Effectiveness of Social Media Sentiment Analysis Tools with the Support of Emoticon/Emoji

68

Duncan C. Peacock and Habib Ullah Khan

68.1 Introduction

Opinions are a central driver of human behaviour. Individuals naturally seek the opinions of others before making decisions, such as buying products and services, investing, and voting in elections. This consultation is being increasingly done using microblogging platforms such as Twitter, posts on social media, discussion forums or reviews on sites like TripAdvisor [1, 2]. Organisations also need feedback on their products and services so that resources can be allocated efficiently to find new investment opportunities, to publicise and improve products, and to anticipate problems. Consequently, interest has grown in a field of study called sentiment analysis to extract meaning from the vast amounts of digital opinion data available. One key feature of a post (or group of posts) that is frequently desired is whether its sentiment polarity is positive, neutral, or negative about a subject. This can be used to give a single sentiment signal, or be aggregated to give an opinion over time [3]. It is vital, therefore, that the increasing number of sentiment analysis tools developed for this purpose classify posts as accurately as possible.

The main approaches used in sentiment analysis are lexicon-based, data- or corpus-based, or a combination of the two. Depending on the algorithm used and the training data, there can potentially be wide differences in the results. For example, unsupervised (lexicon-based) methods can perform better across different subject domains, whereas supervised methods (trained, e.g., on product data), may be better in specialist areas. Analysis of posts made by the wider public must deal with slang, sarcasm, abbreviations, misspellings, grammatical aspects (e.g., multiple exclamation marks), demographics, and technology changes. For example, emojis and emoticons, which are increasingly used on smartphones, can be used to clarify, enhance, or sometimes reverse the sentiment of a post.

Sentiment analysis tools are offered as stand-alone products, but increasingly through APIs as web services. This could potentially offer organisations the chance to compare products, select specialist tools depending on requirements, and benefit from online lexicons and ongoing algorithm development.

Sentiment analysis of short social media messages on microblogging platforms such as Twitter or Instagram is of high interest to organisations that increasingly want to use social media to study the public mood in addition to or in place of traditional methods of obtaining feedback, such as surveys and opinion polls. An increasing number of specialist tools, that can rate the sentiment of a post in a microblog, are being offered to organisations as web services to cater for this need.

Analysis of microblogging messages must be able to handle short messages, varied language use, and specifics such as emoticons, emojis, and hashtags. Emoticons and emojis are increasingly being used in short social media messages and appear to have a significant effect on the sentiment of a tweet and the accuracy of classification. For example, one study [4] suggested that using only the emoticon to rate sentiments could achieve accuracy rates of above 80% [5]. further suggested that emoticon sentiment is likely to be more important than text sentiment and may increase accuracy across subject domains. However, [6, 7], in a limited test, cautiously suggested that there may be classification errors with some sentiment analysis tools in the case in which the emoticon sentiment disagrees with the text sentiment.

D. C. Peacock
University of Liverpool, Liverpool, UK

H. U. Khan (✉)
College of Business and Economics, Qatar University, Doha, Qatar
e-mail: habib.khan@qu.edu.qa

S. Latifi (ed.), *16th International Conference on Information Technology-New Generations (ITNG 2019)*,
Advances in Intelligent Systems and Computing 800,
https://doi.org/10.1007/978-3-030-14070-0_68

Details of the approach used in developing commercial web services for sentiment analysis are not always available; and therefore comparing them is difficult. It seems that the effect of emoticons and emojis should be considered.

68.1.1 Main Aims of the Project

1. To develop a prototype application that can be used to compare web-service-based sentiment analysis tools and prove or disprove the hypotheses.
2. Through this artefact, study inconsistencies in the treatment of emoticons, emojis, and subject area by different sentiment analysis tools.
3. Evaluate whether one tool or method of analysis is more accurate than another by comparing tools against a manually labelled data set.
4. Provide an application that could be expanded in the future into a platform for testing, comparing, and benchmarking sentiment analysis tools and generate test sets for wider study.

68.1.2 Subsidiary Aims

5. Demonstrate a technique for organisations to find the best sentiment analysis tools for their needs.
6. Demonstrate a tool for organisations offering sentiment analysis web services the ability to benchmark their tool against others on the market.

68.2 Literature Review

Sentiment analysis or opinion mining – a subtopic of natural language processing (NLP) – is the study of public opinion, emotions, and attitudes through the analysis of written language [1, 8]. It is a popular area of research with 7000 research articles already written by [9]. Interest from business and other organisations has grown as the amount of digital data, and the use of social media and smart devices has increased [10–12]. There are currently around 319 million monthly active Twitter users, compared with 1.817 billion users of Facebook, 1 billion users of WhatsApp, 600 million users of Instagram, and 877 million users of QQ – a Chinese microblogging platform [13].

