] erfassen syntaktische und lexikalische Phänomene einer speziellen Domäne.“ (Wolfgruber 2015, 64) Das dafür erforderliche Sentiment-Lexikon wurde auf den Texten eines domänenrelevanten Korpus, statt einem generischen Wörterbuch, aufgebaut. Ein solches Modell erfordert umfassende Domänenkenntnis, da relevante Aspekte für die Analyse im Vorhinein definiert und die resultierenden syntaktischen Strukturen händisch kontrolliert werden müssen. Im Gegenzug ist das fertige Modell jedoch in der Lage, innerhalb der eigenen Domäne, detaillierte Angaben zu Meinungsträgern und Meinungsobjekten zu machen.

3.3 Ansatz dieser Arbeit

Die verschiedenen Ansätze für Sentimentanalysen sind umfassend erforscht worden. Forscher haben stark variierend genaue Modelle erstellt, es lässt sich jedoch beobachten, dass gewisse Ansätze Texte genauer klassifizieren als andere. Durch maschinelles Lernen können die genauesten Modelle erzielt werden (K. Lee 2021), ohne Domänen im Detail studieren zu müssen. Hierbei kann zwischen Genauigkeit des Modells, die beim *Deep Learning* am größten sei, und Geschwindigkeit der Implementierung, bei der *Naive Bayes* punktet, abgewogen werden. Für schnelle und einfach anzuwendende Lösungen, wofür Genauigkeit eingebüßt werden kann, sind lexikon-basierte Modelle eine gute Wahl (K. Lee 2021). Auch Hartmann et al (2019) deuten auf genauere Ergebnisse durch maschinelle Lernverfahren wie *Random Forests* und *Naive Bayes* hin. Hartmann et al. (2019) erreichten mit fünf verschiedenen, lexikon-basierten Ansätzen Genauigkeiten von bis zu 68%, mit unüberwachten, maschinellen Lernmethoden erreichten Genauigkeiten zwischen 43% und 94%, maschinelle Lernverfahren waren jedoch, mit wenigen Ausnahmen, konsistent akkurater als lexikon-basierte Verfahren. Wolfgruber (2015) erzielte in ihrer Dissertation mit einem wissensbasierten Ansatz mit Lokalen Grammatiken eine Genauigkeit von rund 97%, punktete also mit hoher Genauigkeit, der Ansatz erforderte beträchtliche Kenntnis des Untersuchungsgegenstandes.

Aufgrund der Daten der *Trnd*-Kampagnen, wird hier ein lexikon-basiertes Modell angewandt. Sentimentanalysen mittels überwachter Lernverfahren benötigen Trainingsdaten mit den korrekten Sentimentwerte als Prüfvariable, was bei den beiliegenden Daten nicht verfügbar ist. Lexikon-basierte Modelle hingegen brauchen keine Trainingsdaten. Sie werden können mittels mit den richtigen Sentimenten markierten Daten optimiert werden, sind jedoch auch ohne Trainingsdaten direkt anwendbar (Liu 2020, 77). Unüberwachten Lernmodellen werden im Gegensatz zu überwachten Lernmodellen keine Vorurteile zu Datenstrukturen antrainiert, sondern erkennen Strukturen in den Daten jedes Mal aufs Neue (Wójcik, Towards Data Science 2019). Jedoch sind diese, sobald in einer gewissen Domäne (oder mehreren) optimiert, nicht oder schlecht für Daten aus anderen Domänen einsetzbar (Liu 2020, 77). Für diese Arbeit liegen Daten zu mehreren heterogenen Kampagnendomänen vor, die jeweils eigene unüberwachte Modelle bräuchten. Wissensbasierte Ansätze würden pro Modell umfassende Domänenkenntnis erfordern, damit die lexikalischen Regeln genau dem Kontext entsprechen können (Schuller und Knaup 2011, 457). Dafür könnte hierbei ein

Sentimentanalysemodell auf Aspektenebene erstellt werden, was zu besonders genauem und dazu noch intuitivem Verständnis der Texte führen würde. Aufgrund der Vielseitigkeit der Kampagnen von Agenturen wie *Trnd* und ihrer verschiedenen Domänen, ist ein wissensbasierter Ansatz mit lokalen Grammatiken nicht der effizienteste Ansatz, es sei denn, die Agentur würde sich auf gewisse Domänen spezialisieren oder viele, immer wiederkehrende Kategorien von Kampagnen leiten.

3.4 Auswahl des Sentiment-Lexikons

Alle zuvor genannten Ansätze können von der Berücksichtigung vorher geschriebener Sentiment-Lexika profitieren. Von den, von Hartmann et al. (2019) getesteten, Lexikon-basierten Sentimentanalysemodellen hat das Analysemodell *VADER* (*Valence Aware Dictionary and Sentiment Reasoner*) von Hutto und Gilbert (2014) für beinahe alle Anwendungsfälle die höchste Genauigkeit erzielt. Ein Rangvergleich von englischsprachigen Sentiment-Lexika sollte im Angesicht der deutschsprachigen Kampagnen von *Trnd* mit Vorsicht betrachtet werden, da die deutsche Variante zu *VADER*, *GerVADER* (K. Tymann, et al. 2021), zwar die gleichen Analyseschritte durchgeht, jedoch in der Literatur kaum mit anderen deutschen Sentimentanalysemodellvarianten verglichen wurde. Bei einem Ergebnisvergleich von *Vader* und *GerVADER* anhand eines übersetzten Korpus wurde gezeigt, dass die Genauigkeit von *Vader* die von *GerVADER* knapp übersteigt, die Autoren begründen das jedoch mit externen Einflüssen durch die automatische Übersetzung (Tymann, Steinkamp, et al. 2020). *VADER* und *GerVADER* zeichnen sich durch ihre Vielzahl an kontextbeachtenden Regeln aus. Es werden beispielsweise Negationen und Intensitätsveränderungen durch Adjektive berücksichtigt. Die Erstellung von *GerVADER* wurde mit speziellem Fokus auf Social-Media-Texte durchgeführt. *GerVADER* ist somit nicht nur auf die gängige Sprache in Mikroblogs optimiert, es ist auch in der Lage, Emoticons und Emojis richtig zu deuten. Das sind beides Eigenschaften, die die Texte in den behandelten Daten teilen.

Ribeiro et al. (2016) vergleichen 24 verschiedene, lexikon-basierte Sentimentanalysemodelle. *VADER* hatte konsistent bemerkenswert gute Ergebnisse, marginal genauer war *SentiStrength* (SentiStrength 2021), welches ebenfalls auf Social-Media-Texten aufgebaut wurde. Fehle, Schmidt und Wolff (2021) vergleichen ebenfalls die Genauigkeit 19 verschiedenen Sentiment-Lexika in verschiedenen Kontexten,

inklusive Social-Media. Die besten Ergebnisse im Social-Media-Kontext lieferten die Sentiment-Lexika *GermanPolarityClues* (Waltinger 2010), *SentiWS* (Remus, Quasthoff und Heyer 2010) und jenem von Rauh (2018), welches jedoch einen Fokus auf politische Sprache legt. *GermanPolarityClues* wurde in einer anderen Studie von *SentiWS* übertroffen (Waltinger 2010, 208). Aufgrund der Performance von *SentiWS* bei der Klassifizierung von Texten in Social-Media-Kontexten wird das darauf aufgebaute Sentimentanalysemodul *GerVADER* (K. Tymann, et al. 2021, 1).

4 Datenverarbeitung

Es gibt zahlreiche Techniken, natürliche Sprachtexte zu bearbeiten, um Analysen nicht nur zu erleichtern, sondern auch ihre Genauigkeit zu erhöhen. Im Folgenden werden die Schritte der Textvorbereitung für Sentimentanalysen im Detail erklärt, sowie ihre Anwendung durch das Modul *GerVADER* anhand des Kampagnenbeispiel dokumentiert. Manche Schritte, die in der Forschung als wichtig angepriesen werden, sind in *GerVADER* nicht inkludiert. Darum wird *GerVADER* im Zuge dieser Arbeit durch einige dieser domänenunabhängigen, lexikalischen Regeln ergänzt.

4.1 Tokenisierung

Als erstes werden ganze Sätze in Arrays an Tokens getrennt. Ein Token ist eine bedeutungsvolle Texteinheit wie ein Wort oder Emoji. Durch ihre Trennung in Sequenzen können sie leichter sowohl als solche als auch einzeln bearbeitet und entfernt werden. Gewisse Sentiment-Lexika, beispielsweise *SentiWS*, entfernen im Zuge der Tokenisierung Satzzeichen, um nur lexikalische Worte zu behalten (Remus, Quasthoff und Heyer 2010). Gewisse Sentiment-Indikatoren würden dadurch verloren gehen, beispielsweise Emoticons und Rufzeichen. Satzzeichen gänzlich zu belassen würde jedoch dazu führen, dass der Algorithmus „sehr,“ und „sehr“ als verschiedene Worte betrachten und erstere darum nicht bewerten würde, wenn es nicht in den Lexika annotiert ist. In *GerVADER* werden für jedes Dokument ein Python-Dictionary mit allen Tokens und jeweils allen Token-Satzzeichen-Permutationen erstellt und ihre Schlüssel als Tokens genutzt. Durch diesen Ansatz wird die mögliche Tokenmenge pro Text zwar sehr groß, die Rechenzeit wird dadurch jedoch nicht verlangsamt, da das Nachschlagen eines Tokens in einem Dictionary eine lineare Zeitkomplexität hat (Terasa 2020), also unabhängig von Dictionary-Größe ist.

