



King Saud University
Journal of King Saud University – Engineering Sciences

www.ksu.edu.sa
www.sciencedirect.com



ORIGINAL ARTICLE

A survey on sentiment analysis challenges

Doaa Mohey El-Din Mohamed Hussein

Faculty of Computers and Information, Cairo University, Cairo, Egypt

Received 12 November 2015; accepted 18 April 2016

KEYWORDS

Sentiment analysis;
Text analysis;
Sentiment analysis chal-
lenges;
Sentiments;
Review structure;
Accuracy

Abstract With accelerated evolution of the internet as websites, social networks, blogs, online portals, reviews, opinions, recommendations, ratings, and feedback are generated by writers. This writer generated sentiment content can be about books, people, hotels, products, research, events, etc. These sentiments become very beneficial for businesses, governments, and individuals. While this content is meant to be useful, a bulk of this writer generated content require using the text mining techniques and sentiment analysis. But there are several challenges facing the sentiment analysis and evaluation process. These challenges become obstacles in analyzing the accurate meaning of sentiments and detecting the suitable sentiment polarity. Sentiment analysis is the practice of applying natural language processing and text analysis techniques to identify and extract subjective information from text. This paper presents a survey on the sentiment analysis challenges relevant to their approaches and techniques.

© 2016 Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Sentiment analysis (Basant et al., 2015) uses the natural language processing (NLP), text analysis and computational techniques to automate the extraction or classification of sentiment from sentiment reviews. Analysis of these sentiments and opinions has spread across many fields such as Consumer information, Marketing, books, application, websites, and Social. Sentiment analysis becomes a hot area in decision-making (Tawunrat and Jeremy, 2015) (Matthew et al., 2015). Hundreds of thousands of users depend on online sentiment reviews. 90% of customer's decisions depended on

Online Reviews in April 2013 (Ling et al., 2014). The main goal of analyzing sentiment is to analyze the reviews and examine the scores of sentiments. This analysis is divided into many levels (Thomas, 2013): document level (Ainur et al., 2010), sentence level (Noura et al., 2010), word/term level (Nikos et al., 2011) or aspect level (Haochen and Fei, 2015). The sequence processes are of sentiment analysis evaluation and detection of the sentiment polarity (Khairullah et al., 2014). This paper focuses on the most important challenges in sentiment evaluation phase that they have a significant effect in sentiment score and polarity detection. The evaluation sentiment drawbacks that Reflected in language coverage. This paper summarizes keys of sentiment challenges (Sujata and Partee, 2014) (Vinodhini and Chandrasekaran, 2012) (Arjun et al., 2012) with respect to the type of review structure. It also divides the challenges into two types to ease to deal with them and focus on the degree of accurate meaning. This research discusses these sentiment challenges, the factors affecting them, and their importance. As a result, a

E-mail address: d.mohey@alumni.fci-cu.edu.eg

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

<http://dx.doi.org/10.1016/j.jksues.2016.04.002>

1018-3639 © 2016 Production and hosting by Elsevier B.V. on behalf of King Saud University.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Please cite this article in press as: Hussein, D.-M.E.D.M. A survey on sentiment analysis challenges. Journal of King Saud University – Engineering Sciences (2016), <http://dx.doi.org/10.1016/j.jksues.2016.04.002>

Table 1 Study for the sentiment challenges relevant to review structure.

Ref. No	Domain oriented	Challenge type	SA challenge	Review structure
Bas et al. (2011)	No (40 different topics)	Theoretical	Negation	Semi-structured adjectives only
Alexander et al. (2011)	Yes, movie reviews	Theoretical	Negation	Un-structured
Amna (2012)	N broader sense domain	Theoretical	Negation	Semi-structured nouns/adjectives/verbs and adverbs-clauses and phrases
Maral (2011)	Yes movies	Theoretical	Negation + domain dependence	Semi-structured adverbs, adjectives
Lifeng (2009)	Yes health/medical domain	Theoretical	Negation	Semi-structured
Robert (2013)	Y	Theoretical and Technical	Negation + bipolar words	Semi-structured, sentences or topics documents
Michael et al. (2010)	N	Theoretical and Technical	Negation + entity features/keywords	Structured or semi-structured
Emanuele et al. (2012)	N	Theoretical + Technical	Negation + huge lexicon	Semi-structured,
Stanislav (2013)	Yes	Theoretical	Domain dependence	Unstructured conjunction with predefined taxonomy of emotional terms
Yulan et al. (2011)	N mutli-domain	Theoretical	Domain dependence	Semi-structured
Hiroshi and Tetsuya (2006)	N	Theoretical	Domain dependence	Structured, objectives expressions
Bing and Liang (2014)	Y	Theoretical	Domain dependence	Un-structured, twitter
Alexandra et al. (2013)	Y	Theoretical	Domain dependence	Structured, news articles
Fangtao et al. (2010)	Y	Theoretical	Domain dependence	Un-structured, online customers reviews
Ouyang et al. (2014)	N	Theoretical	Domain dependence	Unstructured, emotion reviews
Fangtao et al. (2011)	Y, product reviews	Theoretical	Spam and fake detection	Unstructured
Xia et al. (2014)	Y, social media	Theoretical	Spam and fake detection	Unstructured
Qingxi and Ming (2014)	N, online customers reviews	Theoretical	Spam and fake detection	Unstructured
Ahmed et al. (2010)	Y, ecommerce and online security	Theoretical	Spam and fake detection	Semi-structured
Myle et al. (2011)	N, online customer reviews	Theoretical	Spam and fake detection	Unstructured,
Theodoros (2012)	N	Theoretical	Spam and fake detection + negation	Semi-Structured
Alexandra and Ralf (2009)	Y, online news reviews	Theoretical	World knowledge	Semi-structured, unstructured
Marina et al. (2014)	Y, the game on amazon mechanical turk	Theoretical	World knowledge	Unstructured
Svetlana et al. (2014)	Y, tweets	Theoretical	NLP overheads (Short Abbreviations)	Unstructured
Jiliang et al. (2012)	Y, facebook, and twitter	Theoretical	NLP overheads (Short Abbreviations)	Unstructured
Yanfang et al. (2015)	N	Theoretical	NLP overheads (Ambiguity)	Semi-structured
Yunfang and Miaomiao (2010)	N	Theoretical	NLP overheads (Ambiguity)	Structured, adjectives only
Duyu et al. (2014)	Y, social media	Theoretical	NLP overheads (Emotions)	Unstructured
Saif and Peter (2010)	N	Theoretical	NLP overheads (Emotions)	Unstructured
Christine et al. (2013)	Y, tweets	Theoretical	NLP overheads (Sarcasm) + negation	Unstructured
Nathan and Ruihong (2013)	Y, tweets	Theoretical	NLP overheads (Sarcasm)	Unstructured
Subhabrata and Pushpak (2012)	Y, products	Technical	Extracting features or keyword	Semi-structured
Gizem et al. (2012)	Y, trip advisor	Technical	Extracting features or keyword	Semi-structured

Download English Version:

<https://daneshyari.com/en/article/11020664>

Download Persian Version:

<https://daneshyari.com/article/11020664>

[Daneshyari.com](https://daneshyari.com)