Literature Review

By

Kikelomo O. Obayemi

25 April, 2022

Sentiment Analysis, Choosing the Best Machine Learning Tool.

1.0 Introduction

User opinions are generated daily across multiple data sources such as social media platforms, review sites, blogging sites and other online media. These opinions, have an influence on readers, how they rate entities like individuals, product, brands or political events (Ain et al., 2017). However, they are largely unstructured heaps of data consisting of text, images, videos and animations (Ravi and Ravi, 2015). The value in user opinions can be maximised through proper structuring and analysis in a process known as Sentiment Analysis (Ain et al., 2017). According to Mejova (2009), the field of sentiment analysis has been attracting attention in the past decade both in industry and academia. This is owing to the fact the opinions of internet users can be found on practically any subject and are becoming increasingly important. Sentiment analysis has been applied in organisations to estimate the acceptance level of a product. It can used by policy makers to analyse public sentiments on policies and political challenges (Prabowo and Thelwall, 2009).

Hussein (2018) simply defines Sentiment analysis as “the practice of applying natural language processing and text analysis techniques to identify and extract subjective information from text”. It involves the use of advanced techniques to classify opinions about an entity into positive, negative, or neutral thereby helping users to reach an

informed decision (Mitra, 2020). A key aspect of sentiment analysis is sentiment classification and this is usually done at three levels; the document level, the sentence level and the aspect level or feature level (Nandwani & Verma, 2021). At the document level, the entire document is classified as a whole unit. At the sentence level, the opinion in each sentence is classified and at the aspect level, sentiments are classified with respect to a specific entity (Medhat et al., 2014).

According to Ahmad et al. (2017), the techniques used for conducting sentiment analysis can be classified into three broad categories; the lexicon-based technique, the machine learning technique and the hybrid technique which is a combination of lexicon-based and machine learning approaches.