4.2 Part-of-Speech-Tagging

Ein Nachteil von Sentimentanalysen auf Dokument- und Satzebene, das fehlende Verständnis davon, welche Worte im Text Meinungsträger und Meinungsobjekt sind, kann durch Part-of-Speech-Tagging (POS-Tagging) zum Teil umgangen werden (Sanketh 2020). POS-Tagging-Module analysieren Sätze und annotieren Worte als Satzteile wie Adjektive, Nomen, Präpositionen und, bei besonders ambitionierten Modellen, sogar als Subjekt oder Objekt. Somit können ambige Bedeutungen gleich

geschriebener Worte mit unterschiedlichen POS-Tags getrennt werden (Fehle, Schmidt und Wolff 2021, 0). *GerVADER* hat keinen POS-Tagging-Algorithmus eingebaut. Zwar beinhaltet *SentiWS*, auf dem *GerVADER*'s Lexikon basiert, POS-Tags für alle Wörter, das schlussendlich benutzte Lexikon nutzt diese jedoch nicht.

4.3 Stemming und Lemmatisieren

Stemming und Lemmatisieren sind Techniken, die die Vielfalt emotional gleichwertiger Wörter von gleicher Bedeutung zu reduzieren. Beim Stemming werden Worte auf ihren gemeinsamen Wortstamm gekürzt (Fehle, Schmidt und Wolff 2021). Somit werden zum Beispiel die Worte „eating“ und „eats“ auf den Wortstamm „eat“ reduziert. Solche Transformationen basieren auf vordefinierten Regeln und sind anfällig auf Ausnahmefälle und sprachliche Komplexität (Manning, Raghavan und Schütze 2009, 32). Darum verlassen sich Forscher tendenziell auf Lemmatisierungsmethoden, um Wortvarianten zu dezimieren. Hierbei werden Wortvarianten durch ihr lexikalisches Lemma, also ihrer Basis- oder Wörterbuchform, ersetzt (Manning, Raghavan und Schütze 2009, 32). Lemmatisierungsmodule sind in der Lage, Konjugationen und Deklinationen aus morphologisch komplexen Sprachen, wie der deutschen, umzuformen (Fehle, Schmidt und Wolff 2021). Um die Notwendigkeit und gegebenenfalls Fehlerwahrscheinlichkeit von Lemmatisierungsmodulen zu umgehen, inkludieren manche Sentiment-Lexika möglichst viele Wortvariationen, so auch *GerVADER*. Da lexikon-basierte Ansätze grundsätzlich sehr schnell sind, ist die längere Rechenzeit aufgrund der Lexikongröße durchaus tragbar und obendrein womöglich kürzer als die Rechenzeit Lemmatisierungsmoduls.

4.4 Groß-Kleinschreibung

Anders als bei in der englischen Sprache, in der nur Satzanfänge und Eigennamen großgeschrieben werden, sind Nomen in der deutschen Sprache generell groß zu schreiben. Das kann zu möglichen Fehlentscheidungen bei der Sentimentzuweisung führen. Das Verb „würde“ hat kein Sentiment, das Nomen „Würde“ hingegen könnte ein positives Sentiment haben (Fehle, Schmidt und Wolff 2021). Steht das Wort nun am Satzanfang, müssen kontextbeachtende Regeln die Entscheidung treffen. Die strikte Prüfung korrekter Groß-Kleinschreibung verliert an Relevanz im informellen Social-Media-Bereich, in dem Rechtschreibfehler und falsche Groß-Kleinschreibungen gang und gäbe sind. *GerVADER* prüft zudem, ob gewisse, und nicht alle oder keine, Sentiment-

Wörter in einem Satz gänzlich in Großbuchstaben geschrieben sind und verstärkt die Intensität solcher Vorkommnisse (Tabelle 7).

Ich fand das Produkt beschissen.	negative	-0.5267
Ich fand das Produkt BESCHISSEN.	negative	-0.6289

Tabelle 7: Sentiment-Modifikation durch Großbuchstaben

4.5 Stoppwörter

Stoppwörter sind kontextunabhängig häufig vorkommende Wörter, die an sich keine Sentiment-Träger sind. Der Unterschied zu anderen neutralen Worten ist, dass sie lediglich sprachliche Hilfsörter sind, die selbst keinen eigenen Inhalt haben, also Präpositionen, Artikel und Pronomen. Diese werden von vielen Sentimentanalysemodellen markiert, indem alle Wörter eines Textes mit einer vordefinierten Liste an Stoppwörtern abgeglichen werden, und pauschal entfernt, da sie erfahrungsgemäß die Genauigkeit von maschinellen Lernmodellen verringern (Fehle, Schmidt und Wolff 2021). Bei lexikon-basierten Modellen besteht jedoch keinen Vorteil darin, Stoppwörter zu entfernen, da nicht-sentimentale Wörter ohne Bedeutung ohnehin nicht in Sentiment-Lexika annotiert sind.

4.6 Kontextbedingte Einflüsse

Worte werden auf Verneinung untersucht indem in einem gewissen „Fenster“, also einer Reichweite von Tokens, nach verneinenden Wörtern gesucht wird. Fehle, Schmidt und Wolff (2021) haben die optimale „Fenstergröße“ untersucht und ob beide Richtungen erwogen werden sollten. Die Studie ergab, dass bei deutschen Texten eine beidseitige Fenstergröße von 4 die besten Ergebnisse erzielte. Hutto und Gilbert (2014, 221) geben aber an, 90% der Sentiment-Umkehrungen durch Negationen mit *Vader* erkennen zu können. Zusätzlich inkludiert auch Ausnahmefälle: „Never [...] beautiful“ würde ohne diese Aufnahmen die darauf folgenden 4 Worte negieren, „never so beautiful“ wäre jedoch ein Hinweis auf eine Sentimentverstärkung, anstatt einer Verneinung. *GerVADER* nach Tyman et al. (2021) hat keine Übersetzungen für diese Spezialfälle, diese wurden im Zuge dieser Arbeit manuell eingefügt (Tabelle 8). Die Berücksichtigung von Negationen hat laut Fehle, Schmidt und Wolff (2021) in ihren Tests zu Verbesserungen der Modelle geführt.

Das Hotel war schön.	positive	0.5106
Das Hotel war noch nie schön.	negative	-0.4023
Das Hotel war noch nie so schön.	positive	0.7229
Das Hotel war noch nie dermaßen schön.	positive	0.5960
Das Hotel war noch nie derart schön.	positive	0.6802

Tabelle 8: Nutzung von „nie“ und seine Auswirkungen auf Sentimente

Adjektive und Adverbien werden, ähnlich wie Negationen, innerhalb eines gewissen Wortabstandes zum betrachteten Sentiment-Träger, gesucht. Ihr Multiplikator wird aus einem vordefinierten Python-Dictionary an Boostern nachgeschlagen und mit dem Sentimentwert des betrachteten Sentiment-Trägers multipliziert. Die Berücksichtigung von Sentiment-tragenden Adjektiven, Adverbien in den Tests von Fehle, Schmidt und Wolff (2021) hatte ebenfalls positive Auswirkungen auf die durchschnittliche Genauigkeit der Modelle.

VADER und *GerVADER* beachten noch weitere Spezialfälle, die Sentimente beeinflussen. Dazu gehört die Nutzung von „but“, bzw. „aber“. Diese signalisieren, dass die vorgestellten Sentimente weniger signifikant sind, die nachfolgenden jedoch umso signifikanter (Hutto und Gilbert 2014, 221). Somit kann ein fundamentales Problem von Sentimentanalysen auf Satzebene teils umgangen werden: Die riskante Annahme, dass ein Satz nur jeweils ein Sentiment trägt (Liu 2020, 77). Darüber hinaus sucht *VADER* explizit nach „at least“ vor Sentiment-Trägern, um diese ebenfalls mit einem verringernden Multiplikator zu versehen. *GerVADER* hat ursprünglich keine Übersetzung für „but“ und „at least“ integriert, diese wurden ebenfalls händisch im Zuge dieser Arbeit in das Modul eingebaut (Tabelle 9).

Das Essen war gut.	positive	0.4767
Der Service war schlecht.	negative	-0.4588
Das Essen war gut und der Service war schlecht.	neutral	0.0258
Das Essen war gut aber der Service war schlecht.	negative	-0.4497
Service war schlecht, immerhin war das Essen gut.	neutral	-0.0498

Tabelle 9: Nutzung von „aber“ und seine Auswirkung auf Sentimente

Emojis und Emoticons werden bei der Anwendung kontextueller Modifikatoren explizit ihrer Intensität belassen, unter der Annahme, dass Emojis und Emoticons das gesamte Sentiment eines Satzes repräsentieren.

4.7 Sentimentwerte der Sätze

Nach der Optimierung von *GerVADER* wurden die Kommentare Satz für Satz auf Sentimente überprüft und mit erkannten Sentimentwerten annotiert. Die satzweise Verarbeitung war notwendig, um die oben beschriebenen Kontextregeln nicht falsch anzuwenden. Bei einem Eingabetext, wie „Es stört mich nicht. Sehr schön.“, würde *GerVADER* das „nicht“ allen Tokens davor und danach zuschreiben, dabei endet der semantische Einfluss von Negationen im Satz, in denen sie ausgedrückt werden. *GerVADER* errechnet für jedes Dokument einen Sentimentwert, indem die Werte jedes Tokens in einem Lexikon nachgeschlagen und danach summiert werden. Dabei werden alle kontextuellen Regeln angewandt und die Werte dementsprechend verändert und schlussendlich summiert.

