

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/301649355>

A Survey on Sentiment Analysis Challenges

Article in *Journal of King Saud University - Engineering Sciences* · April 2016

DOI: 10.1016/j.jksues.2016.04.002

CITATIONS

639

READS

12,155

1 author:



Doaa Mohey El-Din

Cairo University

32 PUBLICATIONS 862 CITATIONS

SEE PROFILE

Accepted Manuscript

Original Articles

A Survey on Sentiment Analysis Challenges

Doaa Mohey El-Din Mohamed Hussein

PII: S1018-3639(16)30007-1

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

Reference: JKSUES 207

To appear in: *Journal of King Saud University - Engineering Sciences*

Received Date: 12 November 2015

Accepted Date: 18 April 2016



Please cite this article as: Mohey El-Din Mohamed Hussein, D., A Survey on Sentiment Analysis Challenges, *Journal of King Saud University - Engineering Sciences* (2016), doi: <http://dx.doi.org/10.1016/j.jksues.2016.04.002>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A Survey on Sentiment Analysis Challenges

Corresponding Author:

Doaa Mohey El-Din Mohamed Hussein

Faculty of computers and information

Cairo University

Cairo, Egypt

d.mohey@alumni.fci-cu.edu.eg

A Survey on Sentiment Analysis Challenges

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 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 faced 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.

Keywords— *Sentiment analysis, text analysis, sentiment analysis challenges, Sentiments, Review structure, accuracy.*

I. 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. Analyzing of these sentiments and opinions enter into many fields such as Consumer information, Marketing, books, application, websites, and Social. Sentiment analysis becomes a hot area in decision-making (Tawunrat & Jermey, 2015) (Matthew et al., 2015). Hundreds of thousands of users depend on online sentiment reviews. 90% of customer's decisions depends 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 divides 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 & Fei, 2015). The sequence processes of sentiment analysis evaluation and detect 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 & Parteek, 2014) (Vinohini & Chandrasekarn, 2012) (Arjun & Bing, 2012) with respect 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 the importance of them. As the result, a large number of studies and research have monitored the trending new research increasing year by year. In this research, trying to achieve to the most challenges faced sentiment evaluation to be useful for researchers and facilitate the relationships of them. The rest of this paper is organized as follows: Section 2, Empirical Study, section 3, Discussion, and Section 4, conclusion and proposes directions for future work.

II. EMPIRICAL STUDY

This research is based on two comparisons among the forty-seven previous research in sentiment analysis to choose the suitable challenge for each research and to show the effects of them on the sentiment accuracy (Ismat & Ali, 2011). First comparison discusses the relationship between the sentiment analysis challenges and review structure. Second comparison examines a significance of solving the sentiment challenges to improve accuracy.

First comparison: is between the thirty-seven research papers. The target of this comparison is recognizing the relationship between the sentiment challenges and review structure and how to effect on the sentiment results. Sentiment review structure becomes an essential factor which effects on selecting the important challenges should the researchers face in their research. By assuming the types of reviews format as in the following:

- A. **Structured Sentiments** are found in formal sentiment reviews, but it targets the formal issues as books or research. Because the writers are professional and writing sentiments or notices about the scientific or fact issues.
- B. **Semi-Structured Sentiments** are lied on the range between the formal structured sentiments and unstructured sentiments. These require understanding several issues about reviews. This type depends on Pros and Cons are listed separately by the writer and the contents of Pros and Cons are usually short phrases.
- C. **Unstructured Sentiments** are an informal and free text format, the writer don't follow any constraints (Arjun, et al., 2013a). There is no formal separation of Pros and Cons and the content may consist of several sentences, where each sentence contains features and/or opinions. For the example below that unstructured reviews have the potential to provide more abundant and detailed opinion information than its counterpart (Arjun, et al., 2013b). **Explicit Feature:** If a feature f appears in the segment/chunk of a review sentence, the feature is called an explicit feature of a product. For example; in the segment the picture is wonderful, picture is an explicit feature. **Implicit Feature:** If a feature f does not appear in the

segment of review, but is implied, the feature is called an implicit feature of a product. For example; in the segment it is very expensive, the price is an implicit feature, and expensive is a feature indicator. With respect to the importance of sentiment analysis, so this survey discusses the relationship between the review structure and sentiment analysis challenges. We examine the sentiment challenge that appears more with the type of sentiment structure.