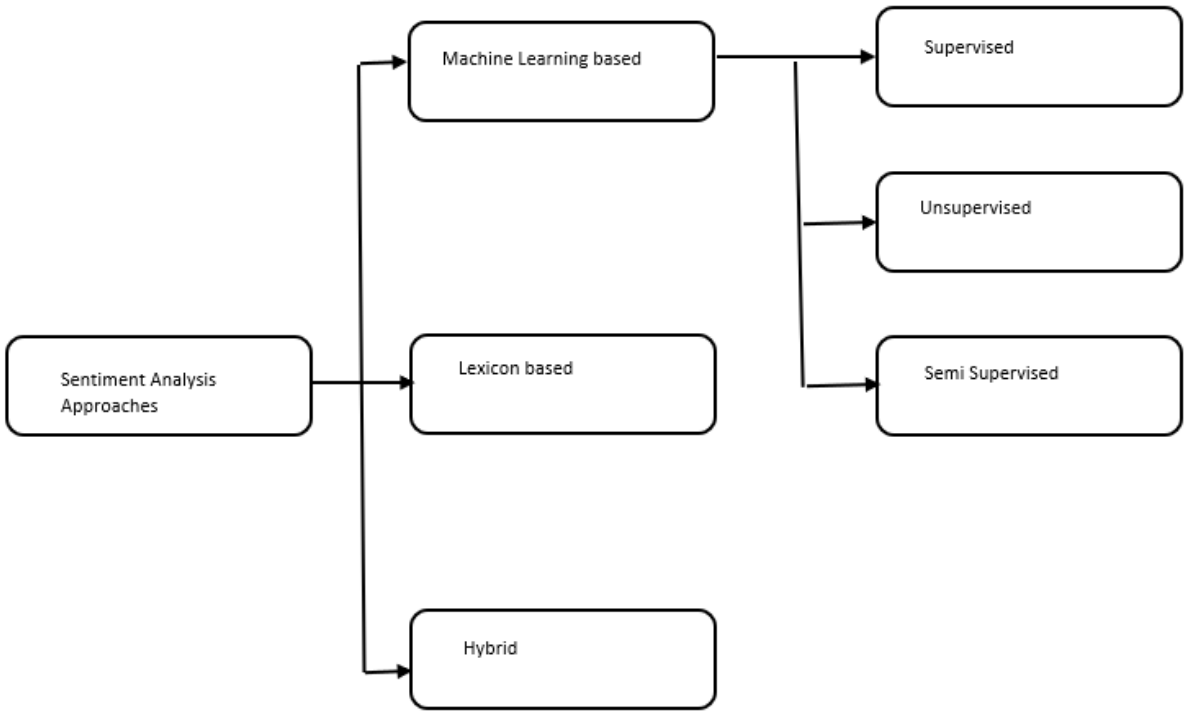

Figure 1: Sentiment analysis approaches and techniques (Madhoushi et al., 2015).

The lexicon-based approach has two methods; the dictionary-based method and the corpus-based method (Madhoushi et al., 2015). The dictionary-based method involves the use of a dictionary which contains words that have already been classified (as positive, negative or neutral) and assigned a polarity value. For example, to analyse a word on a blogging site, we search for it in the dictionary and assign its score to the total polarity score of the site (Annett and Kondrak, 2008). The corpus-based method involves the use of statistical or semantic methods (Madhoushi et al., 2015).

However, in the machine learning approach, classification is done using two sets of data; training data and test data and these classifications are used to predict the polarity of sentiments (Ahmad et al., 2017). Ahmad et al. (2017) further explains that there are three methods under the machine learning approach; the supervised learning which is the most commonly used, the unsupervised learning and the semi-supervised learning. The supervised learning involves the use of labeled data for classification, unsupervised learning uses unlabeled data while semi-supervised combines both labeled and unlabeled data (Tripathy et al, 2017). The supervised learning approaches have been adopted in many researches and have proven quite successful (Madhoushi et al., 2015). Notable supervised learning algorithms are Naïve Bayes, Support Vector Machine (SVM), Maximum Entropy and K-Nearest Neighbors (Devika et al., 2016). Hierarchical clustering, Bayesian Networks and Artificial Neural Networks are all examples of unsupervised learning techniques (Usama et al., 2019). Graph-based methods, multi-view learning and generative models are examples of semi- supervised learning (Madhoushi et al., 2015).

The machine learning approach is more practical for sentiment analysis as it has the ability to handle larger amounts of data (Madhoushi et al., 2015). Mahmood et al. (2020)

notes that the machine learning approach performs better than the lexicon-based approach. This is evident from the many research works that have been conducted on machine learning techniques however most of these works have focused generally on identifying the most effective supervised machine learning algorithms. This literature review critically explores accuracy of machine learning algorithms at all three classification levels. The rest of this paper is structured as follows; section 2 provides more information on the focus of this research and critically reviews other related work. Section 3 provides some concluding thoughts and outlines opportunities for further research.

2.0 Background and Related Work.

Jagtap and Pawar (2013) reports that there are tasks performed at all levels of sentiment classification. The document level classification has one main task of determining the overall sentiment orientation of a document (for example, a blog). There are two tasks at the sentence level; the first task is to determine whether the sentence is subjective (based on sentiments) or objective (factual) and the second one is to classify the sentiments as positive, neutral or negative. At the aspect level, there are three tasks; the first task is to identify features of the entity being commented on, the second task is to determine the polarity of sentiments on these features and the third task involves identifying synonyms of features. Evidently, the aspect level is the most detailed and is useful in extracting information about specific aspects of an entity (Alqaryouti et al., 2020). Imagine a situation where one is satisfied overall with a car but not happy about some aspects of its interior or a situation where a user likes the camera quality of a phone but thinks the battery life

is poor. Aspect level classification is referred to fine-grained analysis and will help to determine exactly what people like or don't like (Bongiwar, 2015).

Medhat et al. (2014) maintains that most applications require the amount of detail that the aspect-level classification offers but for applications like the news sites and blogs, it is important to explore effective techniques at the document and sentence level. Machine learning practitioners and future researchers could use the results of this research to select the most accurate tool at a given classification level. This study will also explore a combination of machine learning approaches to achieve a higher accuracy value.

