Decoding Sentiments: Virtue or Vice through Multilingual Paragraph Analysis

Vishal U[1], Veena M V[2], Poornima R M[3]

Students, Department of Information Science and Engineering[1,2]

Assistant Professor, Department of Information Science and Engineering[3]

Global Academy of Technology, Bangalore, India

Abstract: *This work provides a comprehensive overview of recent developments in sentiment analysis methodologies. It explores innovative approaches, including the integration of rule-based sentiment dictionaries, machine learning techniques, and deep learning solutions for financial sentiment analysis. Emphasis is placed on key preprocessing steps such as tokenization, lowercasing, stop words removal, and punctuation elimination. Feature extraction techniques like Bag-of-Words, Word2Vec, and TF-IDF are discussed, highlighting their roles in representing textual information. The abstract delves into model selection, covering traditional machine learning models like Naive Bayes, Support Vector Machines, and Random Forests, as well as deep learning models such as Recurrent Neural Networks, Long Short-Term Memory networks, and BERT. The abstract explains these algorithms in detail, emphasizing their application in sentiment analysis. Training the model through supervised learning and evaluating its performance using metrics like accuracy, precision, recall, and F1 score are outlined. Additionally, a structured approach to paraphrasing is introduced, underlining its significance in creating meaningful representations of text.*

Keywords: sentiment analysis, preprocessing, deep learning, paraphrasing

REFERENCES

[1]. H. Liu, X. Chen and X. Liu, "A Study of the Application of Weight Distributing Method Combining Sentiment Dictionary and TF-IDF for Text Sentiment Analysis," in IEEE Access, vol. 10, pp. 32280-32289, 2022, doi: 10.1109/ACCESS.2022.3160172.

[2]. P. Durga and D. Godavarthi, "Deep-Sentiment: An Effective Deep Sentiment Analysis Using a Decision-Based Recurrent Neural Network (D-RNN)," in IEEE Access, vol. 11, pp. 108433-108447, 2023, doi: 10.1109/ACCESS.2023.3320738.

[3]. Almalis, Ioannis, Eleftherios Kouloumpris, and Ioannis Vlahavas. "Sector-level sentiment analysis with deep learning." *Knowledge-Based Systems* 258 (2022): 109954.

[4]. H. Q. Abonizio, E. C. Paraiso and S. Barbon, "Toward Text Data Augmentation for Sentiment Analysis," in IEEE Transactions on Artificial Intelligence, vol. 3, no. 5, pp. 657-668, Oct. 2022, doi: 10.1109/TAI.2021.3114390.

[5]. D. Deng, L. Jing, J. Yu and S. Sun, "Sparse Self-Attention LSTM for Sentiment Lexicon Construction," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 27, no. 11, pp. 1777-1790, Nov. 2019, doi: 10.1109/TASLP.2019.2933326.

[6]. J. Yu, K. Chen and R. Xia, "Hierarchical Interactive Multimodal Transformer for Aspect-Based Multimodal Sentiment Analysis," in IEEE Transactions on Affective Computing, vol. 14, no. 3, pp. 1966-1978, 1 July-Sept. 2023, doi: 10.1109/TAFFC.2022.3171091.

[7]. M. Huang, H. Xie, Y. Rao, Y. Liu, L. K. M. Poon and F. L. Wang, "Lexicon-Based Sentiment Convolutional Neural Networks for Online Review Analysis," in IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1337-1348, 1 July-Sept. 2022, doi: 10.1109/TAFFC.2020.2997769.

[8]. Z. Ren, G. Zeng, L. Chen, Q. Zhang, C. Zhang and D. Pan, "A Lexicon-Enhanced Attention Network for Aspect-Level Sentiment Analysis," in IEEE Access, vol. 8, pp. 93464-93471, 2020, doi: 10.1109/ACCESS.2020.2995211.

[9]. Y. -C. Tsai and F. -C. Lin, "Paraphrase Generation Model Integrating Transformer Architecture, Part-of-Speech Features, and Pointer Generator Network," in IEEE Access, vol. 11, pp. 30109-30117, 2023, doi: 10.1109/ACCESS.2023.3260849.

[10]. D. Zeng, H. Zhang, L. Xiang, J. Wang and G. Ji, "User-Oriented Paraphrase Generation With Keywords Controlled Network," in IEEE Access, vol. 7, pp. 80542-80551, 2019, doi: 10.1109/ACCESS.2019.2923057.

[11]. Y. Dong, Y. Fu, L. Wang, Y. Chen, Y. Dong and J. Li, "A Sentiment Analysis Method of Capsule Network Based on BiLSTM," in IEEE Access, vol. 8, pp. 37014-37020, 2020, doi: 10.1109/ACCESS.2020.2973711.

[12]. Punetha, Neha, and Goonjan Jain. "Bayesian game model based unsupervised sentiment analysis of product reviews." *Expert Systems with Applications* 214 (2023): 119128.