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
(Emauele et al.,2012)	N	Theoretical + Technical	Negation+ huge lexicon	Semi- Structured,
(Stansilav, 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 & Tetsuya,2006)	N	Theoretical	Domain Dependence	Structured, objectives expressions
(Bing& Linag, 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 & 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 & 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 et al.,2010)	N	Theoretical	NLP overheads (Ambiguity)	Structured, adjectives only
(Duyu et al.,2014)	Y, social media	Theoretical	NLP overheads (Emotions)	Unstructured,
(Saif & Peter, 2010)	N	Theoretical	NLP overheads (Emotions)	Unstructured,

(Christine et al., 2013)	Y, tweets	Theoretical	NLP overheads(Sarcasm) + negation	Unstructured,
(Nathan & Ruihong, 2013)	Y, tweets	Theoretical	NLP overheads(Sarcasm)	Unstructured
(Subhabrata & Pushpak, 2012)	Y, products	Technical	Extracting keyword Features or	Semi-structured
(Gizem et al., 2012)	Y, Trip advisor	Technical	Extracting keyword Features or	Semi-structured
(Mus'ab et al., 2012)	Y , with aspect level	Technical	Extracting keyword Features or	Semi-structured
(Ivan et al., 2013)	Y, movie and product domains	Technical	Bi-polar words	Unstructured,
(Lucie et al., 2015)	Y, social media	Technical	Bi-polar words	Unstructured,
(Sathya & Kaladevi, 2014)	Y, tweets	Technical	Bi-polar words	Unstructured,
(Chetan & Atul, 2014)	Y, tweets	Technical	Huge Lexicon	Unstructured,
(Erik, 2013)	N, online collaborative media.	Technical	Huge Lexicon	Structured, unstructured
(Andrius et al., 2012)	Y, CNET, IMDB movie reviews	Technical	Huge Lexicon	Semi-structured
(Doaa et al., 2015)	Y, scientific papers	Technical + Theoretical	Lexicon+ Feature extraction+ Negation+ world knowledge	Structured

Table.1 illustrates that the comparison between the forty-one papers in sentiment analysis challenges. The comparison's results declare that there is an essential factor important relevant to the review structure. This factor is domain oriented, that requires having an orientation of the topic domain and its features or keywords to determine the fitting challenge for the research or application. The comparison relies on the relationship between the domain and the review structure. Another result the negation is the most important challenge which has the greatest impact in any sentiment analysis and evaluation whether structured, semi-structured or unstructured review. But the comparison shortcoming requires updatable research constantly to reach the suitable challenges easily and quickly.

Second comparison explains the summary of sentiment challenges and how to improve the accuracy of each one based on the previous works. Its goal is identifying the most significance challenges in sentiment and how to improve its results relevant to the used techniques. Figure.1 explains the proposition of using the techniques with respect to the sentiment analysis (SA) challenges types (Theoretical or technical). According to the comparison between the twenty-six research papers in the Table.2 identifies the usage of each technique. The theoretical challenges use many techniques to improve the results with solving the selective sentiment challenges. The highest techniques usage in theoretical type are part-of-speech (POS) tagging and lexicon-based techniques. Bag-of-words (BOW) technique is the second technique. And the last one is Maximum entropy (ME) technique. But the results different in technical sentiment challenges type, the highest usage technique is n-gram technique, because it is based on phrases and expressions. And the least technique usage here is lexicon-based technique.