Hung und Lin (2013) haben in ihrer Studie erforscht, ob objektive Worte, die häufig in sentimental Sätzen vorkommen, so stark mit diesen Sentimenten assoziiert werden, dass sie selber mit gewisser Wahrscheinlichkeit das Sentiment ihrer Sätze mitbeeinflussen, obwohl sie selbst objektiv und daher ohne eigenständigem Sentiment sind. Durch die Berücksichtigung der Vorkommnisse von objektiven Worten in Sentimentsätzen, konnten die Autoren die Genauigkeit ihres lexikon-basierten Sentimentanalysemodells signifikant erhöhen. Aufgrund dessen wurde ihre Methode in dieser Analyse ebenfalls angewandt.

Nachdem die Sentimentsummen für jeden Satz errechnet wurden, wurden die Häufigkeiten aller Nicht-Sentiment-Trägerworte in jeweils positiven, negativen und neutralen Sätzen berechnet. Es wurden hierbei nur Sentimentsätze mit einer normalisierten Sentiment-Intensität über dem, von Hung und Lin (2013) entdeckten, optimalen Schwellenwert von 0,6 mitgezählt. Es ließen sich nicht nur objektive Worte herausfiltern, sondern auch einige, eindeutig sentimentale Worte, die entweder nicht in den Lexika enthalten oder einfach falsch geschrieben waren, zum Beispiel „suuuuperlecker“ oder „buhu“. Im nächsten Schritt wurden die Wahrscheinlichkeiten jedes Wortes, in einem positiven oder negativen Satz vorzukommen, errechnet. Die höchste dieser Wahrscheinlichkeiten wird als Sentiment-Tendenz des Wortes genutzt, mit negativem Vorzeichen bei höherer Häufigkeit in negativen Sätzen. Hung und Lin (2013) haben ihren Algorithmus an einem Datensatz mit gleich vielen positiven und negativen Sentimentsätzen getestet. Um die ungleiche Verteilung des Datensatzes zu

berücksichtigen, wurden die relativen Häufigkeiten der objektiven Wörter um die Ungleichheit der erkannten Satz-Sentimente beim Vergleich korrigiert. Die positiven oder negativen Tendenzen werden dann dem objektiven Wort als Sentimentwert zugewiesen. Es wurden nur Wörter mit Tendenzen höher als 0,5 behalten. Auch dieser Schwellenwert führte in der Studie von Hung und Lin (2013) zu optimalen Ergebnissen. Die Idee hinter diesem Schwellenwert ist, dass Worte in Sätzen mit schwachen Sentimenten keinen großen Einfluss auf diese Sätze gehabt haben können. Die frisch errechneten Sentimentwerte der objektiven Wörter wurden dann den jeweiligen Sätzen hinzuaddiert.

Die Ergebnisse der Analyse unterscheiden sich drastisch durch diesen Schritt. Ohne ihn wurden im Korpus 503.343 positive, 222.360 neutrale und bloß 48.649 negative Kommentare entdeckt. Nach der Berechnung der Sentiment-Tendenzen für objektive Wörter wurden von den Kommentaren 750.230 als positiv und lediglich 18.426 als negativ eingestuft. Neutralen Kommentaren wurden ausnahmslos Sentimente zugewiesen. Die positive Tendenz der Kommentare hat sich beträchtlich verstärkt, wie in Abbildung 1 zu sehen ist. Die schwarze, durchgehende Dichtefunktion zeigt die Verteilung der Sentimente vor der Berücksichtigung der Tendenzen objektiver Worte, die rote, strichlierte die Verteilung danach. Die Spitze der schwarzen, durchgehenden Linie beim Nullpunkt zeigt die Menge an (annähernd oder gänzlich) neutralen Kommentaren.

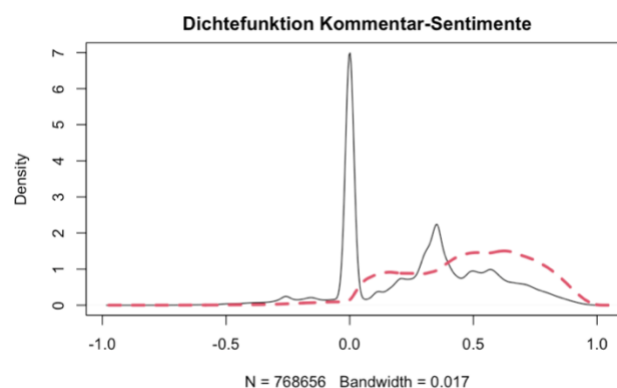


Abbildung 1: Sentiment-Dichtefunktion: vor (schwarz, durchgehend) und nach (rot, strichliert) der Berücksichtigung der latenten Sentimente objektiver Worte

4.8 Normalisierung der Sentimentwerte

Die Autoren von *GerVADER* nutzen für die schlussendliche Eingrenzung der Sentimentwerte in eine Skala von -1 bis 1 eine Normalisierungsformel, die eine

Normalisierungskonstante α mit dem Wert 15 als Parameter nutzt (Formel 1). Um doppelte Normalisierung zu vermeiden, wurden die Werte erst nach der Zusammenfügung der Sätze zu ganzen Kommentaren und der Summierung ihrer Werte normalisiert.

$$x = \frac{x}{\sqrt{x^2 + \alpha}}$$

Formel 1. Normalisierung von Sentimenten

Die Wahl des Alphawertes verändert die Verteilung optisch und macht die kleinen Unterschiede zwischen den zahlreichen, sehr positiven Kommentaren besser erkennbar. In der Abbildung 2 sind für verschiedene Alphas die Histogramme abgebildet. In der Analyse dieser Arbeit wurde Alpha bei 40 festgelegt.

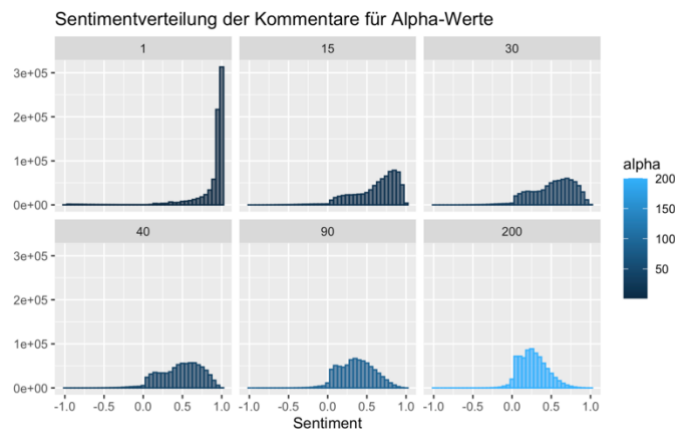


Abbildung 2: Sentiment-Verteilung nach Normalisierung mittels verschiedener Alpha-Werte

5 Analyse

5.1 Verteilung der Sentimente

Anhand der Ergebnisse lässt sich eine ungleiche Verteilung der Sentimente beobachten. Das war, der Literatur nach, auch zu erwarten. Bei Produktbewertungen, einer ähnlichen Domäne wie der hier behandelten, haben Hu et al. (2009) eine J-förmige Verteilungskurve gezeigt. Es soll also sehr viele besonders gute Bewertungen, einige besonders schlechte und kaum dazwischenliegende geben. Das soll dem zentralen Grenzwertsatz widersprechen, laut dem sich eine annähernd normalverteilte Kurve um den Durchschnitt ergeben soll (Hu, Zhang und Pavlou 2009). Hu et al. begründen diese Abweichung unter anderen mit dem sogenannten Under-Reporting-Bias, laut dem Menschen mit besonders guten oder besonders schlechten Erfahrungen mit höherer Wahrscheinlichkeit auch davon berichten würden. Somit sind Sentimente womöglich sehr wohl annähernd normalverteilt, die durchschnittlichen Erfahrungen werden jedoch nicht berichtet.

Im behandelten Kampagnenformat gibt es, wie in Kapitel 1.2 bereits erwähnt, einen systematischen Unterschied zwischen den Teilnehmern und den Autoren von Online-Produktbewertungen: Die Teilnehmer haben sich proaktiv angemeldet, um das jeweilige Produkt testen zu dürfen. Hier wird demnach eine positive Verzerrung der Sentimente erwartet, aufgrund der folgenden Intuition: Ein leidenschaftlicher Joghurtgenießer hat eine höhere Wahrscheinlichkeit, sich für eine Joghurt-Testkampagne zu bewerben. Darum ist ein positiver Bias einem neuen Joghurt gegenüber zu erwarten. Diese Vermutung bestätigt sich durch das Histogramm in Abbildung 3, das die Verteilung der Sentimente aller Kommentare zeigt. Es lässt sich eine starke Tendenz zu positiven Sentimenten beobachten. Die 750.230 positiven Kommentare stellen die 18.426 negativen weit in den Schatten, sodass eine J-Kurve nicht zu erkennen wäre. Von einer annähernd normalverteilten Kurve ist auch nicht zu sprechen. Im Schnitt erkennt das Modell bei den gegebenen Normalisierungsparametern Kommentarsentimente von 0,48. Es gibt darüberhinaus eine signifikante, wenn auch nur schwach positive Korrelation zwischen dem erkannten Sentimente von Kommentaren und ihrer Länge. Eine intuitive Erklärung dafür wäre, dass besonders motivierte oder glückliche Teilnehmer umfassender ausdrücken wollen, um ihre Begeisterung besser zu umschreiben.

