

Opinion Mining von Daten aus Sozialen Medien zur Bewertung von
Produkteigenschaften

Bachelorarbeit

zur Erlangung des akademischen Grades „Bachelor of Science (B. Sc.)“ im Studiengang
Wirtschaftswissenschaft der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Name: Tesfazgi



Vorname: Awet



Prüfer: Prof. Dr. Michael H. Breitner

Hannover, den 10.08.2016

Inhaltsverzeichnis

Abbildungsverzeichnis.....	III
Tabellenverzeichnis.....	IV
Abkürzungsverzeichnis.....	IV
1. Einleitung.....	1
2. Grundlagen.....	2
2.1. Web 1.0.....	2
2.2. Web 2.0.....	2
2.2.1. Soziale Netzwerke.....	3
2.2.2. Kunden Produktrezensionen / Produktreviews.....	4
2.3. Opinion Mining.....	5
2.3.1. Opinion / Meinung.....	5
2.3.2. Unterscheidung subjektive / objektive Aussage.....	6
2.3.3. Unterschied regulären Meinung und der komparativen Meinung.....	7
2.3.4. Meinungskomponenten.....	7
2.4. Methoden zur Meinungsanalyse.....	8
2.4.1. Machine Learning Ansätze.....	9
2.4.2. Lexika Ansätze.....	11
2.4.2.1. Wörterbuch-basierter Ansatz.....	11
2.4.2.2. Korpus-basierter Ansatz.....	12
2.4. Pre Processing.....	12
2.4.1. Part of speech (POS)-tagging.....	12
2.4.2. Stemming und Lemmatization.....	13
2.4.3. Stop Words entfernen.....	14
2.5. Level der Meinungsklassifizierung.....	14
2.5.1. Dokument Level Meinungsklassifizierung.....	15
2.5.2. Satz Level Meinungsklassifizierung.....	15
2.5.3. Eigenschaft Level Sentiment Klassifikation.....	16

2.6. Eigenschaftskategorien.....	17
2.6.1. Explizite & häufig genannte Eigenschaften	17
2.6.2. Seltene Eigenschaften.....	18
2.6.3. Implizite Eigenschaften	19
3. Aktueller Stand der Forschung.....	19
3.1. Methodik nach Webster und Watson	20
3.2. Identifikation relevanter Literatur.....	21
4. Der letzte Stand der Technik im Opinion Mining.....	25
4.1. Methoden und Modelle für Opinion Mining bei Produktreviews.....	25
4.2. Methoden und Modelle für Opinion Mining bei Sozialen Medien	31
5. Diskussion.....	37
6. Schlussteil.....	42
Literaturverzeichnis.....	V
Internetquellen.....	XV

1. Einleitung

Mit der zunehmenden Menge textueller Daten im Web 2.0 wächst auch die Notwendigkeit der maschinellen Auswertung dieser Informationen, beispielsweise um in Texten geäußerte Meinungen aufzuspüren. Zahlreiche Blogs, Soziale Netzwerke, Forums, Internetshops, News-Plattformen und andere Internetdienste bieten die Möglichkeit der öffentlichen Meinungsäußerung von Besuchern und Kunden bezüglich verschiedenster Themen, Produkte und Dienste. Opinion Mining befasst sich als Disziplin mit der Aufgabe der Auswertung von Meinungsäußerungen und versucht diese zu erkennen, verstehen, analysieren und zusammenzufassen. Unternehmen, Politiker und andere Interessenten erhoffen sich hierdurch die Stimmungen der breiten Masse und entsprechender Zielgruppen zu verstehen und relevante Implikationen herausziehen zu können. Das Ziel der vorliegenden Bachelorarbeit ist es zu präsentieren wie Opinion Mining funktioniert und zu analysieren wie weit die aktuelle Forschung in Bezug auf Opinion Mining und Meinungsanalyse in Sozialen Medien ist. Besonders soll hierbei betrachtet werden, wie weit Erkenntnisse im Bereich der Meinungsanalyse zur Eigenschaftsbasierten Produktbewertung bestehen und diese auf Soziale Medien angewandt werden können.

Der Aufbau der vorliegenden Bachelorarbeit ist entsprechend folgendermaßen strukturiert: In Abschnitt 2 werden zunächst wichtige Grundlagen aufgearbeitet und vorgestellt, um ein Fundament zum Verständnis der präsentierten Konzepte zu kommunizieren. So wird die Geschichte des Internets kurz rekonstruiert und die einher gekehrte Natur der Meinungsäußerung unterstrichen. Meinungen werden auf ihre Unterschiede untersucht und ihre Komponenten werden vorgestellt. Die verschiedenen Methoden und Level der Meinungsanalyse werden zusätzlich präsentiert. In Abschnitt 3 wird der aktuelle Stand der Forschung im Opinion Mining präsentiert. Hierfür wird eine Konzeptmatrix dargestellt, die zur visuellen Präsentation der verschiedenen Konzepte und Ansätze der Autoren dienen soll. In Abschnitt 4 werden die verschiedenen Konzepte, Ansätze und Methoden der Autoren vorgestellt und in Abschnitt 5 kritisch diskutiert. Zusätzlich werden in Abschnitt 5 noch neue Anwendungsmöglichkeiten und Adaptionismöglichkeiten unter verschiedenen Ideen von Autoren diskutiert. Abschließend soll die Arbeit in Abschnitt 6 zusammengefasst werden.

6. Schlussteil

Die vermehrte Anzahl an Meinungsäußerungen im Internet haben u.a. dazu geführt, dass Opinion Mining eine stark untersuchte Disziplin im Text Mining wurde. Forscher untersuchen verschiedenste Aspekte von Meinungsäußerungen und haben sich hierbei an Methoden und Techniken aus Disziplinen wie Natural Language Processing, Information Extraction, Data Mining und Machine Learning bedient. Das Ziel der vorliegenden Bachelorarbeit war es zu präsentieren wie Opinion Mining funktioniert und zu analysieren wie weit die aktuelle Forschung in Bezug auf Opinion Mining und Meinungsanalyse in Sozialen Medien fortgeschritten ist. Besonders sollte hierbei betrachtet werden wie weit Erkenntnisse im Bereich der Meinungsanalyse zur eigenschaftsbasierten Produktbewertung bestehen und diese auf Soziale Medien angewandt werden können. Zu diesem Zweck wurden Arbeiten untersucht, die im Zeitraum zwischen 2002 und 2016 veröffentlicht wurden. Hierbei wurde beobachtet welche Überlegungen, Konzepte und Methoden die Autoren präsentiert haben. Diese umfassen verschiedene Machine Learning-, Lexika- und Hybride-Ansätze. Aggregiert wurden diese Erkenntnisse in einer Konzeptmatrix in Abschnitt 3. Die Arbeit liefert einen allgemeinen Überblick über verwendete Methoden des Opinion Mining und erklärt ihre Funktionalitäten.