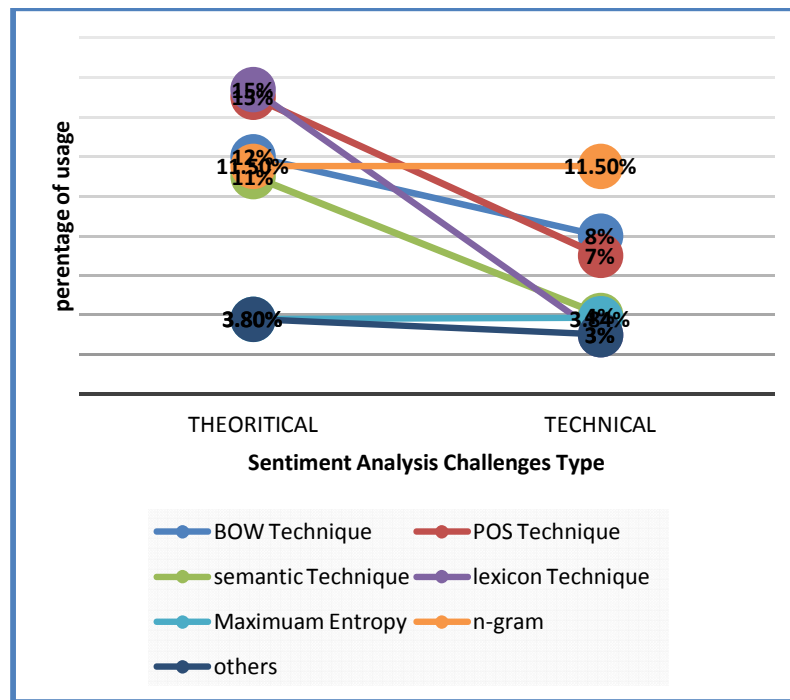


Figure. 1: SA Challenges Type and technique Used

Table.2 examines several parameters relevant to the sentiment analysis challenges. These parameters are lexicon type, domain oriented, lexicon type, dataset, the technique used and the accuracy results. This comparison summarizes the effect of sentiment challenges solutions in analyzing and evaluating sentiment analysis accurately. The lexicon type in comparison in Table.2 refers to the language of the dataset and the size of the dataset. There are several available lexicon as Senti, HowNet, and Wordnet. The used lexicon has the sentiment word and polarity. The polarity differs in the sentiment classification polarity level. This classification of polarity divide into several class levels such as two levels (Positive, and Negative polarities), three levels as in the hierarchical level, or four level (-, Neutral, +, Mixed), and more specified classification into five levels (Very Negative, Negative, Neutral, Positive, Very Positive polarities) (Doaa, 2016). The comparison's strengths are that 1) the facility of understanding the hot area research, 2) illustrating the most effect challenges on the accuracy results, 3) recognizing the propagation of use each sentiment analysis technique, and 4) discuss the relationship among the domain dependence, lexicon type and the accuracy results (Doaa et al., 2015). The results of this comparison is very important in choosing the suitable technique to solve the sentiment challenges to reach the highest accuracy.

Table .2: Study to several parameters effects on the sentiment challenges

Ref.No.	SA Type	Challenge	SA Challenge	Technique Used	Domain oriented	Lexicon type	Data Set	Accuracy
(Yulan et al.,2011)	Theoretical		Domain Dependence	Naïve Bayes and support vector machines from WEKA5	N multi domain	46English	the two datasets, the movie review (MR) data and the multi domain	90%
(Ivan et al., 2013)	Technical		Bi-polar words	combination of features (n-grams) and preprocessing techniques (unsupervised Stemming and phonetic transcription).	Y	English Facebook	Facebook dataset containing 10,000 posts	69%
(Hiroshi & Tetsuya,2006)	Theoretical		Domain Dependence	Deep sentiment analysis method analogous to machine translation	N	Japanese	polar clauses conveying goodness and badness in a specific domain	94% (25 to 33%)
(Maral, 2011)	Theoretical		Negation+ domain	BOW term	Y	two wordlists	2000 movie reviews: 1000	65% with

		dependence	frequencies			positive and 1000 negative	higher recall 83%
[55]	Theoretical	Domain Dependence	SemEval-2013	Y	Tweets and MPQA English	2000 positive words and 4700 negative words, also the popular MPQA	Improve accuracy and F-measure about 13% from base line to reach 69%
(Andrius et al., 2012)	Technical	Huge lexicon	Bag-of-words SVM.	Y, CNET, IMDB movie reviews	pSenti	The first dataset Software Review, second data set Movie Reviews	82.30%
(Saif & Peter, 2010)	Technical	NLP overheads (emotions)	(Naive Bayes, Maximum Entropy, and SVM)	N multi domain	Microblogging lexicon	tweets with emoticons, 1,600,000 training tweets., 800,000 tweets with positive emoticons, and 800,000 tweets with negative emoticons,	accuracy improved for Naive Bayes (81.3% from to 82.7%) and Max-Ent (from 80.5 to 82.7). However, there was a decline for SVM (from 82.2% to 81.6%).
(Alexandra et al., 2013)	Theoretical	Domain Dependence	WordNet-based	Y	News reviews	newspaper articles (the set of 1292 quotes)	82% improve the base line 21%
(Theresa et al., 2005)	Technical	Huge lexicon	Auto.Distinguishing prior and contextual polarity.	N	Multi-perspective Question Answering (MPQA) Opinion Corpus1,	15,991 subjective expressions from 425 documents (8,984 sentences)	75.9%,
(Erik & Marie, 2009)	Technical	Nlp overheads (Multi-lingual)	Integrated approach combining from information retrieval, natural language processing and machine learning	Y	English,, Dutch and French tex	blog, review and forum texts found on the World Wide Web	83% 70% and 68%
(Fangtao et al., 2010)	Theoretical	Domain Dependence	Dependency-Sentiment-LDA-Markov chain	Y	HowNet-Senti-wordnet-MPQA	online customers reviews-HowNet 2700 2009 English translation of positive/negative Chinese SentiWordNet 4800 2290 Words with a positive or negative score- MPQA 4152 2304 MPQA subjectivity lexicon	70.7
(Duyu et al., 2014)	Technical	Nlp overheads (emotions)	fine-grained emotions	Y	Chinese lexicon	35,000 tweets about Sichuan earthquake	80%,
(Bas et al., 2011)	Theoretical	Negation	Part of speech (POS)	40 different topics	OpenNLP	Dutch language	71.23% for negation (Precision improves with 1.17%)

