ISSN: 2455 - 9172
EIJO Journal of Engineering, Technology And Innovative Research (EIJO–JETIR)
Einstein International Journal Organization (EIJO)
Available Online at: www.eijo.in
Volume – 10, Issue –3 , May - June - 2025 , Page No. : 01 - 07

Sentiment Analysis on Social Media Using Transformer Models

[1]Ankit Kumar Taneja, Assistant Professor, Department of Computer Science, Arya College of Engineering, Jaipur

[2]Parth Sharma, Research Scholar, Department of Computer Science, Arya College of Engineering, Jaipur

[3]Naveen Banjara, Research Scholar, Department of Computer Science, Arya College of Engineering, Jaipur

Abstract

This paper explores the application of transformer models for sentiment analysis on social media platforms. Social media presents unique challenges for sentiment classification due to informal language, slang, and sarcastic expressions. Transformer models like BERT, RoBERTa, and DistilBERT, through their contextual understanding and attention mechanisms, outperform traditional methods in capturing sentiment from noisy text. The study outlines the methodology of applying these models, including data preprocessing, model training, and evaluation. It also discusses the challenges posed by bias, computation cost, and ambiguity in social media language. The paper concludes by highlighting future trends, including multilingual sentiment analysis, few-shot learning, and model explainability, positioning transformers as the future of real-time opinion mining.

Keywords: Sentiment Analysis, Social Media, Transformer Models, BERT, RoBERTa, NLP, Emotion Detection, Self-Attention, Fine-Tuning, Sarcasm Detection, Multilingual Analysis

Introduction

Social media platforms like Twitter, Facebook, Instagram, and Reddit have become major sources of public opinion and sentiment. The massive volume of user-generated content on these platforms offers valuable insights into public mood, brand perception, political opinion, and emerging societal trends. Sentiment analysis, a subfield of Natural Language Processing (NLP), focuses on identifying and categorizing opinions expressed in text to determine whether the attitude is positive, negative, or neutral.

Traditional sentiment analysis methods relied on lexicons, rule-based systems, and classical machine learning techniques such as Naïve Bayes or Support Vector Machines (SVM). While these methods performed reasonably well for simple texts, they struggled with informal language, sarcasm, slang, emojis, and the complex syntactic structures commonly found in social media posts.

The advent of deep learning brought a paradigm shift to sentiment analysis. Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), and Convolutional Neural Networks (CNNs) started delivering better results due to their ability to learn contextual and sequential patterns in text. However, these models still had limitations when it came to understanding long-term dependencies or parallelizing computations effectively.

Transformer models, introduced with the landmark paper *"Attention Is All You Need"* (Vaswani et al., 2017), revolutionized NLP by introducing a self-attention mechanism that allows models to focus on relevant parts of the input text regardless of their position. Transformers such as BERT (Bidirectional Encoder Representations from Transformers), RoBERTa, XLNet, and DistilBERT have since dominated benchmarks for numerous NLP tasks, including sentiment analysis.

The application of transformer models to social media sentiment analysis presents several advantages: contextual understanding, robustness to noisy data, scalability, and state-of-the-art accuracy. These models can handle informal language, abbreviations, emojis, and even code-mixed texts (e.g., Hinglish), making them highly suitable for analyzing user sentiment in real-world online conversations.

This paper explores how transformer models are applied to sentiment analysis on social media platforms. It discusses data preprocessing techniques, model selection, training strategies, evaluation metrics, challenges, and future directions. Through empirical analysis and literature review, this study demonstrates that transformer-based sentiment analysis significantly outperforms traditional approaches, especially on social media text.

Overview of Transformer Models

Transformer models have redefined the landscape of natural language processing by enabling machines to understand and generate human language with remarkable accuracy. The key innovation of transformer architecture is the self-attention mechanism, which allows the model to weigh the importance of different words in a sentence irrespective of their distance from one another. This capability is especially useful in analyzing complex and context-dependent texts, such as those found on social media.

