Comparative Analysis of Machine Learning Models for Climate Change Sentiment Classification: A Multi-Method Evaluation Study

Climate change discourse significantly impacts public policy formation and environmental awareness (Carmichael & Brulle, 2017). Understanding public sentiment trends is essential for policymakers and environmental organizations to develop effective communication strategies and gauge support for climate initiatives. While previous studies have analyzed climate sentiment on Twitter and Reddit over shorter periods, this study provides a comprehensive 12-year longitudinal analysis comparing machine learning approaches for sentiment classification using specialized forum data with sustained, in-depth climate discussions.

The analysis examined sentiment patterns using data from the RC Groups forum spanning 2011-2024, comprising 517,468 comments. RC Groups was selected for its unique characteristics: sustained user engagement over decades, technical discussions by hobbyists with diverse climate perspectives, and focused debates around "Do you think mankind's burning of fossil fuels causes global warming?" Unlike broader social media platforms, this specialized community provides deeper, more substantive climate discussions ideal for longitudinal sentiment analysis. Four machine learning approaches were evaluated: VADER and TextBlob (rule-based methods), and BERT and ClimateBERT (deep learning methods). VADER uses valence-aware lexicon matching optimized for social media text (Hutto & Gilbert, 2014), while BERT applies a fine-tuned transformer architecture with bidirectional context understanding (Devlin et al., 2019). Ground truth validation was established through manual annotation of 1,000 randomly selected samples across positive, negative, and neutral categories.

Rule-based methods significantly outperformed transformer-based approaches. TextBlob achieved the highest performance with 43.36% accuracy, demonstrating balanced precision and recall across sentiment categories. VADER showed consistent 41% accuracy across all metrics. Deep learning methods underperformed expectations, with BERT achieving 33% accuracy and ClimateBERT recording 29% accuracy despite climate-specific pre-training. The transformer models' poor performance stems from their training on formal text corpora, making them poorly suited for informal forum discussions with domain-specific terminology and conversational patterns. Transformer models exhibited strong negative sentiment bias, misclassifying 60% of positive and neutral comments.Temporal analysis revealed significant sentiment evolution, with positive sentiment peaking in 2013 and declining in recent years, contrasting with shorter-term studies that miss these long-term trends. Error analysis identified sarcasm as the primary performance bottleneck, affecting approximately 35% of misclassified samples across all models (Joshi et al., 2017).

This study demonstrates that climate sentiment analysis requires careful model selection based on domain-specific characteristics rather than defaulting to state-of-the-art transformer models. Rule-based approaches proved more robust due to their ability to handle domain-specific lexicons and informal language patterns. The superior performance of simpler methods indicates that climate sentiment analysis benefits from explicit sentiment indicators rather than complex contextual understanding. These findings are crucial for selecting appropriate algorithms for climate communication analysis and environmental policy assessment applications.

Keywords: Sentiment Analysis, Climate Change, Machine Learning Comparison, Social Media Analysis, Natural Language Processing.

References

[1] Carmichael, J. T., & Brulle, R. J. (2017). Elite cues, media coverage, and public concern: An integrated path analysis of public opinion on climate change, 2001–2013. *Environmental Politics*, 26(2), 232-252.

[2] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL-HLT*.

[3] Hutto, C., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. *ICWSM*, 8(1), 216-225.

[4] Joshi, A., Bhattacharyya, P., & Carman, M. J. (2017). Automatic sarcasm detection: A survey. *ACM Computing Surveys*, 50(5), 1-22.

[5] Webersinke, N., Kraus, M., Bingler, J. A., & Leippold, M. (2022). ClimateBERT: A pretrained language model for climate-related text. *arXiv preprint arXiv:2110.12010*.