(Ouyang et al., 2014)	Theoretical	Domain Dependence	Emotion Dependency Tuple (EDT-improved (BOW) TF-IDF and cross entropy , space vector model	N	Chinese	COAE2014 dataset	60%
(Lucie et al., 2015)	Technical	Bi-polar words	n-gram (uni and bi-grams)	Y	HL and MPQA lexicon.	data set of 1,600 Facebook messages	70%
(Qingxi & Ming, 2014)	Theoretical	Spam and fake reviews	Combine lexicon and use shallow dependency parser	N, online customers reviews	SentiWordNet and MPQA	store#364,	85.7% for sentiment method but word counting approach 76.7%
(Emitza & Walid, 2014)	Technical	Feature and keywords extraction	POS tagging with fine-grained app	N, 7 applications	SentiStrength	7 apps from the Apple App Store and Google Play Store	91%
(Alexandra & Ralf, 2009)	Theoretical	World knowledge	Adding word polarity scores from sentiment lexicons.	Y	context-dependent lexicon	6500 answers on game reviews	Improve acc 60% to 80%
(Lifeng, 2009)	Theoretical	Negation	parse Tree and dependency	Y	English, health /medical domain	dataset that consists of 1000 sentences	Between 79.2% to 82% with different four methods
(Mohammed et al., 2014)	Technical+ Theoretical	Domain dependence+ NLP overheads (multi-language)	Lexicon-based method depends on POS tagging	N	16 domain Lexicon-based tool for Arabic opinion mining.	Deal with emoticons, chat language, Arabizi,	93.9%
(Chetan & Atul, 2014)	Technical	Huge lexicon	lexicon based technique	Y	6,74,412 tweets	the polarities of the words in the dictionary are set according to a specific domain,	73.5 %
(Jiang & Min, 2011)	Theoretical	Domain-dependence	n-gram	N, 7 domains	Chinese reviews b	560 Chinese review	65%
(Alexander et al., 2011)	Theoretical	Negation	POS Technique (Word Sense Disambiguation, Sentiment analysis)	Y	WORDNET	1;000 positive and 1;000 negative English movie	98:7%-
(Doaa et al., 2015)	Technical + Theoretical	Lexicon+ Feature extraction+ Negation+ world knowledge	Enhancement BOW model	Y, scientific papers	New lexicon	Three datasets (training set, test set and the verified set) 1000, 5000, and 10.000	83.5%
(Walter & Mihaela, 2011)	Technical + Theoretical	Extracting Features keywords +domain dependence or	character n-grams instead of terms	Y	German Hotel reviews	Corpus of 1559 hotel reviews crawled from the web.	83%
(Myle et al., 2011)	Theoretical	Spam/ fake reviews	POS tagging similarities and n-gram algorithm	N, online customer reviews	LIWC	800 opinions	nearly 90%

The comparison's conclusion in Table.2, the relationship between the sentiment analysis challenges type and the importance of its presence in the new search. Other results from the second comparison declare in the figure.2 that the percentage of Average of accuracy enhancement related to the compared research papers. Although the Negation is the most affected in any sentiment type as the results in comparison one in the Table.1 that means it has a big number of research. That makes the result of it is lower here.

$$AVG(ACC.) = \sum_{i=0}^n \frac{\text{Accuracy of each paper}}{\text{number of papers (n)}}, \quad (1).$$

That means the lowest Average of accuracy which is the highest rate research area with the bi-polar words with 69.5. Then domain dependence and NLP overheads have the ranked second. And the Negation challenge has the third rank.