The ability of microblogs to give instant feedback is valuable in many domains. Applications have been written that give organisations the ability to, for example, assess whether a target audience is happy (a positive sentiment polarity) or unhappy (a negative sentiment polarity) using a sample of tweets from the live Twitter stream on a desired subject. This could be used to track the sentiment of a brand or feeling about a product. Other uses include politics (judging reactions to policies or predicting election results), financial markets (tracking sentiment on stocks) and tracing the spread of a disease [3, 14].

Given the amount source material, a solid foundation in the subject was provided by sources such as a widely-cited book [1] and survey papers [9]. These followed the evolution of the subject since it was identified around the year 2000, including the strengths and weaknesses of the different techniques – supervised, unsupervised or hybrid – used in sentiment analysis research. A series of conferences called SemEval [15] tracks ongoing developments of computation techniques in semantic evaluation and has a competition to improve techniques in sentiment analysis applications, such as the ‘support vector machine classifier and hashtag’ used successfully by [16]. Twitter is of particular interest because of the availability of data and the ease-of-use of the public API. At the recent 2016 SemEval conference, Twitter research was the most popular [15], but techniques refined on Twitter could be applied to other platforms. For these reasons, Twitter was chosen as the data source in this project.

The first recorded emotional icon or emoticon in digital communication was a smiley ‘:-)’ used in 1982 at Carnegie Mellon University to indicate that a piece of text was a joke [17]. They began (and are still commonly used) as text characters indicating facial expressions. More recently, icons such as have been increasingly used in place of the text – either through substitution or by allowing the user to select one from a list of icons. Emojis (meaning picture character in Japanese) are a step further, allowing short messages to be sent with pictograms showing concepts such as celebrations, weather, vehicles, thumbs up, and so on [15].

There is some confusion about the terms emoticon, emoji, and smiley, and they are often used interchangeably. However, there are differences in the history, usage, and technical implementations of the text and pictorial variants. Hence, following [15], they are defined in this paper as follows:

- *Emoticons* are pictures made up from the standard ASCII character set used to indicate a facial expression. For example, the smiley emoticons ‘:)’ (read sideways) and ‘ (^_^) ’, and the sad emoticon ‘:(’.
- *Emojis* are pictorial evolutions of emoticons that allow a wider range of ideas (such as weather, directions, and vehicles) as well as emotions. They are stored as Unicode characters with the first set introduced in 1995 [18].

Since the introduction of smartphones (and their addition to popular apps), emojis have become increasingly popular [15] found 4% of tweets and up to 50% of Instagram messages contained emoticons or emojis, whereas [19], found a 20% occurrence in a database of Japanese tweets [20]. noted that the Sina microblog contains a larger number of emoticons and emojis than Twitter. Academic studies such as those by [4, 5, 15] have investigated how emoticons and emojis could be used to improve the accuracy of sentiment analysis classification tools. One challenging area is where emoji or emoticon sentiment is different from the text, perhaps indicating sarcasm.

Several downloadable sentiment analysis tools, such as SentiStrength (2016), have been developed and methods have been created to compare their performance [15] and set benchmarks [21]. However, sentiment analysis is also increasingly being offered in the form of commercial web service APIs. many of which are aimed at Twitter. Factors such as increasing emoticons/emoji use, online dictionaries, machine classification, and ongoing development may make the performance of such tools differ from generalised tools, and vary in relation to each other over time [3, 22, 23]. included some web services in their benchmarking tests; however, to the author's knowledge, no comparison framework specifically aimed at Twitter-based web services exists. There is a need for a specialist service so that organisations with specific social sentiment analysis requirements can find the best tool for their needs.

68.2.1 Sentiment Analysis Research

Liu [1] described how sentiment analysis research chiefly consists of breaking a piece of text down into its constituent parts at three levels:

- Document level – A document is assumed to refer to one entity or subject (like a product), and a positive or negative sentiment is calculated for the whole document.
- Sentence (or even clause) level – where sentences within a document are classified first as subjective or objective and then as positive, negative, or neutral.
- Entity/aspect level – attempts to analyse exactly which aspects of the entity (price, size, etc.) are being rated.

As sentiment analysis concerns natural language, many difficult research problems have had to be addressed, including:

- Comparative opinions, for example, 'iPhones are better than Samsungs'.
- Sentences that mean different things in different subject domains, for instance, 'this vacuum cleaner sucks' [1].
- Sarcasm (particularly in the political sphere).
- Sentences with sentiment words that express no feelings (i.e., factual) and sentences with no sentiment words that express an opinion, for example, 'Can you recommend a good restaurant?'.

