

Peer Review

Review of: "Sentiment Analysis on Social Media"

Bochra Hadj Kilani¹

1. University of Carthage, Carthage, Tunisia

This paper is all about a really important topic:

The potential of sentiment analysis to address contemporary issues, such as environmental concerns and public health challenges, is a subject of significant interest. The integration of social media into the policy-making process is a notable development, and the present research topic is particularly timely and relevant.

It's a bit hard to tell if it's a research paper or a literature review, since it mixes both. It talks about some of the methods (like dataset pre-processing and machine learning models) like it's an actual study, but it doesn't go into enough detail about the methods, the results, or the evaluation. It also looks at other studies, but it doesn't really talk about their findings in a critical way, which makes it a bit weak as a literature review.

If the goal is a research paper:

you could clearly define the research question and hypothesis, have a well-structured methodology section with all the details about the dataset, model descriptions, and evaluation metrics. Then, it's so important to have an explicit results and discussion section.

And if you're planning to do a literature review:

Just make sure to remove any ambiguous mentions of methodology that suggest an experiment was conducted.

Focus on synthesizing previous research, comparing methods, and identifying trends and gaps.

At the end, you could leave some recommendations for future research. Just remember to keep it positive and encouraging, and don't imply that your paper conducted experiments.

2- Methodology

give more details about the dataset size, how the data was collected, and the sentiment lexicons that were used. And, try to explain why you chose Naïve Bayes over other deep learning models like LSTMs or transformers.

The methodology section could be clearer, with missing elements like data sources, sample size, and model validation.

The study looks at tweets from 2018 to 2022 and uses different machine learning (ML) methods, but it doesn't give exact details on:

Dataset Size and Composition – The study mentions analyzing tweets, but it doesn't say how many were used or how they were distributed.

The paper mentions analyzing tweets between 2018 and 2022, but it would be helpful to know more about the specific details of this data collection. For example, it would be useful to know how many tweets were collected, whether it was a sample of thousands or millions, and what hashtags, keywords, or topics were used for filtering the tweets.

It would also be helpful to know which API or tool was used for data collection, such as the Twitter API, etc...

Additionally, it would be beneficial to understand the geographical scope of the data, whether it is from a specific country or worldwide.

Example of missing detail:

"We investigate how the perception of people has changed from January 2018 to December 2022 and analyze the differences

in the sentiment of the people, if any, pre-pandemic (2018-2019) and post-pandemic (2020-2022)."

What's missing? How many tweets? What filtering criteria?

References Section

The paper cites only 9 references, which is inadequate for either a literature review or a research study.

Conclusion

The paper provides important information about how to analyze sentiments on social media. It focuses on environmental issues and the effects of the pandemic. However, the study could be improved by adding more details about the research method, discussing related studies more in-depth, and analyzing how well the model worked. Making these improvements would make the paper more useful in both academic and practical ways.

Final Recommendation: Major Revision Needed

References: (most cited paper)

Focus on these topics:

- 1- Foundational Studies in Sentiment Analysis
- 2- Studies on COVID-19 and Social Media Sentiment
- 3- Sentiment Analysis in Environmental and Urban Research

To ensure no conflict of interest, I searched for the most important and most cited papers, which serve as examples of widely recognized works but are not an exhaustive list of all relevant literature in sentiment analysis, machine learning, and social media research to provide objective and widely recognized references

[1][2][3][4][5]

References

1. [△]May El Barachi, Manar AlKhatib, Sujith Mathew, Farhad Oroumchian. (2021). A novel sentiment analysis framework for monitoring the evolving public opinion in real-time: Case study on climate change. *Journal of Cleaner Production*, vol. 312, 127820. doi:10.1016/j.jclepro.2021.127820.
2. [△]Dawei Li, Yujia Zhang, Cheng Li. (2019). Mining Public Opinion on Transportation Systems Based on Social Media Data. *Sustainability*, vol. 11 (15), 4016. doi:10.3390/su11154016.

3. [△]Rabindra Lamsal. (2020). Design and analysis of a large-scale COVID-19 tweets dataset. *Appl Intell*, vol. 51 (5), 2790–2804. doi:10.1007/s10489-020-02029-z.
4. [△]Sakun Boon-Itt, Yukolpat Skunkan. (2020). Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study. *JMIR Public Health Surveill*, vol. 6 (4), e21978. doi:10.2196/21978.
5. [△]Usman Naseem, Imran Razzak, Matloob Khushi, Peter W. Eklund, et al. (2021). COVIDSenti: A Large-Scale Benchmark Twitter Data Set for COVID-19 Sentiment Analysis. *IEEE Trans. Comput. Soc. Syst.*, vol. 8 (4), 1003–1015. doi:10.1109/tcss.2021.3051189.

Declarations

Potential competing interests: No potential competing interests to declare.