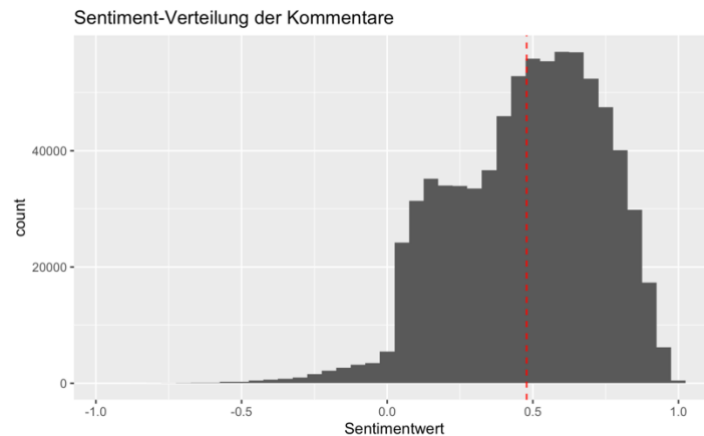


Abbildung 3: Sentiment-Verteilung der Kommentare

Für die Analyse jedes Blogs und jeder Kampagne wurden die durchschnittlichen Sentimente aller zugehöriger Kommentare errechnet. Eine absolute Sentimentsumme zu errechnen wäre nicht zielführend, da die Sentimentwerte mit der Wortmenge signifikant korrelieren. Wenig überraschend führte das dazu, dass keinem Blog oder Kampagne eine durchschnittlich negative Valenz zugeschrieben wurde. Beim Shapiro-Wilk-Test zeigte sich dennoch, dass die Sentimente pro Blog (Abbildung 4) und pro Kampagne (Abbildung 5) nicht annähernd normalverteilt sind. Die Kommentare jedes Blogbeitrags haben im Schnitt ein Sentiment von ca. 0,49, also etwas positiver als dem Durchschnitt der Kommentare. Die Verteilung der durchschnittlichen Sentimente pro Kampagne hat ähnliche Form und Mittelwert.

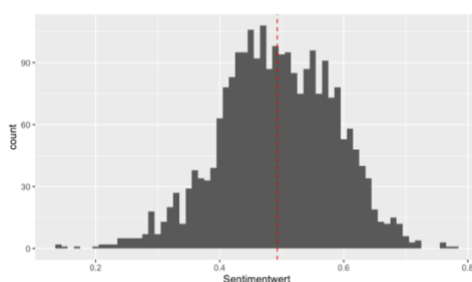


Abbildung 4: Sentiment-Verteilung der Blogbeiträge

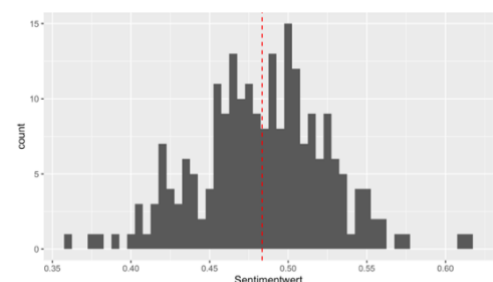


Abbildung 5: Sentiment-Verteilung der Kampagnen

Den Erwartungen entsprechend gibt es eine signifikante, positive Korrelation zwischen den durchschnittlichen Kundenbewertungen der Produkte und den durchschnittlichen, aus den Kommentaren extrahierten, Sentimenten. Mit einem Wert von lediglich 0,33 ist sie jedoch schwächer als vermutet, wobei beachtet werden muss, dass hier alle negativen und positiven Erfahrungen mit den Produkten zusammengefasst wurden. Für eine aussagekräftigere Analyse müssten die einzelnen Teilnehmerbewertungen herangezogen

werden, um die Genauigkeit des Modells anhand der Daten messen zu können. Zwischen den durchschnittlichen Kampagnen-Sentimenten und den durchschnittlichen Weiterempfehlungsquoten gibt es eine ebenfalls signifikante, wenn auch schwache, positive Korrelation von 0,2. Interessanterweise korrelieren die durchschnittlichen Sentimente, sowie die Bewertungen der Kampagnen leicht negativ mit der Anzahl an Blogeinträgen, die über die Kampagnenlaufzeit gepostet werden. Die Kampagnenlaufzeit hingegen hat keine signifikante Korrelation mit den Metriken.

Jede Kampagne wird von einem oder manchmal mehreren Kampagnenmanagern geleitet, die für das Schreiben der Blogbeiträge und die Leitung der Teilnehmer verantwortlich sind. Die Agentur kann die Kampagnenmanager anhand einiger Metriken einordnen. In den Daten ist einsehbar, welche Kampagnenmanager die höchsten Weiterempfehlungsquoten anregen, bei welchen durchschnittlich die meisten Kommentare hinterlassen werden, was das Engagement der Teilnehmer widerspiegelt, und welche durchschnittlichen Sentimente die Kommentare enthalten.

Die Kampagnen der betrachteten Daten sind in 10 verschiedene Kategorien unterteilt. Die meisten Kampagnen bewerben Produkte „Beauty- und Pflege“, die Nächsthäufigen sind die Kategorien „Food & Snacks“, „Haushalt & Reinigung“ und „Technik & Online“. Die durchschnittlich höchsten Weiterempfehlungsraten bekommen Kampagnen der Kategorie „Freizeit & Reisen“, am besten bewertet werden Kampagnen der Kategorie „Getränke“. Die Kommentaranzahlen der Teilnehmer sind in der Kategorie „Haushalt & Reinigung“ am höchsten, gefolgt von „Beauty & Pflege“ und „Gesundheit“.

Dass die Valenzen der Sentimente einen Einfluss auf Verkaufszahlen im Zuge von WOM-Kampagnen haben, wurde bereits angesprochen. Zudem kommt, dass die Varianzen der Sentimente in verschiedenen Domänen verschiedene Einflüsse auf die Effektivität von WOM-Kampagnen haben (Mahesh und Saji 2015). In der Unterhaltungsbranche können hohe Varianzen der Kunden-Sentimente zu besserem Abschneiden von Performances und Filmen führen (Martin, Barron und Norton 2007). In Domänen, in denen die Zielgruppe jedoch risiko-averser ist, wie der Kategorie „Reisen & Freizeit“, können hohe Varianzen in Kunden-Sentimenten jedoch mit weniger Buchungen oder schlechteren Bewertungen zusammenhängen (Ye, Law und Gu 2009). Andere WOM-Forschungen haben aufgedeckt, dass der Einfluss von Sentiment-

Varianzen auf Verkaufszahlen zusätzlich von der durchschnittlichen Kundenbewertung des Produktes oder der Dienstleistung abhängt (Sun 2012). Bei höheren, durchschnittlichen Bewertungen sollen hohe Sentiment-Varianzen mit niedrigeren Verkaufszahlen verbunden sein, wobei bei durchschnittlich schlechter bewerteten Produkten eine hohe Sentiment-Varianz ein Indiz auf qualitative Nischenprodukte sein soll. Die Kategorie „Freizeit & Reisen“ hat in den *Trnd*-Daten die niedrigste Varianz der Kundensentimente pro Kategorie, zudem auch die zweithöchsten, durchschnittlichen Kundenbewertungen. Wäre die Varianz höher, sollte Kampagnenmanagern dies auffallen, entsprechende Kommunikation der Ergebnisse an die jeweiligen Brandmanager wäre vorteilhaft. Die Kategorie „Technik & Online“ hingegen verzeichnet relativ niedrige, durchschnittliche Kundenbewertungen, verglichen mit anderen Kategorien. Die erkannten Sentimente der Kommentare haben jedoch relativ hohe Varianz. Die Kampagnen der Kategorie „Getränke“ hingegen sind von allen Kampagnen im Schnitt am besten bewertet, ihre Kundensentimente haben jedoch von allen Kategorien die höchste Varianz, was laut Sun (2012) nicht vorteilhaft ist. Auch diese Erkenntnisse sollten Teil einer Reportage der Kampagnenmanager an die Brandmanager sein.

Auf WOM-Plattformen wie *Trnd* finden sich sowohl einmalig teilnehmende als auch immer wiederkehrende Nutzer. Der Median der Anzahl der Kampagnen, an denen Nutzer teilnehmen, liegt bei 3. Der Mittelwert von 6 ist weniger aussagekräftig, im Boxplot der Abbildung 6 sieht man, dass hier viele Ausreißer vorliegen. Grund dafür sind mehrfach verwendete Benutzernamen, auf der Plattform wurden Kampagnen im maximal dreistelligen Bereich durchgeführt. Während die Teilnehmer im Schnitt 190 Wörter schreiben, gibt es zahlreiche, übermotivierte Nutzer, die Texte mit über 3000 Wörtern verfassen. Wie oben bereits erwähnt haben solche langen Texte eine Tendenz zu positiven Sentimenten, mit einer signifikanten Korrelation von 0,26.

