

Paper Title: Improving Sentiment Analysis in Social Media by Handling Lengthened Words

Paper Link: <https://doi.org/10.1109/ACCESS.2023.3238366>

1. Summary

- 1.1. **Motivation:** The paper explains how people express strong feelings online, especially with words like "happyyyy" or "badddd." These are called "lengthened words." These words are essential for showing real emotions, but current computer systems don't pay much attention to them. So, creating a new system that understands these words better, makes it more accurate when computers try to figure out how people feel on the internet.
- 1.2. **Contribution:** As people often use lengthened words to express emotions online, especially in the context of social media communication, these words enhance the accuracy of sentiment analysis systems. In the current digital age, where social media plays a crucial role in shaping opinions, understanding the true sentiments expressed by individuals is valuable. This research offers a more nuanced approach to sentiment analysis, which can be applied to better comprehend public reactions, opinions, and emotions shared on various online platforms.
- 1.3. **Methodology:** The study uses a multi-phase framework for sentiment analysis. To begin, it tokenizes raw data from several sources. The following phase involves removing stop words with no meaningful significance. The method then normalizes the remaining words and provides sentiment scores based on a reference database. Furthermore, the algorithm evaluates the tail end of words, which frequently contain lengthened words, and gives an increased tail score. These scores are used to get an amplified sentiment score for each document. The technique ensures a nuanced examination of indicated feelings, particularly when people use longer phrases to indicate emotions on social media.
- 1.4. **Conclusion:** In summary, the study highlights the influence of intentional word lengthening on sentiment analysis and presents a sentiment extraction technique. Recognizes the necessity for optimization, takes spelling accuracy into account, and intends to investigate other features and languages in further work.

2. Limitations

- 2.1. **Assumption of Correct Spellings:** The proposed system assumes that spellings in the informal chats are correct. This limitation may affect the accuracy of sentiment analysis, especially considering the prevalence of intentionally misspelled or abbreviated words in online communication.
- 2.2. **Optimization Challenges:** The paper acknowledges that the proposed system is still in the process of optimization to achieve results on par with the gold standard. The ongoing efforts to enhance performance indicate a current limitation in reaching optimal accuracy and effectiveness in sentiment recognition.

3. Synthesis

The study highlights how important it is to record linguistic details to improve sentiment analysis accuracy, especially when it comes to social media. In addition to overcoming the shortcomings of current approaches, the suggested system advances our understanding of online public sentiment and opinion. The study emphasizes how sentiment analysis is constantly changing and how important it is to understand the variety of expressions that are common in digital communication.