The Transformer Architecture

Introduced by Vaswani et al. in 2017, the original transformer architecture consists of an encoder and a decoder, both built using layers of self-attention and feed-forward neural networks. For sentiment analysis, especially classification tasks, only the encoder is typically used. The architecture removes the recurrence mechanism found in RNNs and LSTMs, enabling better parallelization and efficiency during training.

Popular Transformer Models for Sentiment Analysis

Several transformer-based models have been fine-tuned for sentiment classification:

- **BERT (Bidirectional Encoder Representations from Transformers)**: Developed by Google, BERT is pretrained using masked language modeling and next sentence prediction. Its bidirectional nature allows it to capture context from both directions, making it powerful for sentiment tasks.
- **RoBERTa (Robustly Optimized BERT Pretraining Approach)**: RoBERTa improves upon BERT by training on more data for longer durations and removing the next sentence prediction objective. It typically achieves better performance than BERT on classification tasks.
- **DistilBERT**: A smaller and faster version of BERT, it maintains 95% of BERT’s performance while being significantly more efficient—ideal for real-time applications.
- **XLNet**: Combines the best of BERT and autoregressive models by using permutation-based language modeling, capturing bidirectional context while maintaining autoregressive training benefits.
- **ALBERT**: An optimized version of BERT that reduces memory consumption through parameter sharing and factorized embeddings.

Pretraining and Fine-Tuning

Transformer models are generally pretrained on large corpora such as Wikipedia, BooksCorpus, or Common Crawl. For sentiment analysis, these models are then fine-tuned on labeled datasets such as Sentiment140, IMDb, or domain-specific Twitter sentiment corpora. Fine-tuning adjusts the model weights to perform well on sentiment prediction using relatively little labeled data.

Overall, transformer models bring unparalleled contextual awareness and robustness to sentiment analysis, particularly for the noisy, diverse, and informal language styles seen on social media platforms.

Methodology and Data Preprocessing

The success of transformer-based sentiment analysis on social media hinges not only on model architecture but also on how the data is preprocessed and the training pipeline is set up. Social media text presents unique challenges including informal language, slang, abbreviations, emojis, hashtags, and non-standard grammar.

Data Collection

The data used for sentiment analysis typically comes from social media platforms like Twitter, Reddit, or Facebook. Common benchmark datasets include:

- **Sentiment140** – 1.6 million tweets labeled as positive or negative.
- **SemEval datasets** – for nuanced sentiment tasks including sarcasm and emotion detection.
- **Twitter US Airline Sentiment Dataset** – focused on sentiment toward airlines.

APIs like the **Twitter API** or **Pushshift API** (for Reddit) are often used to gather real-time or historical data. This raw data then goes through a robust preprocessing pipeline.

Data Preprocessing Steps

1. **Text Cleaning**: Remove URLs, user mentions (@user), hashtags, special characters, and extra spaces. Hashtags may also be split into individual words using libraries like wordsegment.
2. **Emoji Handling**: Emojis are either removed or converted to text using tools like emoji.demojize(), which translates emojis into descriptive words (e.g., □ → "smiling_face").
3. **Stopword Removal and Tokenization**: While stopwords (e.g., “the”, “is”) may be removed in traditional models, transformer models like BERT are pretrained on full sentences, so retaining stopwords can help preserve context.
4. **Lowercasing and Lemmatization**: Standardize the text while preserving semantic meaning.
5. **Handling Code-Mixing and Spelling Variations**: Spelling correction tools or code-mixed language support (like Hinglish) is considered when the dataset includes multilingual input.

Evaluation Metrics

Evaluating the performance of sentiment analysis models requires a set of well-defined metrics to determine how effectively the model classifies sentiment from text. Given the often imbalanced and subjective nature of social media sentiment datasets, multiple evaluation criteria are used to obtain a comprehensive assessment.

Accuracy

Accuracy is the most basic metric, defined as the ratio of correctly predicted labels to the total number of predictions. While intuitive, it can be misleading in imbalanced datasets where one class dominates (e.g., neutral sentiment may form the majority).

Precision, Recall, and F1-Score

These metrics provide a better understanding of model performance in each class.