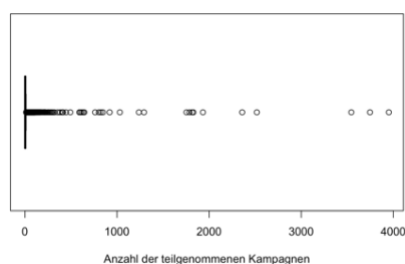


Abbildung 6: Boxplot zur Anzahl der Kampagnen jedes Teilnehmers

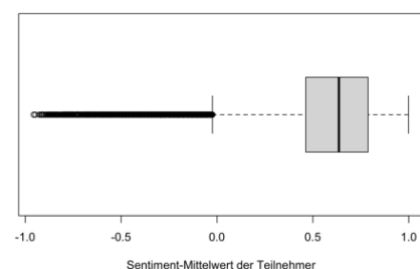


Abbildung 7: Boxplot zum mittleren Sentiment jedes Teilnehmers

Durch die Markierung der Kommentare mit Sentimenten können nörgelnde Nutzer erkannt werden, erkannt durch konsequent negative Kommentare mit wenigen angesprochenen Produktaspekten. Bei der Suche nach solchen „Nörglern“ wurde entdeckt, dass ein großer Teil der negativen Kommentare mit Lieferungsproblemen zusammenhängt, einige Nutzer haben bei jeder Teilnahme Pech und erhalten ihre Pakete spät oder gar nicht. Tabelle 10 zeigt alle Kommentare eines Benutzers welcher sich ausschließlich über Lieferschwierigkeiten beklagt. Der Kampagnenmanager hätte hiermit, einen weiteren Kritikpunkt entdeckt, den es zu behandeln gilt.

compound	comment_text
-0.47	Mein Paket ist leider auch noch nicht da und hat noch den gleichen Status vom 12.08.. :(
0.11	wie lang wird es denn bei mir noch dauern? Status ist bei mir noch vom 12.08.. hat sich noch nicht's getan und warte schon Geduldig :)
0.07	Schade.. auch ich habe noch kein Paket erhalten und der Status ist noch vom 12.08.. wurde nicht irgendwo genannt das heute die letzten Pakete rausgehen sollen? :/
0.08	Ihr seid nicht allein, der Status hat sich bei mir auch noch nicht verändert und warte noch..
-0.13	Bei mir noch das gleiche.. am 12.08. vbermittelt aber noch kein weiterer Status.. dann mal weiter abwarten :/
-0.15	Hab mein Paket leider auch noch nicht.. ab wann kommt denn die Sendungsnummer zur selbstkontrolle ins Profil ?
-0.33	Schade, DHL ist vorbei und leider noch kein Paket dabei.
-0.23	Die Sendungsnummer ist im profil leider auch noch nicht hinterlegt

Tabelle 10: Alle Kommentare von Teilnehmer „blackstar1487“

5.2 Entwicklung der Sentimente

Da in den meisten Kampagnen über ihre Laufzeit mehrere Blogbeiträge gepostet werden, können die Entwicklung von Engagement der Teilnehmer und durchschnittliche Sentimente in deren Kommentaren visualisiert werden. Es gibt eine signifikante, leicht negative Korrelation (-1,8) zwischen der Nummer des aktuellen Blogeintrages und der erkannten Sentimente der Kommentare zum Beitrag. Das betrifft alle Kampagnenmanager zusammengefasst. In Abbildung 8 sind die Sentimente der Kommentare pro konsekutivem Blogeintrag für Kampagne abgebildet. Abbildung 9 zeigt die Entwicklung der Sentimente mehrerer Kampagnen.

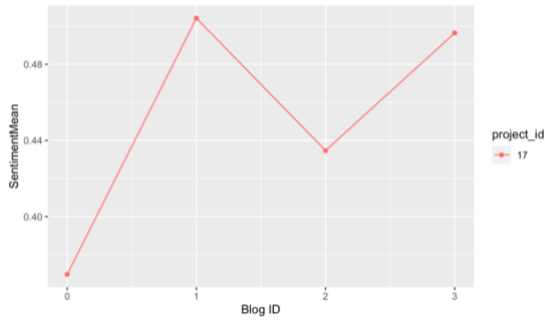


Abbildung 8: Sentiment-Entwicklung für Kampagne 17

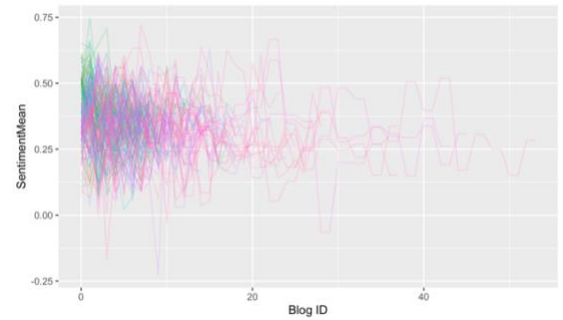


Abbildung 9: Sentiment-Entwicklung für 50 Kampagnen

Somit können die Kampagnenverläufe verschiedener Kampagnenmanager verglichen werden. Abbildung 10 zeigt die Entwicklungen der Kommentar-Sentimente in den Kampagnen des Kampagnenmanagers *shirley*.

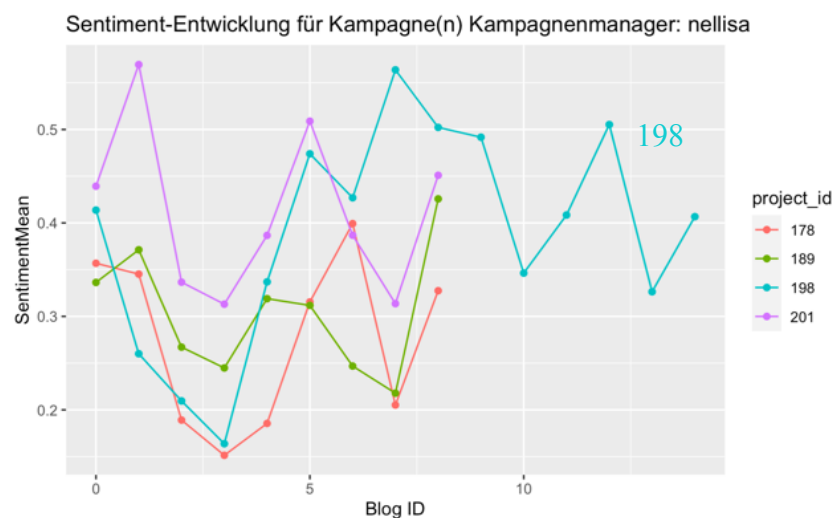


Abbildung 10: Sentiment-Entwicklung für Kampagnenmanager "nellisa"

Mit diesen Einsichten haben Kampagnenmanager nun die Möglichkeit, festzustellen, an welchen Zeitpunkten emotionale Wendepunkte in der Kundenzufriedenheit stattfanden. Der Kampagnenmanager *shirley* hat, wie in Abbildung 10 am türkisen Graphen (198) zu sehen ist, verhältnismäßig schwach positive Kommentar-Sentimente nach dem ersten Blogeintrag verzeichnet. Ab dem fünften Blogeintrag wurden diese jedoch dauerhaft positiver. Der Tiefpunkt beim dritten und vierten Blogeintrag ist auch bei seinen anderen Kampagnen zu sehen. Diese Analyse kann dem Kampagnenmanager Anlass geben, die Kommentare der Kampagnen zu diesen Zeitpunkten stichprobenartig oder gezielt die besonders negativen Kommentare zu analysieren, um einen möglichen Grund dafür zu finden. Gründe können, unter der Annahme, das Sentimentanalysemodell sei nicht von systematischem Bias geplagt, in der Formulierung der Fragen oder Anforderungen in

seinem Blogeintrag liegen. Solche Tiefpunkte können auch ein Hinweis auf generelle Veränderungen in der Kundenzufriedenheit mit einem Produkt sein.

Wie in im Boxplot der Abbildung 11 ersichtlich, haben Kommentare der Kampagne 198 von Kampagnenmanager *shirley* ein Median-Sentiment von knapp über 0,5, was über dem Kampagnen-Mittelwert von 0,48 (rote, strichlierte Linie) liegt. Das statistische Minimum ist jedoch nicht positiv, es sind also nicht alle Ausreißer negativ.

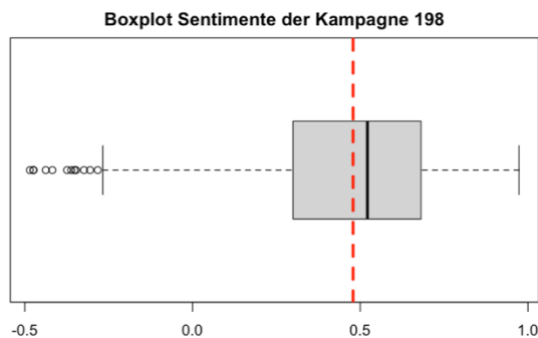


Abbildung 11: Sentimente der Kampagne 198

In der Abbildung 10 war für die Kampagne 198 (türkis) ein starkes Sentiment-Minimum beim vierten Blogeintrag verzeichnet. Eine intuitive Begründung für diesen Tiefpunkt wäre, dass es besonders viele negative Kommentare und/oder wenige positive gegeben hat. Die türkisen, durchgehenden Linien in Abbildungen 12 und 13 zeigen die durchschnittlichen, erkannten Sentimente pro Blogeintrag der Kampagne 198, die orange, strichlierte Linie in Abbildung 12 hingegen zeigt die Anzahl der negativen Kommentare für denjenigen Blogeintrag. Zwischen Blogeinträgen 2 und 4 sind relativ viele, negative Kommentare verzeichnet. Die Maxima der orangen, strichlierten Linie in Abbildung 12 stimmen mit einem der Minima der türkisen Linie überein. In Abbildung 13 zeigt sich, dass besonders positiver Kommentar-Sentimente, welche durch die pinke, strichlierte Linie dargestellt werden, zu Kampagnenbeginn ebenfalls relativ selten vorkommen. Bemerkenswert ist auch, dass sowohl negative Kommentare (Abbildung 12, orange, strichliert) und besonders positive Kommentare (Abbildung 13, pink, strichliert) sehr häufig vorkommen. An der Linie in Abbildung 14 ist zu erkennen, dass bei Blog 7 ein grundsätzliches Maximum an Kommentaren herrscht, am Histogramm in Abbildung 15 sieht man jedoch, dass die Sentimente der Kommentare für Kampagne 198 Blog 7 nicht anders als sonst verteilt sind, es gibt hier nicht mehr stark-emotionale Kommentare als sonst.

