

ABSTRACT

Memes are a popular internet form and spread really fast on many social media platforms. People can express their ideas, critics, interests, or aversions through memes. But in some cases, other people may interpret memes differently and feel bad about it. This variation in meme interpretation is a challenge in sentiment analysis, as a meme can be judged negative or positive by different individuals. Therefore, there is a need for an automated system that can consistently predict the sentiment polarity of memes. A meme is a multimodal content that could consist of visual and textual components, which is suitable for sentiment polarity analysis study. To model a system that effectively leverages multimodal features, the model needs to understand the meme features. This study proposes joint deep learning model—BERT and Densenet121—that concatenate text, image, and cluster features based on the extracted face encodings. To assess better context of the texts, BERT was trained with sarcasm dataset. The widely used SemEval 2020 Task 8: Memotion Analysis dataset was used in this study due to its comprehensive annotation of meme-based sentiment and sarcasm, which aligns with this study’s approach. The result demonstrates that the model achieved 0.3047 Macro-F1 score and 0.3738 accuracy, outperform official baseline model by 40%, with improved ability to detect positive and neutral sentiments and reduce sarcastic false negatives compared to the base model. This highlights the effectiveness of integrating sarcasm detection into the model framework for robust sentiment classification in memes.

Keywords: *sentiment classification, meme, deep learning, clustering, sarcasm detection, memotion*