Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/32/20230218

Employing the BERT model for sentiment analysis of online commentary

Bowen Li[1,4], Xiaolu Liu[2] and Ruijia Zhang[3]

[1]International Business School, Henan University, Zhengzhou, 451460, China
[2]College of Computer Science and Technology, National University of Defense Technology, Changsha, 410073, China
[3]College of Artificial Intelligence, Tianjin University of Science and Technology, Tianjin, 300457, China

[4]224240823@henu.edu.cn

Abstract. The objective of this research is to carry out a tone and semantic sentiment analysis of network comments on new media platforms by leveraging the BERT model. With the burgeoning popularity of social media, network comments, rich in emotional and tonal features, have emerged as a significant part of the online culture. Accurate interpretation and analysis of these comments' sentiment and semantic meanings are paramount to grasping online public opinion and user psychology. In this study, the BERT model, lauded for its bidirectional encoding and contextual understanding capabilities, is selected to scrutinize the sentiment and tone of network comments on new media platforms. Through a process of pre-training and fine-tuning, the sentiment attitudes and polarity of comments are accurately identified along with their conveyed tonal features, such as joy, anger, and sarcasm. Conducting an accurate tone and semantic sentiment analysis of network comments on new media platforms facilitates a profound understanding of user preferences and trends in public opinion. This can assist in optimizing content recommendations, enhancing user experiences, and increasing the operational effectiveness of new media platforms. The outcomes of this research will bear significant implications for studies and applications in online culture, offering invaluable references and guidance in related domains.

Keywords: sentiment analysis, BERT model, new media platforms.

1. Introduction

As social media gains traction and the internet continues to rapidly evolve, online comments have become a crucial element of daily communication. They are typified by concise language, varied expressions, and often harbor an abundance of emotional information. Accordingly, sentiment analysis of online comments has grown into a challenging but practically significant task. The ability to discern users' sentiment towards specific topics, products, or events can yield valuable insights for areas such as social media marketing and public opinion monitoring.

An effective approach to enhancing various natural language processing tasks involves language model pre-training [1]. The BERT (Bidirectional Encoder Representations from Transformers) model, a pre-trained deep bidirectional representation model, captures contextual information from both left

DOI: 10.54254/2755-2721/32/20230218

and right contexts, enabling rich language representations to be learned from unlabeled text. BERT has achieved remarkable success across a multitude of natural language processing tasks, including question answering and language inference [2]. This research, therefore, seeks to harness the BERT model for sentiment analysis of online comments, aiming to refine its performance through fine-tuning. The proposed experimental methodology to conduct sentiment analysis on online comments involves first curating a representative dataset comprising notable online comment data from various recent social media platforms. Subsequently, the dataset undergoes preprocessing, involving the extraction of pertinent features and the incorporation of specialized tokens. The processed dataset is then employed to train the BERT model, followed by fine-tuning to boost the accuracy of sentiment classification. The primary aim of this experimental phase is to optimize the performance of the BERT model on sentiment analysis tasks, particularly pertaining to online colloquialisms. By improving the accuracy of sentiment classification, it concurrently seeks to enhance the model's capability to comprehend and capture nuanced emotions, such as surprise or disgust. This goal is addressed by implementing several techniques, including dataset preprocessing and filtering, adjustments to the BERT model architecture parameters, refinement of the loss function algorithm, optimization of training batch configurations, and the implementation of measures to counter overfitting. The model demonstrated exemplary performance on the test dataset, meeting the predetermined accuracy and F1 score benchmarks. This affirms the effectiveness of the approach for sentiment classification tasks. Additionally, a confusion matrix is utilized to analyze the model's predictive outcomes across various sentiment categories, offering deeper insights into its performance. Sentiment analysis was conducted on online comments using the BERT model, introducing a novel mixed loss function, CombinedLoss. This function melds weighted cross-entropy loss with standard cross-entropy loss and includes L2 regularization to combat overfitting. By leveraging different weights during model training to counter data imbalances across sentiment categories, performance improvement was observed across all categories. After extensive training and testing, the model yielded satisfactory results in the context of sentiment analysis tasks. This research illustrates that a combined loss function, incorporating weights and L2 regularization, can effectively enhance sentiment analysis performance when fine-tuning the BERT model. Importantly, this methodology proves valuable for tackling class imbalance scenarios and improves the model's recognition of minority sentiment categories. This work presents a feasible approach to the sentiment analysis of online comments and lays the groundwork for further exploration of sentiment analysis applications in social media marketing and public opinion monitoring. The findings of this research can act as a beneficial reference for researchers and practitioners in related fields, spurring the development and application of sentiment analysis techniques in practical environments.