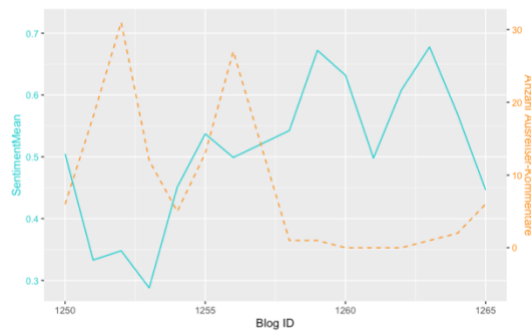


Abbildung 12: durchschnittliche (türkis, durchgehend) und besonders negative (orange, strichliert) Sentimente der Kampagne 198

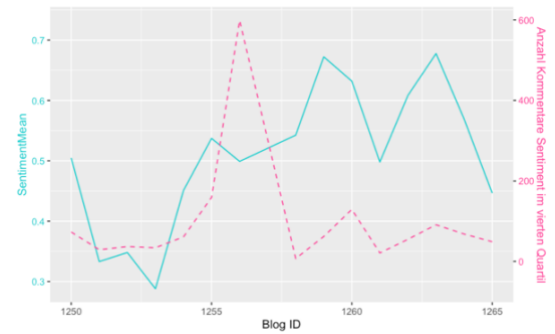


Abbildung 13: durchschnittliche (türkis, durchgehend) und besonders positive (pink, strichliert) Sentimente der Kampagne 198

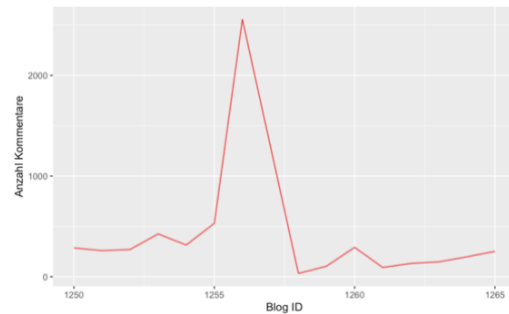


Abbildung 14: Anzahl der Kommentare der Kampagne 198

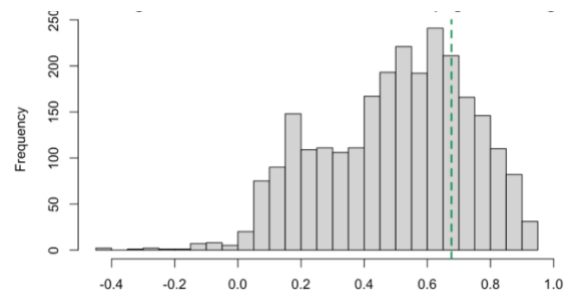


Abbildung 15: Verteilung der Sentimente der Kampagne 198 Blogeintrag 7

5.3 Sentiment-Extreme

Besonders starke Sentimente, in beide Richtungen, aber vor allem negative, sollten die Aufmerksamkeit von Kampagnenmanagern erhalten, hier könnten sich wunde Punkte in der Kundenzufriedenheit zeigen. Ihre Bearbeitung für die Marke von Vorteil sein könnte. Tabelle 11 enthält einige Kommentare, deren Sentimente sehr negativ eingeschätzt wurden. Die Kommentare der ersten, zweiten und fünften Reihe enthalten tatsächliche Kritiken an dem Produkt, bzw. Beschwerden darüber. Der Autor des sechsten Kommentars hingegen klagt über die Ausführung der Lieferung. Der dritte Kommentar bringt klare Kritiken an dem Produkt, schließt jedoch mit einem positiven Fazit ab. Das ist ein Beispiel für einen kritisierenden Kommentar, dessen Sentiment vermutlich eher positiv sein sollte. Da die Message jedoch faktisch negative Umstände umschreibt, ist die negative Annotation nicht unangebracht.

comment_author	compound	wordCount	comment_text
	-0.8928	734	Wir haben den Easy Pull nun auch schon reichlich geprüf und leider hat er nicht so gut bei uns abgeschnitten. Leider reißen die Blätter immer ziemlich unkontrolliert ab mal mitten drinnen, dann hat man nur eine Ecke oder mit etwas Glück mal ein ganzes Blatt. Die Blätter finden wir auch ein bisschen zu klein im Gegensatz zu den Konkurrenzblättern. Das Design gefällt uns leider für diesen Preis auch nicht so gut. Es ist halt Plastik und das ganze für knappe 11,-€ finden wir ein bisschen zu teuer. Desweiteren heißt bei uns auch das Silikon Pad nicht so gut, es ist uns schon beim abreißen passiert das der Spender mit umgekippt ist. Nun haben wir ja noch ein bisschen Zeit vielleicht kann uns der Spender ja doch noch überzeugen ;)
	-0.8841	394	Mein Paket ist heute angekommen. Habe es natürlich gleich geöffnet und den Zewa Easypull vorbereitet und ausprobiert. Ich habe ihn jetzt mal noch nicht fest geklebt und so hat er mich noch nicht überzeugt. Ich habe es zweimal versucht einhängend ein Papierstück abzureißen und beides mal hat es nicht richtig funktioniert. Ich werde es mal weiter probieren und auch mal wenn es fest geklebt ist.
	-0.8681	201	Leider kann auch ich nicht mit einer Hand abreißen. Außerdem heißt es bei mir NICHT kopfüber. Und das Haft Pad nutzt mir nichts, weil ich fast nur gedölte Oberflächen habe. Gefällt mir trotzdem ganz gut!
	-0.8249	302	Hallo zusammen. Habe mir leider meinen Esstisch ein wenig demoliert. Habe einen Tisch aus alten schiffsplanken esypull daraufgestellt zum vorführen und am Ende hatte ich dann einen kl Teil der Farbe vom Tisch am Pö. das fand ich jetzt nicht so witzig. Also gebt Obacht wo ihr ihn aufstellt. Lg Silvi
	-0.7952	465	Ich habe mein Paket gestern erhalten und den schwarzen Spender aufgestellt. Es hat eine Zeit gedauert, bis ich das erste Blatt aus der Rollenmitte erwisch habe und dann ist es gleich gerissen. Nach dem Einsetzen der Rolle sind noch ein paar Blätter in dem Behälter abgerissen, das ist nicht sehr optimal :- (Bei dem Papier teile ich Mullchens Meinung. Bin schon gespannt, wie sich das Papier im Alltagstest schlägt, der erste Eindruck ist noch nicht überzeugend.
	-0.7719	119	ich habe mein pöckchen noch nicht erhalten seit dienstag im selben lieferstatus :- (das kann doch nicht richtig sein :- (

Tabelle 11: besonders negative Kommentare

Tabelle 12 zeigt die negativsten Kommentare der „Beauty & Pflege“-Kampagne 198, welche ein paar Schwachpunkte des Analysemodells zeigen. Die negativen Sentimente der Beispielskommentare aus der Tabelle für die Blogs 1250 und 1251, die ersten Blogs der Kampagne, scheinen nicht mit den Produkten selbst in Verbindung zu stehen. Hingegen merken sie an, dass die Teilnehmer Probleme bei der Anteilnahme hatten. Einige andere Kommentare erwähnten Hautprobleme und -irritationen, welche als negative Sentiment-Indikatoren gelesen werden, auch wenn sie für ein Hautpflegeprodukt sprechen könnten. Gerade bei Heilprodukten kann das Umschreiben der Probleme, die das Produkt gelöst hat, bei der Anpreisung der Wirksamkeit helfen. Hierbei wäre ein wissensbasiertes Modell, dass domänenspezifische Grammatiken berücksichtigt, überlegen.

blog_id	comment_author	compound	wordCount	comment_text
1250		-0.694	327	Schade - hab leider kein Ticket und könnte dringend solch eine Creme gebrauchen. Meine Haut ist so sensibel, dass selbst meine Kosmetikerin schon verzweifelt ist, da ich sogar auf die Linie für Neurodermatiker allergisch reagiere. Vielleicht ist ja noch ein Plätzchen frei und es verirrt sich ein Ticket in meinen Posteingang;)
1250		-0.6577	179	ahhhh was besser gibt es nicht. Super spannend vor allem weil ich später in der Apotheke auch in solcher Beratung tätig bin, nur leider die Bewerbungsphase verpasst= (blödd gelaufen
1251		-0.8421	178	Hallo Ich habe mich verlickt und jetzt:-(kann den fehler nicht verbessern . Das macht mich traurig . Ich würde so gerne dabei . Man müsste die chance haben etwas zu verbessern .
1251		-0.6077	52	Schade zwei fehler gemacht, leider nicht dabei. Mist
1251		-0.4588	126	Ich bin gespannt und meine Haut auch, Ich werde jetzt den Wissenstest ausfüllen und drücke meiner ramponierten Haut die Daumen
1253		-0.6478	102	Im Badezimmer und im Sommer wird alles in den KÜhlschrank verfrachtet. Ist leider sehr umständlich :o(
1256		-0.6597	619	Guten Morgen Laura, ich habe gerade mal ein bisschen auf der Avène Seite gestöbert und bin unter ""Ratgeber bei Rötungen"" auf eine Erklärung meiner angespannten und roten Haute gestoßen. Und zwar leide ich offensichtlich an Rosacea. Noch treten nur selten Pusteln auf, doch ich werde auf jeden Fall eure Antirougeurs testen, da ich noch immer keine Creme gefunden habe, die mein Hautproblem lindert. Ich hoffe sehr, das Problem damit in den Griff zu bekommen denn ich leide schon sehr unter diesen Rötungen, vor allem, weil ich bis vor ca. einem Jahr keinerlei Probleme hatte und dann auf einmal kamen diese Rötungen...