[13]. M. E. Basiri et al., "Improving Sentiment Polarity Detection Through Target Identification," in IEEE Transactions on Computational Social Systems, vol. 7, no. 1, pp. 113-128, Feb. 2020, doi: 10.1109/TCSS.2019.2951326.

[14]. S. M. Al-Ghuribi, S. A. Mohd Noah and S. Tiun, "Unsupervised Semantic Approach of Aspect-Based Sentiment Analysis for Large-Scale User Reviews," in IEEE Access, vol. 8, pp. 218592-218613, 2020, doi: 10.1109/ACCESS.2020.3042312.

[15]. P. Vyas, M. Reisslein, B. P. Rimal, G. Vyas, G. P. Basyal and P. Muzumdar, "Automated Classification of Societal Sentiments on Twitter With Machine Learning," in IEEE Transactions on Technology and Society, vol. 3, no. 2, pp. 100-110, June 2022, doi: 10.1109/TTS.2021.3108963.

[16]. L. Huang, Z. Dou, Y. Hu and R. Huang, "Online Sales Prediction: An Analysis With Dependency SCOR-Topic Sentiment Model," in IEEE Access, vol. 7, pp. 79791-79797, 2019, doi: 10.1109/ACCESS.2019.2919734.

[17]. X. Li, F. Zeng and C. Yao, "A Semi-Supervised Paraphrase Identification Model Based on Multi-Granularity Interaction Reasoning," in IEEE Access, vol. 8, pp. 60790-60800, 2020, doi: 10.1109/ACCESS.2020.2984009.

[18]. J. Liu et al., "Noun Compound Interpretation With Relation Classification and Paraphrasing," in IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 9, pp. 8757-8769, 1 Sept. 2023, doi: 10.1109/TKDE.2022.3208617.

[19]. K. Mrinalini, P. Vijayalakshmi and T. Nagarajan, "SBSim: A Sentence-BERT Similarity-Based Evaluation Metric for Indian Language Neural Machine Translation Systems," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 30, pp. 1396-1406, 2022, doi: 10.1109/TASLP.2022.3161160.

[20]. J. Yu, K. Chen and R. Xia, "Hierarchical Interactive Multimodal Transformer for Aspect-Based Multimodal Sentiment Analysis," in IEEE Transactions on Affective Computing, vol. 14, no. 3, pp. 1966-1978, 1 July-Sept. 2023, doi: 10.1109/TAFFC.2022.3171091.

[21]. Zhang, Lumin, et al. "User-level sentiment evolution analysis in microblog." *China Communications* 11.12 (2014): 152-163.hn

[22]. N. Li, C. -Y. Chow and J. -D. Zhang, "SEML: A Semi-Supervised Multi-Task Learning Framework for Aspect-Based Sentiment Analysis," in IEEE Access, vol. 8, pp. 189287-189297, 2020, doi: 10.1109/ACCESS.2020.3031665.

[23]. J. Z. Maitama, N. Idris, A. Abdi, L. Shuib and R. Fauzi, "A Systematic Review on Implicit and Explicit Aspect Extraction in Sentiment Analysis," in IEEE Access, vol. 8, pp. 194166-194191, 2020, doi: 10.1109/ACCESS.2020.3031217.

[24]. L. -C. Yu, J. Wang, K. R. Lai and X. Zhang, "Refining Word Embeddings Using Intensity Scores for Sentiment Analysis," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 26, no. 3, pp. 671-681, March 2018, doi: 10.1109/TASLP.2017.2788182.

[25]. S. Rida-E-Fatima et al., "A Multi-Layer Dual Attention Deep Learning Model With Refined Word Embeddings for Aspect-Based Sentiment Analysis," in IEEE Access, vol. 7, pp. 114795-114807, 2019, doi: 10.1109/ACCESS.2019.2927281.

[26]. J. He, A. Wumaier, Z. Kadeer, W. Sun, X. Xin and L. Zheng, "A Local and Global Context Focus Multilingual Learning Model for Aspect-Based Sentiment Analysis," in IEEE Access, vol. 10, pp. 84135-84146, 2022, doi: 10.1109/ACCESS.2022.3197218.

[27]. K. R. Mabokela, T. Celik and M. Raborife, "Multilingual Sentiment Analysis for Under-Resourced Languages: A Systematic Review of the Landscape," in IEEE Access, vol. 11, pp. 15996-16020, 2023, doi: 10.1109/ACCESS.2022.3224136.

[28]. K. Chakraborty, S. Bhattacharyya and R. Bag, "A Survey of Sentiment Analysis from Social Media Data," in IEEE Transactions on Computational Social Systems, vol. 7, no. 2, pp. 450-464, April 2020, doi: 10.1109/TCSS.2019.2956957.

[29]. S. Bengesi, T. Oladunni, R. Olusegun and H. Audu, "A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets," in IEEE Access, vol. 11, pp. 11811-11826, 2023, doi: 10.1109/ACCESS.2023.3242290.

[30]. Kalamatianos, Georgios, et al. "Towards the creation of an emotion lexicon for microblogging." *Journal of Systems and Information Technology* 20.2 (2018): 130-151.