Both supervised and unsupervised approaches are used for sentiment analysis, and there are challenges at each level. For example, at the level of extracting aspects from a body of text, [1] noted how a supervised model trained on a test set from one domain (e.g., product reviews) might not perform as well in a different domain. Unsupervised (lexicon or dictionary based) methods can perform better across different domains. This was supported in research by, among [3, 5].

68.2.2 Sentiment Analysis on Microblogging Platforms

Twitter-like microblog posts differ from sources traditionally used in sentiment analysis in several ways:

- Tweets are limited to 140 characters, meaning that they are usually short and to the point. Other platforms may not be as limited, but there is more of a focus on short messages.
- Emoticons and emojis are used both to enhance the sentiment of a tweet [5] and to indicate a joke or sarcasm [24].
- Language use is more casual, less composed, uses slang, and can vary by subject [25]. noted that other emotional signals, such as certain word pairings, exist. However, [26] found a low correlation between the emotional words used in social media and the emotional state of the user, suggesting that using words alone is not sufficient to identify sentiment.
- Volume, speed, variation, and noisiness of data.
- The use of hashtags, both for subject identification and for sentiment annotation [16].
- A group view rather than Individual views on a topic is the target of research [15]
- Other features such retweets, follows, and mentions [27].

Giachanou and Crestani [27] added that there are also specific issues in processing microblogging messages in areas like topic identification, tokenization, and data sparsity (incorrect language and misspellings). This has led to two further approaches to sentiment analysis: hybrid models that combine lexical and machine-learning methods, and graph-based models that include social networking features.

Note: Research is still in progress.