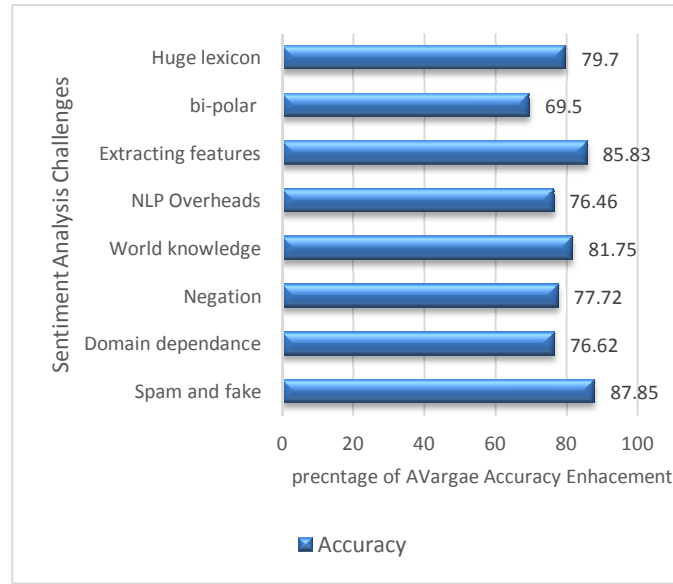


Figure.2: The improvement in accuracy results in sentiment analysis challenges

Figure.3 presents the highest improvement in accuracy for each sentiment analysis challenges related to the second comparison. Negation has the highest accuracy percentage that can support the result of the first comparison because there is no research in sentiment don't need to understand the negative reviews whether explicit or implicit. And the least score in accuracy is bi-polar words research, so we recommend to increase the research in it.

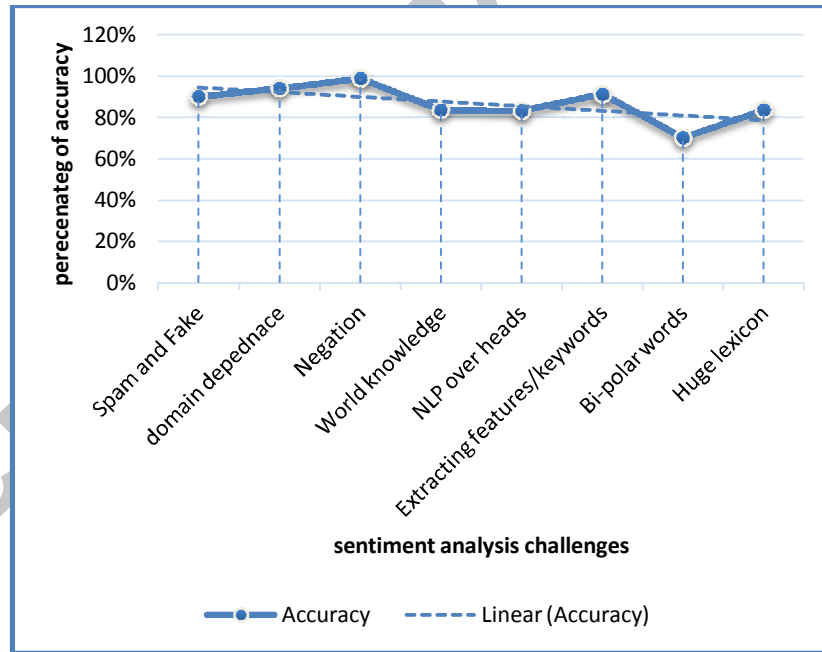


Figure. 3: The highest accuracy to each sentiment analysis challenge

III. CONCLUSION & FUTURE WORK

This survey discusses the importance and effects of sentiment analysis challenges in sentiment evaluation based on two comparisons among forty-seven papers. The first comparison is based on the relationship between the sentiment review structure and the sentiment analysis challenges. The result of this comparison reveal another essential factor to recognize the sentiment challenges which is domain-dependence. Moreover, the negation challenge became popular in all types of review structured just

differs in implicit or explicit meaning. This comparison result provides a facility to the effects of each sentiment challenges on the review structure types. We conclude that the topic nature and the review structure determines the suitable challenges for the evaluation sentiment reviews. Then the second comparison relies on the sentiment analysis challenges relevant to the accuracy rate. The results of them present the importance of sentiment challenges in evaluating the sentiments and how to select the fitting challenge to improve accuracy. We find the relationship between the proportion of sentiment techniques usage in theoretical and technical types to solve sentiment challenges. Another result explains the hot area of research is a theoretical type of sentiment challenges. That reflects on the results of the average of accuracy based on the number of researches in each challenge. The more research in a sentiment challenge, the less Average of accuracy rate. The future work is the expansion of the comparison circle larger with the new research continuously.

