

Identifikation kognitiver Effekte in Online-Bewertungen

Valentina Stuß & Michaela Geierhos, Universität Paderborn

Wie und ob die Meinungsbildung das Verhalten der Verbraucher¹ und Dienstleistungsnehmer beeinflusst und wie Wissensverarbeitung bei Menschen abläuft, sind die Fragen, mit denen sich Sozial- und Kognitionspsychologen beschäftigen (Stürmer, 2009; Wolff, 1993). Fehlerhafte Urteile bei der Meinungsbildung, auch kognitive Effekte genannt, werden in bestimmten Situationen wie der Bewertung von Produkten oder in Anspruch genommenen Leistungen besonders deutlich. Sie lassen sich in der Wissensorganisation und Gedächtnisstruktur des menschlichen Gehirns begründen. Mögliche Gründe für verzerrte Meinungen sind falsche Wahrnehmungen, die auf Basis kognitiver Informationsorganisation und -verarbeitung, aber auch aufgrund emotionaler Betroffenheit und/oder persönlicher Voreingenommenheit entstehen können.

Im Web 2.0 bieten Bewertungsportale Internetnutzern eine Plattform für den multi-direktionalen Erfahrungsaustausch an. Zahlreiche Rezensionen der Konsumenten, Kunden, Patienten oder anderer Zielgruppen bilden einen riesigen Datenbestand und stellen eine wertvolle Informationsquelle dar, die eine wissenschaftliche Auseinandersetzung von mehreren Seiten erfordert. Nicht nur Sozial- und Kognitionspsychologen, sondern auch Computerlinguisten, haben ein wissenschaftliches Interesse daran, Erkenntnisse mittels Informationsextraktion und Stimmungserkennung (Kim & Hovy, 2004; Hatzivassiloglou & McKeown, 1997) aus diesen digitalen Textsammlungen zu gewinnen. Eine interdisziplinäre Analyse psychologischer Phänomene, auch kognitive Effekte genannt, im Kontext schriftlich geäußerter Online-Meinungen wurde jedoch bisher nicht durchgeführt. Darum ist es das Ziel, mögliche kognitive Effekte zu definieren und computergestützt zu identifizieren.

Die Abweichung von der Norm

Je nach Art und Weise der dargebotenen Informationen, können diese unterschiedlich wahrgenommen werden, was zu den sogenannten Wahrnehmungsdefekten oder Informationspathologien (kognitiven Effekten) und somit zu fehlerhaften Meinungsbildern führen kann (Schneider, 2013). Ein kognitiver Effekt liegt immer dann vor, „wenn relevante Informationen nicht beschafft, nicht (korrekt) übermittelt, nicht produziert oder nicht (korrekt) verarbeitet werden, obwohl dies eigentlich möglich wäre.“ (Schneider, 2013:13f.). Es gibt eine Reihe von kognitiven Effekten, wie z. B. den Framing-Effekt, der dann auftritt, wenn inhaltsgleiche, aber auf verschiedene Weise dargestellte Informationen unterschiedlich bewertet werden. Insbesondere in Online-Bewertungen werden allgemeine Verfälschungen wie der Framing-Effekt oder auch Negativitätsverzerrungen auftreten.

„It does not appear possible today to group all of the phenomena that have been qualified as cognitive biases under one and the same definition. [...] As such, a bias is detected when derivation from norm is observed. [...] Occasional and accidental errors are obviously not part of the issue of cognitive biases.“ (Caverni et.al., 1990:7f.)

Um kognitive Effekte letztendlich automatisiert in Online-Bewertungen identifizieren zu können, stellt sich die Frage, was als Norm in einer Bewertung gilt. Da sich diese meist aus einem Kommentar und einem Notenwert zusammensetzt, ist zu vermuten, dass Textinhalt und Note kongruent sind und damit in ihrer Polarität (positiv, neutral, negativ) übereinstimmen. Jedoch können aufgrund von zufälligen Individualfehlern (wie z. B. Verständnisprobleme beim Benotungssystem eines Bewertungsportals) nicht alle Nichtübereinstimmungen grundsätzlich als kognitive Effekte interpretiert werden. Logisch für

¹ Aus Gründen der leichteren Lesbarkeit wird auf eine geschlechtsspezifische Differenzierung verzichtet. Entsprechende Begriffe gelten im Sinne der Gleichbehandlung für beide Geschlechter.

eine Norm wäre ebenfalls, dass die Mehrheit an Bewertungen diese Norm erfüllen sollte. Das bedeutet, dass nur eine kleine Anzahl Bewertungen von kognitiven Effekten überhaupt betroffen sein können.