2. Related works

Sentiment analysis encompasses the examination of sentiments, opinions, attitudes, and emotions expressed toward specific entities such as topics, products, individuals, and organizations. The goal is discerning the author's viewpoint [3]. A vast body of research has been conducted in this domain, exploring diverse approaches from rule-based methods and bag-of-words techniques to machine learning algorithms [4]. For instance, Peter D. Turney introduced the concept of semantic orientation for unsupervised classification, analyzing sentiments in comments to determine positive or negative orientation [5]. A machine learning-based method combined with semantic sentiment analysis for extracting predictions of suicidal ideation using Twitter data was proposed by Marouane Birjali et al. Furthermore, Penalver-Martinez et al [6] implemented a semantic ontology approach to boost feature extraction effectiveness and applied vector analysis techniques for movie review sentiment analysis.

The crux of most machine learning-based sentiment analysis research is the enhancement of feature extraction algorithms [7]. Notably, Zichao Yang et al. proposed a hierarchical attention network for boosting document-level sentiment analysis via optimized feature extraction algorithms [8]. Moreover, Soujanya Poria et al. refined the feature extraction algorithm by computing the mutual information between features and sentiment categories, thus selecting the most informative feature set to boost sentiment analysis efficacy. In addressing these challenges, this research applies techniques such as

DOI: 10.54254/2755-2721/32/20230218

dataset preprocessing and filtering, fine-tuning of BERT model parameters, enhancement of the loss function algorithm, training batch configuration, and mitigation of data overfitting risks among other approaches.

3. Proposed methods

The objective was to conduct sentiment analysis of internet slang utilizing the BERT model. A selection of representative internet slang data was gathered from contemporary social platforms. Following data integration, the dataset underwent pre-processing, involving feature extraction from the data and the inclusion of special tokens. The processed dataset was then fed into the BERT model for training, with subsequent fine-tuning of the model to enhance its accuracy in sentiment determination. The comprehensive workflow of the model is illustrated in Figure 1.

