

Titel der Arbeit

Unter Umständen positiv? Analyse kontextbedingter Ambivalenz in Kundenmeinungen

Management Summary: Kurzfassung der Arbeit (max. 500 Wörter)

Kaufinteressenten, die sich beispielsweise ein neues Smartphone zulegen möchten, vertrauen nicht mehr allein auf die Angaben oder Beratung des Händlers, sondern recherchieren zunächst, welches Fabrikat und Modell von anderen empfohlen wird. Der Online-Handel hat bereits erkannt, dass Kaufentscheidungen häufig von positiven Produktbewertungen abhängen, reagiert deshalb mit passenden Kaufempfehlungen und bietet über Web-Shops eine Plattform für den Erfahrungsaustausch via Rezensionen an. Studien haben gezeigt, dass 70% der Kaufinteressenten den Empfehlungen und Bewertungen anderer Kunden und Kundinnen im Internet vertrauen. Insbesondere für den Online-Handel ist folgende Fragestellung von Interesse: Welche positiven Gründe führen zu einer Produktweiterempfehlung und welche negativen Urteile verhindern sie? Mit einer automatisierten Textanalyse und Klassifikation dieser Produktbewertungen in positive und negative Aspekte, können Online-Händler entsprechende Analysen ihrer Warenverkäufe in Abhängigkeit der vorliegenden Erfahrungsberichte vornehmen.

Die computergestützte Klassifikation von Meinungen als Ergebnis der automatisierten Erkennung von Stimmungen in Texten - auch als Sentimentanalyse bekannt - stellt Softwareentwickler vor Herausforderungen. Denn bereits kleine Unterschiede in Sätzen können im Hinblick auf die Stimmung einer Aussage einen gravierenden Unterschied verursachen. Ein Vergleich der Aussagen „Das ist gut“, „Das ist weniger gut“ und „Das ist nicht gut“ veranschaulicht dieses Problem. Für eine zuverlässige Sentimentanalyse müssen kontextbedingten Ambivalenzen zuverlässig erkannt und durch Algorithmen korrekt interpretiert werden. Ziel dieser Bachelorarbeit war deshalb die Identifikation sprachlicher Phänomene, welche situationsabhängig (z. B. "weniger gut" vs. "weniger schlecht") einen negativen oder positiven Einfluss auf die Wertungen von Kunden (z. B. "gut", "schlecht") haben. Insbesondere wurden hierbei Verneinungen, Ironie und Sarkasmus sowie Verstärkungen und Abschwächungen von Aussagen analysiert. Ein weiteres Ziel dieser Bachelorarbeit bestand in der praktischen Umsetzung dieser gewonnenen Erkenntnisse. Es wurde eine Software zur Klassifikation von Stimmungen in Kundenmeinungen entwickelt, welche erstmals eine automatisierte Echtzeit-Statistik über den Textinhalt von Produktbewertungen ermöglicht. Die Resultate der Software - angewendet auf klassische Produktbewertungen (SFU Review Corpus) - bestätigen, dass die Berücksichtigung sprachlicher Phänomene in Bewertungstexten in einen positiven Einfluss auf die Leistungsfähigkeit von Algorithmen zur Sentimentanalyse hat und so eine feingranulare Bestimmung von positiven und negativen Aspekten in Kundenmeinungen erst möglich wird.

Stellungnahme zu den folgenden bewertungsrelevanten Aspekten der Arbeit

1. Relevanz für den Handel: Ist das Thema bereits im Handel erforscht worden? Welche neuen Erkenntnisse für den Handel hält die Arbeit bereit? (max. 200 Wörter)

Zwar gibt es seit ein paar Jahren Dienstleister wie z.B. TrustYou, die sich insbesondere in der Tourismusbranche auf die kommerzielle maschinelle Inhaltsauswertung von Hotelbewertungen spezialisiert haben, um das Online-Reputations-Management ihrer Kunden nachhaltig zu verbessern. Allerdings gibt es nichts Vergleichbares im B2C-Bereich für den Handel. Bereits jetzt werden im E-Commerce Angaben über vergebene Sterne oder Noten für Produkte genutzt, doch fehlt weiterhin die computergestützte Textanalyse zur Ursachenforschung für besonders gute oder schlechte Bewertungen in den dazugehörigen Erfahrungsberichten. Persönliches Review-Management (d.h. manuelle Auswertung von Produktrezensionen) ist angesichts der Fülle an Online-Kundenmeinungen - wie sie beispielsweise bei Amazon digital vorliegen - nicht mehr denkbar. Markus Dollmann hat in seiner Bachelorarbeit empirisch anhand der von ihm entwickelten Software zur Erkennung von positiven und negativen Kundenkommentaren in unterschiedlichen Produktkategorien (Bücher, Computer, Auto, Küche & Haushalt, Telefonie, Hotel, Musik und Film) gezeigt, dass Online-Händler nicht dieselbe Konfiguration zur Analyse von Produktbewertungen für ihren kompletten Produktkatalog einsetzen können, sondern dass sich die Leistungsfähigkeit der Software bei verschiedenen Produktkategorien gravierend unterscheidet. Während ohne Anpassung der Softwareeinstellungen im besten Fall 84% der Kundenmeinungen ("Auto") richtig klassifiziert werden konnten, waren es im schlechtesten Fall lediglich 62% ("Musik").