Als kognitive Effekte bei Online-Bewertungen werden somit diejenigen Bewertungsfehler bezeichnet, die sowohl von der Erwartungsnorm abweichen als auch bewusst von den Rezensenten gemacht wurden, um ein bestimmtes Ziel zu erreichen. Diese führen zu einer verzerrten Meinung, die sich durch einen Widerspruch im Freitext zur entsprechend vergebenen Note aufspüren lässt. Insbesondere um Inkonsistenzen im Bewertungsverhalten aufzudecken, ist es erforderlich ausreichend repräsentative Meinungen als Referenzkorpus zu haben. Nur so kann sichergestellt werden, dass ein Vergleich der zu untersuchenden Meinungen mit der zu erwartenden Norm für jede Ausprägung eines kognitiven Effekts durchgeführt werden kann.

Die Identifikation kognitiver Effekte

Datenbasis

Dieser Studie liegen zufällig ausgewählte Datensätze der Portale jameda und DocInsider aus den Jahren 2009 bis 2013 zugrunde. Das Korpus umfasst 217.841 individuelle Erfahrungsberichte, die allesamt nach einem Arztbesuch in Deutschland verfasst und online gestellt wurden. Jede Arztbewertung besteht dabei aus einer Überschrift, dem eigentlichen Text und bis zu 17 verschiedenen numerischen Bewertungskategorien (u.a. Behandlung, Vertrauensverhältnis, Wartezeit, Barrierefreiheit).

Methodik

Bei der Identifikation kognitiver Effekte geht es zunächst darum, bestimmte Muster, die Meinungen zu den vordefinierten Kategorien ausdrücken, und deren Polarität (Turney, 2002) zu erkennen. Diese Muster sind aus den frei formulierten Texten der Nutzer zu extrahieren. Im Rahmen einer Sentiment Analyse werden alle in der jeweiligen Meinung angesprochenen Themen identifiziert und entsprechend ihrer Polarität klassifiziert (Kim & Hovy, 2006). Diese Muster werden mittels regelbasierter, morpho-syntaktischer Verfahren, sogenannten lokalen Grammatiken (Gross, 1997) extrahiert und mithilfe von Transduktoren annotiert. Lokale Grammatiken sind stets modular aufgebaut, was in Abbildung 1 durch eine Subgrammatik zur Erkennung von Zeitangaben mit positiver Polarität der Kategorie „Wartezeit (Praxis)“ illustriert wird.

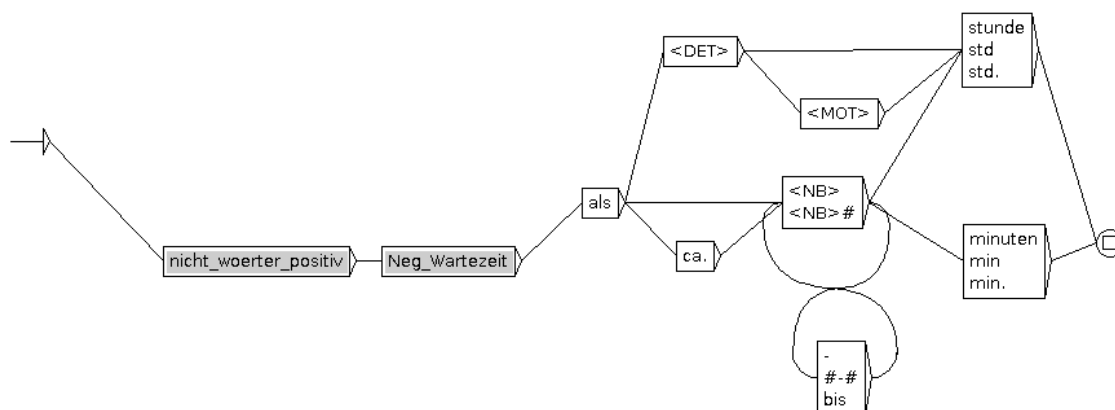


Abbildung 1: Lokale Grammatiken zur Erkennung der Wartdauer in der Arztpraxis

Durch Suche der in dieser Grammatik beschriebenen Muster entsteht auf der hier verfügbaren Datenbasis folgende Konkordanz (vgl. Abbildung 2). Hierfür werden unter anderem Lexika miteinbezogen, die beispielsweise in der in Abbildung 1 dargestellten lokalen Grammatiken über <DET> eingebunden werden, und hier auf alle Determinatoren zugreifen können. Darüber hinaus werden reguläre Ausdrücke zur Erkennung von Zahlen mittels <NB> und Wörter mittels <MOT> eingesetzt. Zudem stellen Lexika wichtige Ressourcen bei der Analyse der Sentiments dar und enthalten Polaritätswörter wie z. B. positive oder negative Adjektive, oder das aus dem Korpus selbst gewonnene Fachvokabular.