REFERENCES

- Ahmed, A., Zhu, Z., David, Z., Hsinchun, C., Jay, F. N., 2010. Detecting Fake Websites: The Contribution of Statistical Learning Theory, *MIS Quarterly* Vol. 34 No. 3, pp. 435-461.
- Aimur, Y., Yisong, Y., Claire, C., 2010. Multi-level Structured Models for Document-level Sentiment Classification, *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1046–1056, MIT, Massachusetts, Association for Computational Linguistics, USA.
- Alexander, H., Paul, v. I., Bas, H., Flavius, F., Uzay, K., 2011. Determining Negation Scope and Strength in Sentiment Analysis, *Proceedings of the 2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2011)*, IEEE Computer Society.
- Alexandra, B., Ralf, S., 2009. Rethinking Sentiment Analysis in the News: from Theory to Practice and back, Troyano, Cruz, Díaz (Eds.): *WOMSA'09*, pp. 1-12.
- Alexandra, B., Ralf, S., Mijail, K., Vanni, Z., Erik, v. d. G., Matina, H., Bruno, P., Jenya, B., 2013. Sentiment Analysis in the News, in *proceeding Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*.
- Amna, A., 2012. Negation Identification and Calculation in Sentiment Analysis, (IMMM 2012) *The Second International Conference on Advances in Information Mining and Management*.
- Andrius, M., Dell, Z., Mark, L., 2012. Combining Lexicon and Learning based Approaches for Concept-Level Sentiment Analysis, *WISDOM'12*, Beijing, China.
- Arjun, M., Bing, L., Junhui, W., Natalie, G., Nitin J., 2011. Detecting Group Review Spam, *WWW-2011 poster paper*.
- Arjun, M., Bing, L., Natalie, G., 2012. Spotting Fake Reviewer Groups in Consumer Reviews, *International World Wide Web Conference (WWW-2012)*, Lyon, France.
- Arjun, M., Vivek, V., Bing, L., Natalie, G., 2013. What Yelp Fake Review Filter Might Be Doing, *Proceedings of The International AAAI Conference on Weblogs and Social Media (ICWSM-2013)*, Boston, USA.
- Bas, H., Paul, v. I., Alexander, H., Flavius, F., Uzay, K., 2011. Accounting for Negation in Sentiment Analysis, *DIR 2011*, Amsterdam, The Netherlands.
- Basant, A., Namita, M., Pooja, B., Sonal Garg2, 2015. Sentiment Analysis Using Common-Sense and Context Information, *Hindawi Publishing Corporation Computational Intelligence and Neuroscience*.
- Bing, X., Liang, Z., 2014. Improving Twitter Sentiment Analysis with Topic-Based Mixture Modeling and Semi-Supervised Training, *proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Short Papers)*, Association for Computational Linguistics, Baltimore, Maryland, USA.
- Chetan, K., Atul, M., 2014. A Scalable Lexicon Based Technquie for Sentiment Analysis", *International Journal in Foundations of Computer Science & Technology (IJFCST)*, Vol.4, No.5.
- Christine, L., Florian, K., Antal, v. d. B., 2013. The perfect solution for detecting sarcasm in tweets #not, *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 29–37, Association for Computational Linguistics, Atlanta, Georgia.
- Doaa, M. E., Hoda, M.O. M., Osama, I., 2015. Online Paper Review Analysis" *International Journal of Advanced Computer Science and Applications (IJACSA)*, Volume 6, Number 9.
- Doaa, M.E., "Enhancement Bag-of-Words Model for Solving the Challenges of Sentiment Analysis" *International Journal of Advanced Computer Science and Applications(IJACSA)*, Volume 7, Issue 1, 2016.
- Duyu, T., Bing, Q., Ting, L., Qihui, S., 2014. Emotion Analysis Platform on Chinese Microblog", *CoRR journal*.
- Emanuele, L., Jonathon, R., Lilja, Ø., 2012. Representing and Resolving Negation for Sentiment Analysis, *Published in Data Mining Workshops (ICDMW)*, 2012 IEEE 12th International Conference on.
- Emitza, G., Walid, M., 2014. How Do Users Like This Feature? A Fine Grained Sentiment Analysis of App Reviews, *22nd IEEE International Requirements Engineering Conference*.
- Erik, B., Marie-Francine, M., 2009. A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts, published in *journal Information Retrieval*, Volume 12 Issue 5.
- Erik, C., 2013. An Introduction to Concept-Level Sentiment Analysis, F. Castro, A. Gelbukh, M.G. Mendoza (Eds.): *MICAI 2013, Part II*, LNAI 8266, pp. 478–483, Springer.
- Fangtao, L., Minlie, H., Xiaoyan, Z., 2010. Sentiment Analysis with Global Topics and Local Dependency, *AAAI Conference on Artificial Intelligence*.
- Fangtao, L., Minlie, H., Yi, Y., Xiaoyan, Z., 2011. Learning to Identify Review Spam, *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*.
- G.Vinodhini, RM.Chandrasekaran, 2012. Sentiment Analysis and Opinion Mining: A Survey, *International Journal of Advanced Research in Computer Science and Software Engineering*, Volume 2, Issue 6.
- Gizem, G., Berrin, Y., Dilek, T. , Y'ücel, S., 2012. New Features for Sentiment Analysis: Do Sentences Matter?, In *SDAD 2012 The 1st International Workshop on Sentiment Discovery from Affective Data*, pages 5.
- Haochen, Z., Fei, S., 2015. Aspect-Level Sentiment Analysis Based on a Generalized Probabilistic Topic and Syntax Model, in *Proceedings of the Twenty-Eighth International Florida Artificial Intelligence Research Society Conference*, Association for the Advancement of Artificial Intelligence.
- Hiroshi, K., Tetsuya, N., 2006. Fully Automatic Lexicon Expansion for Domain-oriented Sentiment Analysis, *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, pages 355–363, Sydney, c2006 Association for Computational Linguistics.
- Ismat. M. E., A.S., Ali , 2011. Comparative study of accuracy in distance measurement using: Optical and digital levels, *Journal of King Saud University - Engineering Sciences*, Volume 23, Issue 1, Pages 15–19.
- Ivan, H., Toma's, P., Josef, S., 2013. Sentiment Analysis in Czech Social Media Using Supervised Machine Learning, *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 65–74, Atlanta, Georgia, Association for Computational Linguistics.
- Jiang, Y., Min, H., 2011. Using Topic Sentiment Sentences to Recognize Sentiment Polarity in Chinese Reviews, *CIPS-SIGHAN Joint Conference on Chinese Language*.
- Jiliang, T., Xufei, W., Huiji, G., Xia, H., Huan, L., 2012. Enriching short text representation in microblog for clustering, *Frontiers of Computer Science*, Higher Education Press and Springer-Verlag Berlin Heidelberg.