- **Precision**: The number of true positive predictions divided by all positive predictions. High precision means fewer false positives.
- **Recall**: The number of true positives divided by the total actual positives. High recall means fewer false negatives.
- **F1-Score**: The harmonic mean of precision and recall. A balanced metric when both false positives and false negatives are important.

For multi-class sentiment (positive, negative, neutral), these metrics are calculated **per class** and averaged using:

- **Macro average** (equal weight to each class),
- **Micro average** (aggregate of contributions),
- **Weighted average** (accounts for class imbalance).

Confusion Matrix

A confusion matrix provides a visual representation of true vs. predicted sentiment labels. It helps identify where the model is making the most errors (e.g., confusing neutral and negative tweets).

AUC-ROC

For binary sentiment classification, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) is useful in visualizing model performance across thresholds. A higher AUC indicates better separability between classes.

Cross-Validation

K-fold cross-validation is often employed to ensure that the model's performance is not specific to a particular train-test split. This increases the reliability of performance metrics.

Error Analysis

Beyond metrics, qualitative error analysis is essential. By manually reviewing misclassified examples, one can determine if errors are due to sarcasm, ambiguous language, or out-of-vocabulary terms. This guides future data augmentation or model improvements.

These metrics collectively offer a well-rounded view of model performance and help in selecting the most robust transformer configuration for real-world deployment.

Challenges and Limitations

Despite their strong performance, transformer models face several challenges when applied to sentiment analysis on social media. These limitations need to be considered for effective deployment.

1. Informal and Noisy Language

Social media text is often grammatically incorrect, includes slang, code-switching (mixing languages), emojis, and abbreviations. Even though transformer models can handle some noise due to pretraining on web text, extreme cases still lead to misinterpretation.

2. Sarcasm and Irony

Detecting sarcasm is a longstanding challenge in sentiment analysis. A tweet like "Great job, government □" might express negative sentiment despite using words typically associated with praise. Transformers may misclassify such nuanced expressions.

3. Class Imbalance

Most datasets have a skewed distribution, with a majority of neutral or positive sentiments. This imbalance affects model training, resulting in poor performance on minority classes unless addressed through data augmentation or weighted loss functions.

4. Computational Cost

Transformers, especially large ones like BERT-large or XLNet, require significant memory and processing power. This limits their deployment on mobile devices or in low-resource settings. Model distillation (e.g., DistilBERT) helps, but there's still a trade-off between performance and efficiency.

5. Data Privacy and Bias

Pretrained models may contain hidden biases due to the datasets they were trained on (e.g., gender, race, or political bias). Sentiment predictions can thus reflect these biases if not carefully handled, leading to ethical concerns in deployment.

6. Limited Domain Adaptation

Even though transformers are pretrained on a broad corpus, they may still underperform on domain-specific data (e.g., technical product reviews, regional dialects). Fine-tuning helps, but may require labeled data which is often limited.

7. Real-time Sentiment Analysis

Real-time analysis requires models to be both accurate and fast. Transformers, especially in their original form, are not optimal for such settings without hardware acceleration or model optimization.

Addressing these challenges requires research in model compression, adversarial training, and specialized datasets, along with interdisciplinary efforts involving linguists and domain experts.

Conclusion and Future Scope

Sentiment analysis using transformer models has brought significant advances to the field of Natural Language Processing. The ability of models like BERT, RoBERTa, and XLNet to understand context and semantics makes them ideally suited for processing informal and nuanced social media text.

Transformer-based models outperform traditional machine learning and deep learning methods in terms of accuracy, contextual understanding, and generalization. Their strength lies in bidirectional context capture and scalability across languages and domains.

Future Scope

The future of sentiment analysis using transformers is promising. Areas of development include:

- **Multilingual Sentiment Models**: Transformers like mBERT and XLM-R are being fine-tuned to analyze sentiment across multiple languages, including code-mixed data.
- **Zero-shot and Few-shot Learning**: Leveraging models like GPT and T5 to perform sentiment analysis with minimal labeled data.
- **Emotion Detection and Sarcasm Handling**: Moving beyond polarity classification to detect emotions and sarcasm through multi-task learning and richer datasets.
- **Explainability**: Building interpretable sentiment models using attention visualizations or post-hoc methods like LIME and SHAP.