Wilson et al. (2005) acknowledges that a lot of research work has been done at document level and focuses on phrase level (same as sentence level). This study recognizes contextual polarity in sentiments and classifies them accordingly. Various kinds of negations were considered and a large data subset was used. However, the algorithm used was developed by the researchers and not widely tested. Jagdale et al. (2019) in their research focuses on sentence level classification using only datasets from Amazon and also limits their study to comparisons between two algorithms; Naïve Bayes and SVM concluding that the latter performs better. Similarly, Basani et al. (2019) limits their study to Naïve Bayes and SVM but was able to show that even though the SVM has better accuracy at aspect level, the Naïve Bayes has a faster execution time. Though Khairnar and Kinikar (2013) introduced a third algorithm known as maximum entropy in their study and agrees that SVM does significantly better than the other two, a small dataset was used.

Pang et al. (2002) on the other hand conducted a feature level analysis study, (which is another name for aspect level) and found out that Naïve Bayes is a more accurate solution

when the feature space is small. Jagtap and Pawar (2013) conducted a similar study to the last two researchers, although with a focus on sentence level classification. In this study, supervised machine learning algorithms were used and findings revealed that the SVM algorithm has the best performance while the Naïve Bayes has the worst. The concern with Jagtap and Power's study is that their approach was based on a dataset of positive and negative opinions, neutral opinions were not included. Bhatia et al. (2015) extends the studies by considering social science applications that may produce data with another class of emotion such as "politeness". Whilst this study is commendable due to its wider data set, the method proposed, RST discourse parsing, offers more improvements for lexicon-based analysis and not machine learning techniques.

3.0 Conclusion and Future Work

Despite the amount of research that has been conducted on machine learning-based sentiment analysis, the findings of this review show that;

- Most researchers focus on assessing the performance of an algorithm at one aspect level.
- Many researchers also limit their studies on accuracy to two or three algorithms. A good number of research work selected an algorithm without considering that other algorithms might perform certain sub-tasks better. As clearly pointed out by Ravi and Ravi (2015), sentiment analysis is a multidimensional problem involving various sub-tasks.
- The quality and range of datasets used for analysis is also a problem. Poorly labelled data is cheaper and easier to gather but could result in accuracy errors.

Therefore, there exists an opportunity to extend research to include discussions on algorithm performance at all three levels of sentiment classification. A robust analysis which will utilize high quality web data from multiple sources and compare various supervised machines learning techniques whilst outlining their strengths and shortcomings will be done. It is expected that machine learning practitioners will find the results of this work useful and more practical in selecting the right tools for any level of sentiment analysis.

Future research could explore a fourth classification level known as concept level. Concept level analysis goes beyond mere word-level analysis and uses a novel approach that allows a more efficient passage from unstructured information to structured information in any domain (Cambria et al, 2013). Broader sentiment classification to include “sarcasm” and “irony” detection proves to be an avenue for future research. It is also worth exploring a combination of machine learning and lexicon-based approaches to get a better accuracy on all tasks that make up sentiment analysis.

4.0 References

Ahmad, M., Aftab, S., Muhammad, S.S. and Ahmad, S. (2017). Machine learning techniques for sentiment analysis: A review. *Int. J. Multidiscip. Sci. Eng*, *8*(3):27.

Ain, Q.T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B. and Rehman, A. (2017). Sentiment analysis using deep learning techniques: a review. *Int J Adv Comput Sci Appl*, *8*(6):424.

Alqaryouti, O., Siyam, N., Monem, A.A. and Shaalan, K. (2020). Aspect-based sentiment analysis using smart government review data. Applied Computing and Informatics. DOI: https://doi.org/10.1016/j.aci.2019.11.003

Annett, M. and Kondrak, G. (2008). A comparison of sentiment analysis techniques: Polarizing movie blogs. In Conference of the Canadian Society for Computational Studies of Intelligence. Springer, Berlin, May. 25-35.