- Khairullah, K., Baharum, B., Aumagzeb, K., Ashraf, U., 2014. Mining opinion components from unstructured reviews: A review, *Journal of King Saud University- Computer and Information Sciences*, Volume 26, Issue 3, Pages 258-275.
- Lifeng, J., 2009. The Effect of Negation on Sentiment Analysis and Retrieval Effectiveness, *CIKM'09*, Hong Kong, China.
- Ling, P., Geng, C., Menghou, Z., Chunya, L., 2014. What do seller manipulations of online product reviews mean to consumers? , (HKIBS Working Paper Series 070-1314). Hong Kong: Hong Kong Institute of Business Studies, Lingnan University.
- Lucie, F., Eugen, R., Daniel, P., 2015. Analysing domain suitability of a sentiment lexicon by identifying distributionally bipolar words", *Proceedings of the Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. EMNLP*.
- Maral, D., 2011. Scope of Negation Detection in Sentiment Analysis, *Proceeding In: Dutch-Belgian Information Retrieval Workshop, DIR 2011*, Amsterdam, the Netherlands.
- Marina, B., Claudiu, C. M., Boi, F., 2014. Acquiring Commonsense Knowledge for Sentiment Analysis using Human Computation, *proceeding WWW'14 Companion*, Seoul, Korea.
- Matthew, J. K., Spencer, G., Andrea, Z., (2015). Potential Applications of Sentiment Analysis in Educational Research and Practice – Is SITE the Friendliest Conference?" In D. Slykhuis & G. Marks (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2015*, Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).
- Michael, W., Alexandra, B., Benjamin, R., Dietrich, K., Andres, M., 2010. A Survey on the Role of Negation in Sentiment Analysis, *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, pages 60–68, Uppsala.
- Mohammed, N. A. , Izzat, M. A., Amal, H. G. , Heider, A. W., Mohamad, M. H. ,2014. Opinion Mining and Analysis for Arabic Language, (IJACSA) *International Journal of Advanced Computer Science and Applications*, Vol. 5, No. 5.
- Mus'ab, H., Ahmet, K., Dilek, T., Berrin, Y., Yucel, S., 2012. An Aspect-Lexicon Creation and Evaluation Tool for Sentiment Analysis Researchers?, *Proceedings of the 2012 European Conference on Machine Learning and Knowledge Discovery in Databases - Volume Part II, of ECML PKDD'12*, pages 804--807, Berlin, Heidelberg.
- Mye, O., Yejin, C., Claire, C., Jeffrey, T., 2011. Finding Deceptive Opinion Spam by Any Stretch of the Imagination, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 309–319, Portland, Oregon.
- Nathan, G., Ruihong, H., 2013. Sarcasm as Contrast between a Positive Sentiment and Negative Situation, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*.
- Nikos, E., Angeliki, L., Georgios, P., Konstantinos, C., 2011. ELS: a word-level method for entity-level sentiment analysis, *WIMS '11 Proceedings of the International Conference on Web Intelligence, Mining and Semantics*.
- Noura, F., Elie, C., Rawad, A.A., Hazem, H., 2010. Sentence-level and Document-level Sentiment Mining for Arabic Texts, in *proceeding IEEE International Conference on Data Mining Workshops*.
- Ouyang, C., Zhou, W., Yu, Y., Liu, Z., Yang, X., 2014. Topic Sentiment Analysis in Chinese News, *International Journal of Multimedia & Ubiquitous Engineering*;12/5/2014, Vol. 9 Issue 11, p385.