- **Edge Deployment**: Research into compressing transformer models for deployment on mobile and IoT devices without sacrificing performance.

As transformer models become more efficient and accessible, they are poised to become the de facto standard for sentiment analysis across industries—ranging from marketing and finance to politics and public health.

References

1. Kokab, S. T., Asghar, S., & Naz, S. (2022). Transformer-based deep learning models for the sentiment analysis of social media data. Array, 14, 100157.
2. Rodríguez-Ibánez, M., Casánez-Ventura, A., Castejón-Mateos, F., & Cuenca-Jiménez, P. M. (2023). A review on sentiment analysis from social media platforms. Expert Systems with Applications, 223, 119862.
3. Prasanthi, K. N., Madhavi, R. E., Sabarinadh, D. N. S., & Sravani, B. (2023, April). A novel approach for sentiment analysis on social media using BERT & ROBERTA transformer-based models. In 2023 IEEE 8th International Conference for Convergence in Technology (I2CT) (pp. 1-6). IEEE.
4. Tan, K. L., Lee, C. P., Anbananthen, K. S. M., & Lim, K. M. (2022). RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network. IEEE Access, 10, 21517-21525.
5. Bozanta, A., Angco, S., Cevik, M., & Basar, A. (2021, December). Sentiment analysis of stocktwits using transformer models. In 2021 20th IEEE international conference on machine learning and applications (ICMLA) (pp. 1253-1258). IEEE.
6. Almalki, S. S. (2025). Sentiment Analysis and Emotion Detection Using Transformer Models in Multilingual Social Media Data. International Journal of Advanced Computer Science & Applications, 16(3).
7. Mandal, R., Chen, J., Becken, S., & Stantic, B. (2023). Tweets topic classification and sentiment analysis based on transformer-based language models. Vietnam Journal of Computer Science, 10(02), 117-134.
8. Naseem, U., Razzak, I., Musial, K., & Imran, M. (2020). Transformer based deep intelligent contextual embedding for twitter sentiment analysis. Future Generation Computer Systems, 113, 58-69.
9. BANDARUPALLI, G. (2025). Enhancing Sentiment Analysis in Multilingual Social Media Data Using Transformer-Based NLP Models A Synthetic Computational Study. Authorea Preprints.
10. Mukherjee, S., & Das, S. (2023). Application of transformer-based language models to detect hate speech in social media. Journal of Computational and Cognitive Engineering, 2(4), 278-286.
11. Bashiri, H., & Naderi, H. (2024). Comprehensive review and comparative analysis of transformer models in sentiment analysis. Knowledge and Information Systems, 66(12), 7305-7361.
12. Gong, X., Ying, W., Zhong, S., & Gong, S. (2022). Text sentiment analysis based on transformer and augmentation. Frontiers in Psychology, 13, 906061.
13. Lubis, A. R., Fatmi, Y., & Witarsyah, D. (2023, September). Comparison of Transformer Based and Traditional Models on Sentiment Analysis on Social Media Datasets. In 2023 6th International Conference of Computer and Informatics Engineering (IC2IE) (pp. 163-168). IEEE.
14. Pipalia, K., Bhadja, R., & Shukla, M. (2020, December). Comparative analysis of different transformer based architectures used in sentiment analysis. In 2020 9th international conference system modeling and advancement in research trends (SMART) (pp. 411-415). IEEE.

15. Ashayeri, M., & Abbasabadi, N. (2024). Unraveling energy justice in NYC urban buildings through social media sentiment analysis and transformer deep learning. Energy and Buildings, 306, 113914.
16. Kannan, M. J., Chauhan, P., Chauhan, V., Gautam, A., Dwivedi, S., Bisht, S. S., & Sharma, S. (2025). Transformer Models for Deep Learning-Based Sentiment Analysis of Social Media Comments Using Hybrid CNN-BiLSTM Networks.
17. AlBadani, B., Shi, R., Dong, J., Al-Sabri, R., & Moctard, O. B. (2022). Transformer-based graph convolutional network for sentiment analysis. Applied Sciences, 12(3), 1316.