Tabelle 12: besonders negative Kommentare der Kampagne 198

6 Zusammenfassung

Sentimentanalysen wurden bereits in vielerlei Kontexten angewandt, diese Arbeit behandelt die wenig erforschte Nutzung solcher Analysen für Seeding-Kampagnen und Potenzial bei der Messung derer Effektivität. Um die impliziten Informationen in den Kommentaren von Produkttestern bei solchen Kampagnen ausnutzen zu können, wurde eine lexikon-basierte Sentimentanalyse durchgeführt. Grundlage für das Analysemodell war das Modul *GerVADER*, welches durch, von der Literatur empfohlene, Regeln erweitert wurde. Die Kommentare wurden Satz für Satz auf ihre Sentimente untersucht. Somit konnten Einblicke in die Emotionalitäten der Tester geboten werden, was bei Evaluierung der Effektivität einer Kampagne hilft. Es wurden positive Korrelationen zwischen den Kundenbewertungen, den Weiterempfehlungsquoten und den Sentimenten der Kommentare entdeckt. Die Sentimente korrelierten außerdem positiv mit der Länge der Kommentare und negativ mit der Nummer des Blogs, auf die sie antworteten. Durch die Quantisierung von Sentimenten konnte außerdem ermittelt werden, welche Kampagnenmanager emotionale Reaktionen ihrer Teilnehmer hervorbrachten. Darüber hinaus konnten für verschiedene Domänen die Varianzen der Sentimente ermittelt werden, sowie deren Implikationen für die Effektivität der Kampagnen. Außerdem ließen sich die emotionalen Erfahrungen über den Kampagnenverlauf und Anomalien durch solche Analysen visualisieren. Letztlich erlaubte die Sentimentanalyse, Schwachpunkte der Kampagnen und ihrer Durchführung ersichtlich zu machen, was Kampagnenmanager und Brandmanager über davor unbekannte Kundenprobleme informieren kann.

Da die Bewertungen für die einzelnen Kommentare im verwendeten Datensatz nicht verfügbar waren und keine ausreichend ähnlichen Testdaten, deren Domänen den behandelten ähnlich genug waren, als freie Ressource zur Verfügung standen, konnte die schlussendliche Genauigkeit des Modells nicht genau ermittelt werden. Mit den tatsächlichen Sentimenten annotierte Daten wären erforderlich.

Ein lexikon-basiertes Sentimentanalysemodell stößt selbst optimiert an seine Grenzen, wenn weitere Regeln die Genauigkeit nicht mehr erhöhen können, ohne weiteren Bias zu schaffen (Singh 2018). In weiterer Folge könnten Daten aus deutschen Seeding-Kampagnen mittels maschineller Lernverfahren oder wissensbasierter Modelle analysiert werden, um noch genauere Analysen und dadurch verlässlichere Einblicke zu gewähren. Auch hierfür wären granular annotierte Kommentar-Sentimente von Vorteil.

Mit zuvor annotierten Texten könnten zudem inhaltliche Analysen mit Sentimentanalysen verbunden werden. So könnten die angesprochenen Aspekte einer Bewertung auf Abhängigkeit mit der Emotionalität geprüft werden, um emotional bedeutungsvolle Aspekte zu identifizieren. Die Fragen der Blogautoren können mit Themen-Extraktion auch auf potenzielle Einflüsse auf die Emotionen der antwortenden Nutzer getestet werden.

Anhang

Diese Arbeit, die erhaltenen und verarbeiteten Daten, sowie alle benutzten Module, Lexika und die geschriebenen Code-Notebooks sind unter folgendem Link abrufbar:

<https://www.dropbox.com/sh/fagcux9l5vh4r2k/AABDpLo5n0GZOeykadI6FxW6a?dl=0>

Literaturverzeichnis

- Bansala, Barkha, und Sangeet Srivastavaa. 2018. „On predicting elections with hybrid topic based sentiment analysis of tweets.“
- Chintagunta, P, S Gopinath, und S Venkataraman. 2010. „The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets.“ *Marketing science* Vol. 29(5): 944-957.
2021. *Clickworker*. Zugriff am 20. 11 2021. <https://www.clickworker.de/sentiment-analyse/>.
- Dabhade, Vinay. 2021. *Conducting Social Media Sentiment Analysis. A Working Example*. Zugriff am 11. 11 2021. <https://expressanalytics.com/blog/social-media-sentiment-analysis/>.
- Davis, Dwight. 2019. „AI Unleashes the Power of Unstructured Data.“ *CIO*. Zugriff am 29. 10 2021. <https://www.cio.com/article/3406806/ai-unleashes-the-power-of-unstructured-data.html>.
- Dodds, Peter. 2007. „Influentials, Networks, and Public Opinion Formation.“ *Journal of Consumer Research* Vol. 34(4): 441-458.
- Dost, Florian, Ulrike Phieler, Michael Haenlein, und Barak Libai. 2018. „Seeding as Part of the Marketing Mix: Word-of-Mouth Program Interactions for Fast Moving Consumer Goods.“ *Journal of Marketing*, Vol. 83(2): 62-81.
- Esuli, Andrea, und Fabrizio Sebastiani. 2010. *SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining*. Pisa: Istituto di Scienza e Tecnologie dell’Informazione, Consiglio Nazionale delle Ricerche.
- Fehle, Jakob, Thomas Schmidt, und Christian Wolff. 2021. *Lexicon-based Sentiment Analysis in German: Systematic Evaluation of Resources and Preprocessing Techniques*. Düsseldorf: KONVENS 2021 Organizers.
- Fulgoni, Gian, und Andrew Lipsman. 2015. „Digital Word of Mouth And Its Offline Amplification: A Holistic Approach to Leveraging And Amplifying All Forms of WOM.“ *Journal of Advertising Research*, Vol. 55(1): 18.
- Garg, Rashi. 2018. „Geeks For Geeks.“ *Social Network API*. Zugriff am 10. 10 2021. <https://www.geeksforgeeks.org/social-network-api/>.
- GitHub. 2021. *MIT License*. Zugriff am 30. 11 2021. <https://github.com/git/git-scm.com/blob/main/MIT-LICENSE.txt>.
- Goldenberg, Jacob, Sangman Han, Donald Lehrman, und Jae Weon Hong. 2009. „The Role of Hubs in the Adoption Process.“ *Journal of Marketing* Vol. 73(2): 1-13.

- Gräve, Jan-Frederik. 2019. „What KPIs Are Key? Evaluating Performance Metrics for Social Media Influencers.“ *Social Media + Society*. 7.
- Haenlein, Michael, und Barak Libai. 2017. „Seeding, Referral, and Recommendation: Creating Profitable Word-of-Mouth Programs.“ *California Management Review*, Vol. 59(2): 68–91.
- Hartmann, Jochen, Juliana Huppertz, Christina Schamp, und Mark Heimann. 2019. „Comparing automated text classification methods.“ *International Journal of Research in Marketing* 20-38.
- Harvey, Steve. 2021. „The power of emotional marketing: Once more with feeling.“ *Fabrikbrands*. Zugriff am 1. 12 2021. <https://fabrikbrands.com/the-power-of-emotional-marketing/>.
- Hu, Nan, Jie Zhang, und Paul Pavlou. 2009. „Overcoming the J-Shaped Distribution of Product Reviews.“ *Communications of the ACM* Vol. 52(10): 144-147.
- Hung, Chihli, und Hao-Kai Lin. 2013. „Using Objective Words in SentiWordNet to Improve Word-of-Mouth Sentiment Classification.“ *Intelligent Systems*, Vol. 28(2): 47-54.
- Hutto, C.J. , und E.E. Gilbert. 2014. „VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text.“ *Eighth International Conference on Weblogs and Social Media*.
2021. *Influencer Marketing Hub*. Zugriff am 1. 12 2021. <https://influencermarketinghub.com/influencer-rates/>.
- Jayaswal, Vaibhav. 2020. „Towards Data Science.“ *Text Vectorization: Term Frequency — Inverse Document Frequency (TFIDF)*. Zugriff am 20. 11 2021. <https://towardsdatascience.com/text-vectorization-term-frequency-inverse-document-frequency-tfidf-5a3f9604da6d>.
- Khan, Wahab, Ali Daud, Tehmina Amjad, und Jamal A. Nasir. 2016. „A survey on machine learning models for Natural Language Processing (NLP).“ *Researchgate*. Zugriff am 25. 11 2021. https://www.researchgate.net/publication/311436163_A_survey_on_machine_learning_models_for_Natural_Language_Processing_NLP.
- Khemka, Rishabh. 2020. „Medium.“ *Emojis — From a programmer's eye*. Zugriff am 1. 11 2021. <https://medium.com/bobble-engineering/emojis-from-a-programmers-eye-ca65dc2acef0>.

- Killer, Markus. 2014. „textblob-de Documentation.“ *textblob-de* . Zugriff am 11. 11 2021. <https://textblob-de.readthedocs.io/en/latest/>.
- Komban, Stanely Jose, Raghavendra Prasad Narayan, und Giorgio Orsi. 2019. *Deep Learning Models for Sentiment Analysis*. Zugriff am 24. 11 2021. <https://underthehood.meltwater.com/blog/2019/08/22/deep-learning-models-for-sentiment-analysis/#:~:text=Traditional%20machine%20learning%20methods%20such,analysis%20because%20they%20scale%20well>.
- Lee, Admond. 2018. *Why and How to Use Pandas with Large Data*. Zugriff am 1. 11 2021. <https://towardsdatascience.com/why-and-how-to-use-pandas-with-large-data-9594dda2ea4c#:~:text=Pandas%20is%20very%20efficient%20with,performance%20is%20rarely%20a%20concern>.
- Lee, Kevin. 2021. „Towards Data Science.“ *Sentiment Analysis — Comparing 3 Common Approaches: Naive Bayes, LSTM, and VADER*. Zugriff am 6. 11 2021. <https://towardsdatascience.com/sentiment-analysis-comparing-3-common-approaches-naive-bayes-lstm-and-vader-ab561f834f89>.
- Liu, Bing. 2020. *Sentiment Analysis: mining sentiments, opinions, and emotions*. Cambridge University Press.
- Liu, Bing, und Minqing Hu. 2004. „Mining and summarizing customer reviews.“ *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Seattle, Washington.
- Loria, Steven. 2020. *TextBlob: Simplified Text Processing*. Zugriff am 3. 11 2021. <https://textblob.readthedocs.io/en/dev/>.
- Mahesh, Balan, und Mathew Saji. 2015. „Online word of mouth using text mining: A review of literature and future directions.“ *2015 IEEE Workshop on Computational Intelligence: Theories, Applications and Future Directions (WCI)* .
- Manning, Christopher, Prabhakar Raghavan, und Hinrich Schütze. 2009. *Introduction to Information Retrieval*. Cambridge University Press.
- Marr, Bernard. 2019. „Forbes.“ *What Is Unstructured Data And Why Is It So Important To Businesses? An Easy Explanation For Anyone*. Zugriff am 20. 10 2021. <https://www.forbes.com/sites/bernardmarr/2019/10/16/what-is-unstructured->

data-and-why-is-it-so-important-to-businesses-an-easy-explanation-for-anyone/?sh=2986711c15f6.