- Qingxi, P., Ming, Z., 2014. Detecting Spam Review through Sentiment Analysis, *Journal of Software*, Vol.9, No. 8.
- Robert, R., 2013. Modeling and Representing Negation in Data-driven Machine Learning-based Sentiment Analysis, *ESSEM@ AI* IA*.
- Saif, M. M., Peter, D. T., 2010. Emotions Evoked by CommonWords and Phrases: Using Mechanical Turk to Create an Emotion Lexicon, *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 26–34, Los Angeles, California, Association for Computational Linguistics.
- Sathya, D., Kaladevi, P., 2014. Enhancing Unsupervised Approach For Person name Bipolarization, *International Journal Of Engineering And Computer Science* ISSN:2319-7242, Volume 3 Issue 2, Pages 3982-3985.
- Stanislav, B., 2013. An Approach to Feature Extraction for Sentiment Analysis of News Texts.
- Subhabrata, M., Pushpak, B., 2012. Feature Specific Sentiment Analysis for Product Reviews, *CICLing 2012*, part I, LNCS 78181, Springer-Verlag Berlin Heidelberg.
- Sujata, R., Parteek, K., 2014. Challenges of Sentiment Analysis and Existing State of Art, *International Journal of Innovation and Research in Computer Science (IJIRCS)*.
- Svetlana, K., Xiaodan, Z., Saif, M. M., 2014. Sentiment Analysis of Short Informal Texts, *Journal of Artificial Intelligence Research* 50.
- Tawunrat ,C., Jeremy, E., 2015. Chapter Information Science and Applications, Simple Approaches of Sentiment Analysis via Ensemble Learning, Volume 339 of the series *Lecture Notes in Electrical Engineering , DISCIPLINES Computer Science, EngineeringSUBDISCIPLINESAI , Information Systems and Applications •Computational Intelligence and Complexity*.
- Theodoros, L., 2012. Fake Reviews: The Malicious Perspective, *Natural Language Processing and Information Systems - 17th International Conference on Applications of Natural Language to Information Systems, {NLDB} 2012*, Groningen, the Netherlands.
- Theresa, W., Janyce, W., Paul, H., 2005. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis, *proceeding HLT '05 Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*.
- Thomas, B., 2013. What Consumers Think about brands on social media, and what bunesses need to do about it, *Report, Keep Social Honest*.
- Walter, K., Mihaela, V., 2011. Sentiment Analysis for Hotel Reviews", *Proceedings of the Computational Linguistics-Applications Conference, JACHRANKA*.
- Xia, H., Jiliang, T., Huiji, G., Huan, L., 2014. Social Spammer Detection with Sentiment Information, *proceeding ICDM '14 Proceedings of the 2014 IEEE International Conference on Data Mining*.
- Yanfang, C., Pu, Z., Anping, X., 2015. Sentiment Analysis Based on Expanded Aspect and Polarity-Ambiguous Word Lexicon , (IJACSA) *International Journal of Advanced Computer Science and Applications*, Vol. 6, No. 2.
- Yulan, H., Chenghua, L., Harith, A., 2011. Automatically Extracting Polarity-Bearing Topics for Cross-Domain Sentiment Classification, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, Portland, Oregon, 2011 Association for Computational Linguistics.
- Yunfang, W., Miaomiao, W., 2010. Disambiguating Dynamic Sentiment Ambiguous Adjectives, *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pages 1191–1199, Beijing.