Literature Review Draft

By

Kikelomo Obayemi

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, datasets, review sites, blogging sites and others. These opinions, have an influence on other readers, how they rate entities such as individuals, product, brands or political events (Ain et al., 2017). User opinions across various sources are largely unstructured heaps of data however, when properly structured and analysed, they can be used to predict significant events such as elections in a process known as Sentiment Analysis (Ain et al., 2017). Opinions about an entity can be classified as positive, negative, or neutral with the aid of sentiment analysis algorithms helping users to reach an informed decision about that entity (Mitra, 2020).

According to Ahmad et al. (2017), the techniques used for conducting sentiment analysis can be classified into three broad categories; lexicon-based technique, machine learning technique and hybrid technique which is a combination of lexicon-based and machine learning approach. Machine learning tools are commonly used for opinion mining and sentiment analysis as they are more effective than other approaches. This literature review focuses on the selecting the right machine learning tools based on the levels at which sentiment analysis is performed. Nandwani & Verma

(2021) outlines the three levels at which sentiment analysis is done as the feature level, the aspect level and the document level.

2.0 Related Work

Many related works on this topic provide a general comparison between machine learning algorithms such as Naïve Bayes and Support Vector Machine (SVW). Jagdale et al. (2019) in their research, concluded that the former provided more accurate results for product reviews. Basani et al. (2019) argues that SVM is more accurate however has a shorter execution time when compared to Naïve Bayes. Khairnar and Kinikar (2013) introduces a third algorithm known as maximum entropy in their comparison but also concludes that SVM does significantly better than the other two in text classification.

3.0 Research Focus

This review will focus on the strengths of machine learning algorithm, identifying which one works best for a given level of sentiment analysis. Machine learning practitioners can use the outcome of this research to conduct a more accurate text classification. Data will be collected from a focus group of about 7 professionals consisting of Data Scientists and recent computer science graduates.

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.

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.

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 .639-647. Springer, Singapore.

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.

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.