kommen und musste auch nicht länger als 10 Minuten im Wartezimmer warten. Die
 abfertigung! Man sitzt nicht länger als 10 Minuten im Wartezimmer, im Behandlu
 en und musste meistens nicht länger als eine halbe Stunde warten. Ich habe mic
 Orthopäden und musste nie länger als 15 bis 20min Warten und wurde auch angem
 - sofort dran. Mussten nie länger als 5 Minuten warten. [column name="Datum"]1
 Ich musste ohne Termin nie länger als eine Stunde warten. [column name="Datum"
 konnte, aber wir haben nie länger als eine viertel stunde oder zwanzig minuten
 Im Wartezimmer saß ich noch nie länger als 10 min. Ich werde immer herzlich vo
 ndlich und ich musste noch nie länger als 30 min in der meist gut besuchten p
 : in Grenzen. Ich habe noch nie länger als 30 Minuten gewartet (auch ohne Term

Abbildung 2: Auszug aus der Konkordanz positiver Äußerungen über Wartezeiten in der Praxis

Iterative Evaluation

Begleitend zur Entwicklung des Verfahrens wird eine iterative Evaluation durchgeführt, die als Basis für Entwurfsentscheidungen oder Korrektur derselben dient. Deshalb wird sowohl die Erkennung der Bewertungskategorien in den Bewertungskommentaren als auch die Identifikation der kognitiven Effekte in mehreren Schritten erfolgen. Nach jeder Etappe werden Zwischenevaluationen durchgeführt, um so die Erkennungsprobleme zu identifizieren und die lokalen Grammatiken zu verbessern.

Überbewertung: Positiv + Negativ = Positiv ?

Am Beispiel eines kognitiven Effekts, der aufgrund der Überbewertungen einer Kategorie durch die Hervorhebung einer anderen entsteht, soll illustriert werden, inwiefern dieser automatisiert in Online-Bewertungen mittels oben genannter Methodik identifiziert werden kann. Wie die Bezeichnung des Effekts bereits sagt, geht es hier um eine Bewertungskategorie, deren Benotung möglicherweise positiver ausfällt als dies in der Wirklichkeit der Fall ist. Als exemplarische Kategorie wird „Wartezeit (Praxis)“ herangezogen. In diesem Auszug aus der Datenbasis wurden alle Aussagen und Noten der Patienten zu den Wartezeiten in Arztpraxen durch lokale Grammatiken annotiert. Ein kognitiver Effekt der Überbewertung liegt in diesem Fall vor, wenn die Wartezeit von Patienten besser benotet als die Situation beim Warten in der Praxis von ihnen tatsächlich wahrgenommen wird.

Herr Dr. Brachvogel nimmt sich für seine Patienten ausreichend Zeit, was gelegentlich auch dazu führen könnte, dass man <N Kategorie="WZP_4">länger warten</N> muss!

In dieser Bewertung war die Note 1,0 für die Kategorie „Wartezeit (Praxis)“. Dabei ist auffällig, dass die Wartezeit zwar als „lang“ wahrgenommen wird, jedoch vorher eine andere Kategorie (in diesem Fall „Genommene Zeit“) positiv hervorgehoben wird. So ist es meist bei diesem Effekt der Fall, dass entweder im gleichen Satz oder in unmittelbarer Nähe von dem betreffenden negativen Muster zur Erkennung der verbrachten Zeit im Wartezimmer eine andere Kategorie (oft „Genommene Zeit“, „Behandlung“ oder „Freundlichkeit“) positiv hervorgehoben wird. In anderen Fällen werden auch positive Äußerungen zu mehreren Kategorien aufgelistet und damit zum Schluss die langen Wartezeiten begründet, was ebenfalls teilweise zum genannten Effekt führen kann.

Obwohl die angesprochene Wahrnehmung der Wartezeitdauer und somit die Polarität der obigen Aussage offensichtlich zu sein scheint, könnten Kritiker behaupten, dass die Polarität der Kategorie „Wartezeit (Praxis)“ in dem genannten Kontext nicht negativ (N), sondern positiv (P) wäre. Somit würde ebenfalls die Existenz des hier angesprochenen Effekts infrage gestellt werden. Als Gegenargument sei angemerkt, dass sich bei der Korpusanalyse zeigt, dass es eine Vielzahl ähnlicher Äußerungen gibt, bei denen sich die Patienten zu langen Wartezeiten bereiterklären und diese durch die Fachkompetenz oder Freundlichkeit des Arztes rechtfertigen, jedoch die „Wartezeit (Praxis)“-Kategorie selbst negativ benoten. Dies spiegelt wiederum die bei der allgemeinen Definition von kognitiven Effekten angesprochene Erwartungsnorm wider, von der die Bewertungsfehler abweichen.