References

1. Liu, B.: Sentiment Analysis and Opinion Mining. Morgan & Claypool (Synthesis lectures on human language technologies, 16), San Rafael (2012)
2. Khan, H.U., Gadhoum, Y.: Measuring internet addiction in Arab based knowledge societies: a case study of Saudi Arabia. J. Theor. Appl. Inf. Technol. **96**, (2018)
3. Abbasi, A., Hassan, A., Dhar, M.: Benchmarking twitter sentiment analysis tools. In: ResearchGate. 9th Language Resources and Evaluation Conference. Available at: https://www.researchgate.net/publication/273000042_Benchmarking_Twitter_Sentiment_Analysis_Tools (2014). Accessed 15 Dec 2018
4. Go, A., Bhayani, R., Huang, L.: Twitter sentiment classification using distant supervision. CS224N Proj. Rep. Stanford. **1**, 12 (2009)
5. Hogenboom, A., Bal, D., Frasincar, F., Bal, M., de Jong, F., Kaymak, U.: Exploiting emoticons in sentiment analysis. In: Proceedings of the 28th Annual ACM Symposium on Applied Computing, pp. 703–710. ACM (2013)
6. Teh, P.L., Rayson, P.E., Pak, I., Piao, S.S., Yeng, S.M.: Reversing the polarity with emoticons. In 21st International Conference on Applications of Natural Language to Information Systems. Available at: http://eprints.lancs.ac.uk/79422/1/nldb_2016.doc (2016). Accessed 20 Dec 2018
7. Khan, H.U., Awan, M.A.: Possible factors affecting internet addiction: a case study of higher education students of Qatar. Int. J. Bus. Inf. Syst. **26**(2), 261–276 (2017)
8. Brock, V.F., Khan, H.U.: Big data analytics: does organizational factor matters impact technology acceptance? J. Big Data. **4**(1), 1–28 (2017)
9. Feldman, R.: Techniques and applications for sentiment analysis. Commun. ACM. **56**(4), 82–89 (2013). https://doi.org/10.1145/2436256.2436274
10. Urabe, Y., Rzepka, R., Araki, K.: Comparison of emoticon recommendation methods to improve computer-mediated communication. In: Ulusoy, Ö., Tansel, A.U., Arkun, E. (eds.) Recommendation and Search in Social Networks, pp. 23–39. Springer International Publishing (Lecture Notes in Social Networks) (2015). https://doi.org/10.1007/978-3-319-14379-8_2
11. Heang, J.F., Khan, H.U.: The role of internet marketing in the development of agricultural industry: a case study of China. J. Internet Commer. **14**(1), 1–49 (2015)
12. Bashir, G.M., Khan, H.U.: Factors affecting learning capacity of information technology concepts in a classroom environment of adult learner. 15th International Conference on Information Technology Based Higher Education and Training (IEEE Conference), Istanbul, Turkey, September 8th – September 10, 2016. (Conference Proceeding) (2016)
13. Statista: Twitter MAU worldwide 2016/Statistic, Statista. Available at: https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ (2017). Accessed 10 Jan 2019
14. Najmi, E., Hashmi, K., Malik, Z., Rezgui, A., Khan, H.U.: ConceptOnto: an upper ontology based on Conceptnet. 11th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA' 2014), November 10–13, 2014, Doha, Qatar, pp. 366–372. (Conference Proceeding) (2014)
15. Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F., Stoyanov, V.: SemEval-2016 task 4: sentiment analysis in Twitter. In: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), San Diego, USA. Available at: http://alt.qcri.org/semeval2016/task4/data/uploads/semeval2016_task4_report.pdf (2016). Accessed 20 Dec 2018
16. Mohammad, S.M., Kiritchenko, S., Zhu, X.: NRC-Canada: building the state-of-the-art in sentiment analysis of tweets. arXiv preprint arXiv:1308.6242. Available at: http://arxiv.org/abs/1308.6242 (2013). Accessed 27 Dec 2018
17. Dresner, E., Herring, S.C.: Functions of the nonverbal in CMC: emoticons and illocutionary force. Commun. Theory. **20**(3), 249–268 (2010). https://doi.org/10.1111/j.1468-2885.2010.01362.x
18. Unicode.org: Emoji Versions, v3.0. Available at: http://www.unicode.org/emoji/charts/emoji-versions.html (2016). Accessed 4 Jan 2019
19. Yamamoto, Y., Kumamoto, T., Nadamoto, A.: Role of emoticons for multidimensional sentiment analysis of Twitter. In: Proceedings of the 16th International Conference on Information Integration and Web-based Applications & Services. ACM, pp. 107–115. Available at: http://dl.acm.org/citation.cfm?id=2684283 (2014). Accessed 27 Dec 2018
20. Zhang, L., Pei, S., Deng, L., Han, Y., Zhao, J., Hong, F.: Microblog sentiment analysis based on emoticon networks model. In: Proceedings of the Fifth International Conference on Internet Multimedia Computing and Service. ACM, pp. 134–138. Available at: http://dl.acm.org/citation.cfm?id=2499832 (2013). Accessed 27 Dec 2018
21. Ribeiro, F.N., Araújo, M., Gonçalves, P., André Gonçalves, M., Benevenuto, F.: SentiBench – a benchmark comparison of state-of-the-practice sentiment analysis methods. EPJ Data Sci. **5**(1), (2016). https://doi.org/10.1140/epjds/s13688-016-0085-1
22. Bankole, O.A., Lalitha, M., Khan, H.U., Jinugu, A.: Information technology in the maritime industry past, present and future: focus on Lng carriers. 7th IEEE International Advance Computing Conference, Hyderabad, India, 5–7 Jan 2017. (Conference Proceeding) (2017)
23. Smuts, R.G., Lalitha, M., Khan, H.U.: Change management guidelines that adress barriers to technology adoption in an HEI context. 7th IEEE International Advance Computing Conference, Hyderabad, India, 5–7 Jan 2017. (Conference Proceeding) (2017)
24. Rajadesingan, A., Zafarani, R., Liu, H. Sarcasm detection on Twitter: a behavioral modeling approach. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, pp. 97–106. ACM (WSDM '15), New York, NY (2015). https://doi.org/10.1145/2684822.2685316
25. Hu, X., Tang, J., Gao, H., Liu, H.: Unsupervised sentiment analysis with emotional signals. In: Proceedings of the 22nd International Conference on World Wide Web, pp. 607–618. ACM (WWW '13), New York, NY (2013). https://doi.org/10.1145/2488388.2488442
26. Beasley, A., Mason, W.: Emotional states vs. emotional words in social media. In: Proceedings of the ACM Web Science Conference, pp. 31:1–31:10. ACM (WebSci '15), New York, NY (2015). https://doi.org/10.1145/2786451.2786473
27. Giachanou, A., Crestani, F.: Like it or not: a survey of twitter sentiment analysis methods. ACM Comput. Surv. **49**(2), 28:1–28:41 (2016). https://doi.org/10.1145/2938640
28. Khan, H.U., Uwemi, S.: Possible impact of E-commerce strategies on the utilization of E-commerce in Nigeria. Int. J. Bus. Innov. Res. **15**(2), 231–246 (2018)
29. Khan, H.U., Ejike, A.C.: An assessment of the impact of mobile banking on traditional banking in Nigeria. Int. J. Bus. Excell. **11**(4), 446:463 (2017)
30. Khan, H.U., Alhusseini, A.: Optimized web design in the Saudi culture. IEEE Science and Information Conference 2015, pp. 906–915. London, UK, 28–30 July 2015 (2015)