

A Comprehensive Evaluation and Comparative Analysis of Data Mining Techniques for Sentiment Analysis in Social Media

Sakhawia Kaleem Farogh

Student, Master of Computer Application

Late Bhausahab Hiray S.S Trust's Hiray Institute of Computer Application, Mumbai, India

Abstract: *Sentiment analysis, also known as opinion mining, has emerged as a pivotal field in the realm of social media. With the exponential growth of user-generated content on various platforms, understanding and extracting sentiments have become essential for businesses and organizations. In this research paper, we embark on a meticulous journey, delving into a comprehensive evaluation and comparative analysis of various data mining techniques employed for sentiment analysis in social media. The primary objective of this study is to provide practitioners and researchers with valuable insights into the strengths, limitations, and performance metrics of these techniques. By conducting an extensive evaluation, we aim to shed light on the effectiveness of different data mining approaches in capturing sentiments accurately and efficiently. Furthermore, we explore the challenges and limitations associated with sentiment analysis in social media, addressing the intricacies involved in analysing the vast and dynamic landscape of user-generated content.*

Keywords: sentiment analysis, social media, data mining, techniques, evaluation, comparison

Structure of the Paper:

To provide a comprehensive understanding of sentiment analysis in social media, this research paper adopts a structured approach. The subsequent sections are organized as follows:

Section 1: Introduction – describes the background, significance of the chosen topic and the main purpose of selecting this topic. The section contains the elaborated information on data mining techniques.

Section 2: Literature Review - A thorough review of the existing literature on sentiment analysis in social media, exploring the significance of sentiment analysis and examining previous research studies that have evaluated and compared data mining techniques for sentiment analysis

Section 3: Methodology - An in-depth description of the research methodology, including data collection and preprocessing, feature extraction and representation, sentiment classification techniques, and evaluation metrics.

Section 4: Evaluation and Comparison of Techniques - A meticulous evaluation and comparison of various data mining techniques for sentiment analysis in social media, highlighting their strengths, limitations, and performance metrics.

Section 5: Challenges and Limitations - A discussion of the challenges faced when conducting sentiment analysis on social media data, addressing the complexities involved in capturing and interpreting sentiments accurately.

Section 6: Future Scope and Advancements - An exploration of potential future research directions and untapped avenues for improvement in sentiment analysis in social media, paving the way for continued innovation and development in this dynamic field.

Section 7: Conclusion - A summary of the key findings, implications, and applications of sentiment analysis in social media, along with actionable recommendations for practitioners and researchers to navigate this domain effectively.

REFERENCES

- [1]. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135.
- [2]. Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report*, Stanford, 1(12).
- [3]. Kim, Y. (2014). Convolutional neural networks for sentence classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1746-1751.
- [4]. Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011).
- [5]. Learning word vectors for sentiment analysis. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 142-150.
- [6]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS)*, 5998-6008.
- [7]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4171-4186