- Martin, J, G Barron, und M Norton. 2007. „Choosing to be uncertain: Preferences for high variance experiences.“ *Book Choosing to be uncertain: Preferences for high variance experiences*.
- Medhat, W, A Hassan, und H Korashy. 2014. „Sentiment analysis algorithms and applications: A survey.“ *Ain SHams Engineering Journal*.
- Pang, Bo, und Lillian Lee. 2008. „Opinion Mining and Sentiment Analysis.“ *Foundations and Trends® in Information Retrieval* Vol. 2(1–2): 1-135.
- Pocs, Mate. 2020. „Towards Data Science.“ *Lovecraft with Natural Language Processing — Part 1: Rule-Based Sentiment Analysis*. Zugriff am 11. 11 2021. <https://towardsdatascience.com/lovecraft-with-natural-language-processing-part-1-rule-based-sentiment-analysis-5727e774e524>.
- Prasad, Sunit. 2021. *Analytix Labs*. Zugriff am 24. 10 2021. <https://www.analytixlabs.co.in/blog/types-of-machine-learning/>.
- Rauh, Christian. 2018. „Validating a sentiment dictionary for German political language—a workbench note.“ *Journal of Information Technology & Politics* Vol. 15(4): 319-343.
- Remus, R, U Quasthoff, und G Heyer. 2010. „SentiWS. A Publicly Available German-language Resource for Sentiment Analysis.“ *Proceedings of the 7th International Language Resources and Evaluation (LREC'10)* 1168-1171.
- Ribeiro, Filipe, Matheus Araújo, Pollyanna Gonçalves, Marcos André Gonçalves, und Fabrício Benevenuto. 2016. „SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods.“ *Ribeiro et al. EPJ Data Science (2016)* 5:23 DOI 10.1140/epjds/s13688-016-0085-1.
- Romanyshyn, Mariana. 2013. „Rule-Based Sentiment Analysis of Ukrainian Reviews.“ *International Journal of Artificial Intelligence & Applications*, 103-111.
- Salloum, Said, Mostafa Al-Emran, und Khaled Shaalan. 2017. „Mining Social Media Text: Extracting Knowledge from Facebook.“ *IJCDS Journal* 6(2): 73-81.
- Sanketh, Ruthu S. 2020. „Towards Data Science.“ *POS Tagging Using CRFs*. Zugriff am 11. 13 2021. <https://towardsdatascience.com/pos-tagging-using-crfs-ea430c5fb78b>.
- Sarica, Serhad, und Jianxi Luo. 2021. *Stopwords in Technical Language Processing*. Bd. <https://doi.org/10.1371/journal.pone.0254937>.

- Sarprasatham, Matilda. 2016. „Big Data in Social Media Environment: A Business Perspective.“ In *Social Media Listening and Monitoring for Business Applications*, 10.4018/978-1-5225-0846-5.ch004.
- Schuller, Björn, und Tobias Knaup. 2011. „Learning and Knowledge-Based Sentiment Analysis in Movie Review Key Excerpts.“ *Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues. Lecture Notes in Computer Science* Vol. 6456.
2021. *SentiStrength*. Zugriff am 11. 11 2021. <http://sentistrength.wlv.ac.uk/#About>.
- Siegel, Melanie, Jennifer Deuschle, Barbara Lenze, Marina Petrovic, und Sascha Starker. 2017. „Automatische Erkennung von politischen Trends mit Twitter – brauchen wir Meinungsumfragen noch?“ *NFD Information-Wissenschaft und Praxis* 68(1): 67-74.
- Singh, Seema. 2018. „Towards Data Science.“ *Understanding the Bias-Variance Tradeoff*. Zugriff am 2. 11 2021. <https://towardsdatascience.com/understanding-the-bias-variance-tradeoff-165e6942b229>.
- Sun, M. 2012. „How does the variance of product ratings matter?“ *Management Science* Vol. 8(4): 696-707.
2021. „Systemic Sciences.“ *Latent Semantic Analysis*. Zugriff am 10. 10 2021. <http://systems-sciences.uni-graz.at/etextbook/bigdata/lisa.html>.
- Taboada, Maite, Julian Brooke, Milan Tofiloski, Kimberly Voll, und Manfred Stede. 2011. „Lexicon-Based Methods for Sentiment Analysis.“ *Computational Linguistics* Vol. 37(2): 267–307.
- Terasa, Jan Christoph. 2020. „Python.“ *TimeComplexity*. Zugriff am 11. 11 2021. <https://wiki.python.org/moin/TimeComplexity>.
2021. *Trnd*. Zugriff am 10. 10 2021. <https://www.trnd.com/de/>.
- Tymann, Karsten Michael, Matthias Lutz, Patrick Palsbröcker, und Carsten Gips. 2021. „GerVADER - A German adaptation of the VADER sentiment analysis tool for social media texts.“ *Proceedings of the Conference "Lernen, Wissen, Daten, Analysen"*. <https://github.com/KarstenAMF/GerVADER>.
- Tymann, Karsten, Louis Steinkamp, Oxana Zhurakovskaya, und Carsten Gips. 2020. „Native sentiment analysis tools vs. translation services - Comparing GerVADER and VADER.“ *Lernen, Wissen, Daten, Analysen*.
- Vermeer, Susan, Theo Araujo, Stefan Bernritter, und Guda Van Noort. 2019. „Seeing the wood for the trees: How machine learning can help firms in identifying relevant

- electronic word-of-mouth in social media.“ *International Journal of Research in Marketing* 36 (3): 492-508.
- Vishwakarma, Ashish. 2021. „Difference between Structured, Semi-structured and Unstructured data.“ *Geeks For Geeks*. Zugriff am 20. 10 2021. <https://www.geeksforgeeks.org/difference-between-structured-semi-structured-and-unstructured-data/>.
- Waltinger, Ulli. 2010. „GermanPolarityClues - A Lexical Resource for German Sentiment Analysis.“ Zugriff am 11. 11 2021. <http://www.ulliwaltinger.de/sentiment/>.
- Waltinger, Ulli. 2010. „Sentiment Analysis Reloaded - A Comparative Study on Sentiment Polarity Identification Combining Machine Learning and Subjectivity Features.“ *Proceedings of the 6th International Conference on Web Information Systems and Technologies* Vol. 1.
- Wolfgruber, Marlene. 2015. *Sentiment Analyse mit lokalen Grammatiken. Wissensbasierter Ansatz zur Extraktion von Sentiments in Hotelbewertungen*. Münster: Verlagshaus Monsenstein und Vannerdat.
2021. *Wortschatz Universität Leipzig*. Zugriff am 30. 9 2021. <https://wortschatz.uni-leipzig.de/de/download#sentiWSDownload>.
- Wójcik, Rafał. 2019. „Towards Data Science.“ *Unsupervised Sentiment Analysis – How to extract sentiment from the data without any labels* . Zugriff am 6. 11 2021. <https://towardsdatascience.com/unsupervised-sentiment-analysis-a38bf1906483>.
- . 2019. „Unsupervised Sentiment Analysis. How to extract sentiment from the data without any labels.“ *Towards Data Science*. Zugriff am 1. 11 2021. <https://towardsdatascience.com/unsupervised-sentiment-analysis-a38bf1906483>.
- Ye, Q, R Law, und B Gu. 2009. „The impact of online user reviews on hotel room sales.“ *International Journal of Hospitality Management* Vol. 28(1): 180-182.
- Yousef, Ahmed Hassan, Walaa Medhat, und Hoda K. Mohamed. 2014. „Sentiment Analysis Algorithms and Applications: A Survey.“ *Ain Shams Engineering Journal* Vol. 5(4).
- Zarei, Koosha, Damilola Ibosiola, Reza Farahbakhsh, Zafar Gilani, Kiran Garimella, Noel Crespi, und Gareth Tyson. 2020. „Characterising and Detecting Sponsored Influencer Posts on Instagram.“ Zugriff am 23. 11 2021. <https://later.com/blog/influencer-engagement-rate/>.

Zhang, Zhu, Xin Li, und Yubo Chen. 2012. „Deciphering Word-of-Mouth in Social Media: Text-Based Metrics of Consumer Reviews.“ *ACM Transactions on Management Information Systems* 3(1).

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Bachelorarbeit selbständig verfasst und in der Bearbeitung und Abfassung keine anderen, als die angegebenen Quellen oder Hilfsmittel benutzt, sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe. Die vorliegende Bachelorarbeit wurde noch nicht anderweitig für Prüfungszwecke vorgelegt.

8.12.2021

Datum



Nikolaus Czernin