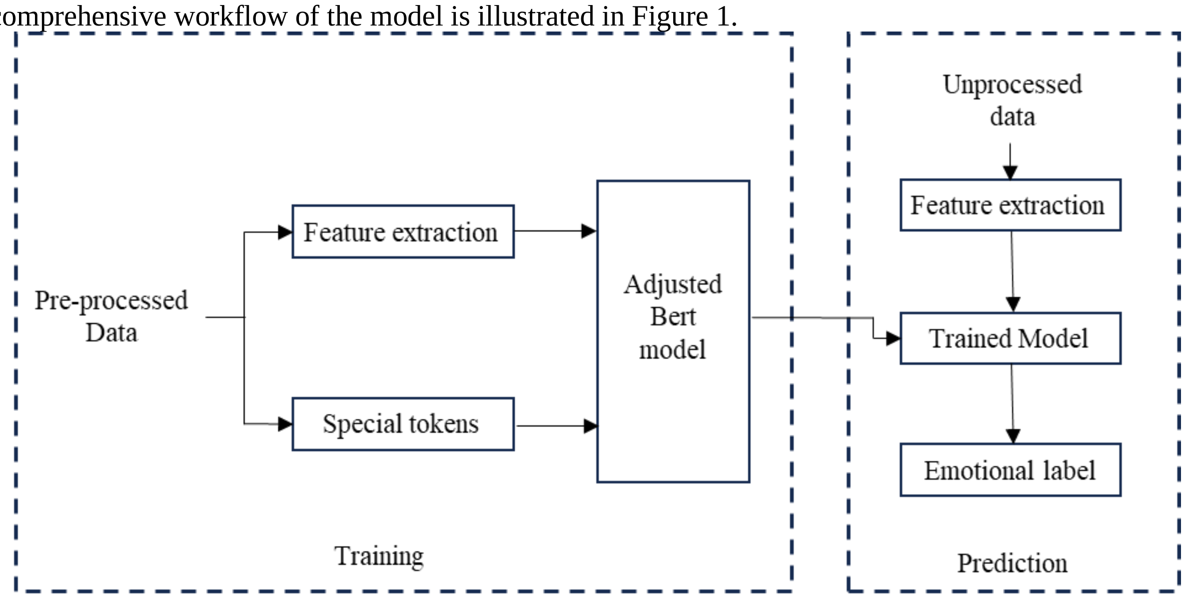

Figure 1. The overall run of the model (Photo/Picture credit: Original).

Table 1. The corpus statistics and label distribution [9].

Emotion	Number	Percentage (%)
Like	4540	11.45
Happiness	9959	25.11
Sadness	14052	35.43
Anger	4562	11.50
Disgust	4876	12.29
Fear	661	1.67
Surprise	1011	2.55
Sum	39661	100.0

3.1. Pre-processing

PyTorch and the Transformers library were utilized for the implementation of the BERT model [1]. A dataset class was designed specifically to transform the text and labels into a format compatible with the BERT model. With the aid of BERT's tokenizer, the text was tokenized and the tokenized results were converted into IDs within the vocabulary. Furthermore, special tokens such as '[CLS]' and '[SEP]' were incorporated, and padding or truncation was applied to ensure a fixed-length input sequence. During the experiment phase, extensive data preprocessing steps were undertaken on the collected internet slang dataset, which ranged from cleaning the data and handling missing values to tokenizing the text. These measures were aimed at enhancing the cleanliness and consistency of the data, with the ultimate goal of improving model performance during both training and prediction phases.

3.2. Dataset
A significant quantity of internet slang data is sourced from an open-source database on GitHub. Stored within the data directory, this dataset serves as an emotion analysis corpus, with each sample meticulously annotated with one sentiment label [9]. The sentiment labels, manually assigned, span seven distinct emotions: 'happiness', 'sadness', 'anger', 'disgust', 'fear', 'surprise', and 'like'. This broad range of data provides a comprehensive portrayal of varying sentiment expressions in internet slang. The dataset has been divided into training, validation, and testing sets in an 8:1:1 ratio and is encoded in UTF-8. The specific quantity and percentage of slang for each emotion are detailed in Table 1 [10].

3.3. Sentiment analysis
The BERT model, an acronym for Bidirectional Encoder Representations from Transformers, is utilized for sentiment analysis. As a pre-trained deep learning model, BERT has been proven to achieve outstanding results in various natural language processing tasks. The 'BERT-base-uncased' variant is selected for this study, incorporating 12 layers of Transformer architecture, 110M parameters, and a lowercase English vocabulary. Such a choice leverages the robust capabilities of BERT in grasping contextual information and discerning sentiment within internet slang.

3.4. Classification
In this classification task, the BERTForSequenceClassification class is utilized, merging the BERT model with an overlaying classification layer. This model configuration encompasses seven categories, aligning with the seven sentiment labels present in the dataset. To optimize model performance during training, the Adam optimizer and the cross-entropy loss function are employed.