Basani, Y., Sibuea, H.V., Sianipar, S.I.P. and Samosir, J.P. (2019) March. Application of sentiment analysis on product review e-commerce. In Journal of Physics: Conference Series 1175(1):012103. IOP Publishing. DOI: https://doi.org/10.1088/1742-6596/1175/1/012103

Bhatia, P., Ji, Y. and Eisenstein, J., (2015). Better document-level sentiment analysis from RST discourse parsing. arXiv preprint arXiv:1509.01599.

Bongirwar, V.K. (2015). A survey on sentence level sentiment analysis. International Journal of Computer Science Trends and Technology (IJCST), 3(3):110-113.

Cambria, E., Schuller, B., Liu, B., Wang, H. and Havasi, C. (2013). Knowledge-based approaches to concept-level sentiment analysis. IEEE intelligent systems, 28(2):12-14.

Devika, M.D., Sunitha, C. and Ganesh, A. (2016). Sentiment analysis: a comparative study on different approaches. *Procedia Computer Science*, *87*:44-49.

Hussein, D.M.E.D.M., 2018. A survey on sentiment analysis challenges. Journal of King Saud University-Engineering Sciences, 30(4):330-338.

Jagdale, R.S., Shirsat, V.S. and Deshmukh, S.N. (2019). Sentiment analysis on product reviews using machine learning techniques. In Cognitive Informatics and Soft Computing. Springer, Singapore. 639-647.

Jagtap, V.S. and Pawar, K. (2013). Analysis of different approaches to sentence-level sentiment classification. International Journal of Scientific Engineering and Technology, 2(3):164-170.

Khairnar, J. and Kinikar, M. (2013). Machine learning algorithms for opinion mining and sentiment classification. International Journal of Scientific and Research Publications, 3(6):1-6.

Madhoushi, Z., Hamdan, A.R. and Zainudin, S. (2015) July. Sentiment analysis techniques in recent works. In 2015 science and information conference (SAI). 288-291. IEEE.

Mahmood A. T, Kamaruddin S. S, Naser R. K. and Nadzir. M. M (2020) A Combination of Lexicon and Machine Learning Approaches for Sentiment Analysis on Facebook. Journal of System and Management Sciences Vol. 10 (3):140-150. DOI: https://doi.org/10.33168/JSMS.2020.0310

Medhat, W., Hassan, A. and Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. Ain Shams engineering journal, 5(4):1093-1113. DOI: https://doi.org/10.1016/j.asej.2014.04.011

Mejova, Y. (2009). "Sentiment analysis: An overview." University of Iowa, Computer Science Department.

Mitra, A., 2020. Sentiment analysis using machine learning approaches (Lexicon based on movie review dataset). Journal of Ubiquitous Computing and Communication Technologies (UCCT), 2(03):45-152.

Nandwani, P. and Verma, R. (2021) A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining, 11(1):1-19.

Pang, B., Lee, L. and Vaithyanathan, S. (2002) Thumbs up? Sentiment classification using machine learning techniques. arXiv preprint cs/0205070.

Prabowo, R. and Thelwall, M. (2009) Sentiment analysis: A combined approach. Journal of Informetrics, 3(2):143-157.

Ravi, K. and Ravi, V. (2015). A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-based systems, 89:14-46

Tripathy, A., Anand A., and Rath S. K. (2017) "Document-level sentiment classification using hybrid machine learning approach." Knowledge and Information Systems 53, 3: 805-831. DOI: https://doi.org/10.1007/s10115-017-1055-z

Usama, M., Qadir, J., Raza, A., Arif, H., Yau, K.L.A., Elkhatib, Y., Hussain, A. and Al-Fuqaha, A. (2019). Unsupervised machine learning for networking: Techniques, applications and research challenges. IEEE access, 7: 65579-65615. DOI: https://doi.org/10.1109/ACCESS.2019.2916648

Wilson, T., Wiebe, J. and Hoffmann, P. (2005), October. Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of human language technology

conference and conference on empirical methods in natural language processing. 347-354.