Insgesamt konnten in ungefähr 22% der Bewertungen, in denen Äußerungen zur Wartezeit in der Arztpraxis gemacht werden, kognitive Effekte identifiziert werden. Dementsprechend handelt es sich bei 78% um keine kognitiven Effekte, wie in folgendem Beispiel illustriert wird:

Immer freundlich, kompetent, empathisch [column Name="Bewertung"] Die <N Kategorie="WZP_4">langen Wartezeiten</N> nimmt man da gerne in Kauf. [...] [column Name="b_WartezeitPraxis"]5.0[/column]

Bestehende Herausforderungen

Der hier vorgestellte interdisziplinär ausgerichtete Forschungsansatz schlägt eine Brücke zwischen den Kognitionswissenschaften und der Computerlinguistik, indem sozialpsychologische Phänomene aus sprachwissenschaftlicher Perspektive maschinell erschlossen werden. Das Ziel dieser Studie ist es, kognitive Effekte über sprachliche Muster zu definieren und in Patientenbewertungen automatisch identifizieren zu können.

Die ersten Probleme und Erkenntnisse der durchgeführten Studie lassen sich wie folgt zusammenfassen:

- Die Genauigkeit der Mustererkennung ist besser als deren Abdeckung; bei der Identifikation von kognitiven Effekten ist dies umgekehrt der Fall. Je höher die Extraktionsgenauigkeit, desto besser und präziser ist die Identifikation der kognitiven Effekte.
- Die Problematik der Mustererkennung hat hauptsächlich mit der Sprachspezifik der Kundenbewertungen zu tun, wobei die Sprachstile von einer gesprochenen Sprache bis hin zu literarischen Ausdrücken reichen.
- Die Verbesserung der Extraktionsqualität kann durch die Erweiterung bzw. Aufnahme zusätzlicher Muster in lokale Grammatiken erreicht werden. Zahlreiche Phänomene müssen kategorienübergreifend behandelt werden, wie z. B. die Extraktion ironischer Aussagen, Emoticons oder Smileys, Tippfehler, Polaritätsbestimmung auf der Satzebene usw.

Literatur

- Caverni, J. P., Fabre, J. M., & Gonzalez, M. (Hrsg.) (1990): *Cognitive biases*. New York u.a.: Elsevier. S. 7–12, 59–68.
- Gross, M. (1997): *The Construction of Local Grammars*. In E. Roche und Y. Schabès (Hrsg.): *Finite-State Language Processing*. Language, Speech und Communication, Cambridge, Mass.: MIT Press. S. 329–354.
- Hatzivassiloglou, V., & McKeown, K. R. (1997): *Predicting the semantic orientation of adjectives*. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics. S. 174–181.

- Kim, S. M., & Hovy, E. (2004): *Determining the sentiment of opinions*. In Proceedings of the 20th international conference on Computational Linguistics. Association for Computational Linguistics. S. 1367–1374.
- Kim, S.-M. & Hovy, E. (2006): *Automatic identification of pro and con reasons in online reviews*. In Proceedings of the Poster Session at the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics, Sydney, Australia, 17.-21. Juli 2006, S. 483–490.
- Schneider, S. (2013): *Das 3-Dimensionsmodell der Wissensrekonstruktion: A priorische Sicherstellung der Güte generierten Wissens*. Forschungsbericht. S. 5–22. http://www.fhkiel.de/fileadmin/data/wirtschaft/dozenten/schneider_stephan/Science/ResearchReport/Schneider_2013_Informationspathologien.pdf (16.09.2014).
- Stürmer, S. (2009): *Sozialpsychologie*. München: Reinhardt. S. 11–16, 69–90, 165–217.
- Turney, P. (2002): *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews*. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Penn., 7.-12. Juli 2002, S. 417–424.
- Wolff, D. (1993): *Der Beitrag der kognitiv orientierten Psycholinguistik zur Erklärung der Sprach- und Wissensverarbeitung*. In: Gienow, W. / Hellwig, K. (Hrsg.): *Prozeßorientierte Mediendidaktik im Fremdsprachenunterricht*. Frankfurt am Main u.a.: Lang. S. 27–41.