2. Umsetzbarkeit im Handel: Wo liegt der Anwendungsnutzen für den Handel? (max. 200 Wörter)

Die von Markus Dollmann entwickelte Software analysiert nicht nur die Grundstimmung in Produktbewertungskommentaren automatisiert und stellt damit dem Handel entsprechende Statistiken zur weiteren Auswertung zur Verfügung (z.B. für Korrelationen von Weiterempfehlungen und Verkaufszahlen oder von Kritik und sogenannten "Ladenhütern"). Außerdem widmet sich seine Arbeit dem meist zu wenig beachteten Thema der Ironie-Erkennung in Produktbestellungen wie z.B. "super Lieferzeit :-(", die fälschlicherweise häufig unerkannt bleibt und fatale Folgen haben kann. Anhand dieser Daten können Verkaufsränge dynamisch angepasst, Rabattaktionen lanciert oder Rücksendequoten prognostiziert werden. Insbesondere deckt seine Software auch durch ihre Echtzeitanalyse auf, wenn sich negative Erfahrungsberichte und Kommentare zu einem Produkt häufen, ohne das der Händler diese ständig mitlesen müsste. Auf diese Weise wird dieser gewarnt, dass es Probleme mit dieser Ware gibt (z.B. Defekte) und dieser mit einer Welle an Retouren rechnen oder die komplette Charge an den Hersteller zurücksenden oder dieses Produkt aus dem Sortiment nehmen muss. Angesichts des umfangreichen Sortiments von Online-Händlern und der Fülle an Produktkommentaren durch deren Kunden würde der Personalansatz für eine gleichwertige manuelle Auswertung höhere Kosten verursachen als der monetäre Nutzen wäre. Somit fungiert die Software von Markus Dollmann als ideales, schnelles und kostengünstiges Warnsystem, das kritische Produkte anhand negativer Bewertungen aufspürt und dem Händler freie Hand bei der Reaktion darauf lässt.

3. Innovationskraft der Idee (max. 200 Wörter)

Die innovative Idee, welche hinter der Software steckt, ist ihre dynamische Adaptivität der Stimmungsanalyse an die jeweilige Produktkategorie. Bisher musste sich ein Händler beispielsweise eine Software zur Auswertung der Kundenmeinungen für seinen Web-Shop kaufen und konnte nur konfigurieren, in welcher Sprache seine Kunden und Kundinnen die Bewertungskommentare verfassen (z.B. auf Deutsch oder Englisch). Nun wechselt die Konfiguration automatisch in Abhängigkeit der Produktkategorie bei der maschinellen Textanalyse, so dass z.B. "gruselig" in Zusammenhang mit einem Horrorfilm als positiv während "gruselig" in Kombination mit Spielzeug als negativ in einem Bewertungstext eingestuft wird. Dieser Herausforderung hat sich Markus Dollmann im Rahmen seiner Bachelorarbeit meisterhaft gestellt, was ein nicht zu unterschätzender Aufwand in der Entwicklungs- und Erprobungsphase ist, bis letztendlich die optimale Konfiguration pro Produktkategorie experimentell auf den zugrundeliegenden Testdaten ermittelt und nun auf jeden beliebigen Online-Shop mit heterogenem Sortiment (wie z.B. Amazon) angewendet werden kann.

4. Angewandte wissenschaftliche Methode (max. 100 Wörter)

Es wurde eine Software zur Erfassung von Stimmungen in Kundenmeinungen, unter Berücksichtigung der identifizierten linguistischen Phänomene programmiert. Dabei wurde ein lexikonbasierter Ansatz verfolgt, welcher Erkennungs- und Behandlungsmöglichkeiten für Negationen, Ironie und Sarkasmus, Verstärkungen und Abschwächungen, modale Aussagen, sowie Konditionalsätzen und Fragen bereitstellt. Wie der Titel der Bachelorarbeit "Unter Umständen positiv? Analyse kontextbedingter Ambivalenz in Kundenmeinungen" verrät, wurde kein statisches sondern ein sich an seine Umgebung anpassendes probabilistisches kontextsensitives Klassifikationsmodell entwickelt, das abhängig von der Produktkategorie passende Parameter einsetzt, um möglichst fehlerfrei positive oder negative Stimmungen zu identifizieren.