3.5. Training and evaluation
After preprocessing the data, we train the BERT model using the preprocessed dataset. The training process involves iterating over the dataset, adjusting the model's weights through backpropagation, and fine-tuning the model's parameters to optimize its performance. We monitor the training progress, including the loss values and accuracy, to ensure the model's convergence.

Once the model is trained, we evaluate its performance on the validation set. This evaluation involves computing various metrics, such as accuracy and F1 score, which provide insights into the model's ability to correctly predict sentiment labels. Additionally, we visualize the confusion matrix to gain a deeper understanding of the model's performance across different sentiment categories. To evaluate the performance of the trained model, we split the dataset into a training set and a validation set. The training set is used to train the model, while the validation set serves as an independent benchmark for evaluating its performance. We employ accuracy as the evaluation metric, which measures the proportion of correctly predicted sentiment labels.

4. Experiments

4.1. Experimental procedure
The goal of these experiments is to enhance the BERT model, boosting its performance in sentiment analysis tasks, especially when dealing with internet slangs. Simultaneously, these enhancements aim to improve the accuracy of sentiment classification and the capability to comprehend complex sentiments such as surprise and disgust. The primary research question posed is: "How can the accuracy of BERT models be improved for sentiment analysis of the internet slang corpus?" This question is approached through several methods including preprocessing and filtering of the dataset, adjusting the model parameters of the BERT model, enhancing the loss function algorithm, setting up training batch configurations, and implementing measures to prevent data overfitting.

To prepare for model training, the text dataset is preprocessed, which includes using BERT's word splitter to separate words and packaging the processed data into a PyTorch dataset. A loss function, entitled CombinedLoss, is employed. It blends a weighted cross-entropy loss with a standard cross-

DOI: 10.54254/2755-2721/32/20230218

entropy loss and is specifically designed to manage category imbalance by assigning higher weights to a few categories. The original dataset is split in an 80/20 ratio to form a training set and a validation set. This split facilitates the evaluation of the model's generalization ability during the training process and helps prevent overfitting. Training parameters are defined, including the learning rate, weight decay, etc., followed by the use of a Trainer for model training. To further guard against overfitting, EarlyStoppingCallback is employed, which sets a condition to cease training early based on the accuracy of the validation set. The model is ultimately evaluated based on the test set, and test data prediction is performed to obtain predictive labels for the model. The trained model is saved for future use. The experimental setup can be observed in Figure 2.

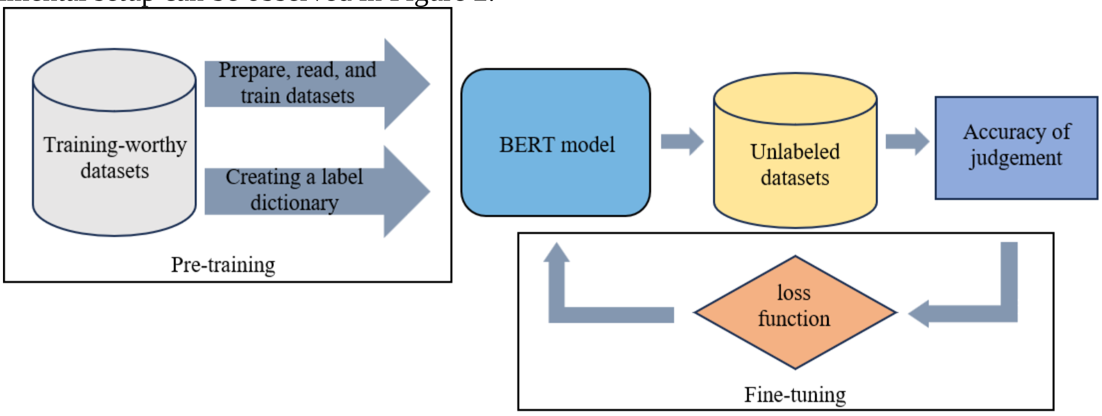

Figure 2. Experimental setup (Photo/Picture credit: Original).

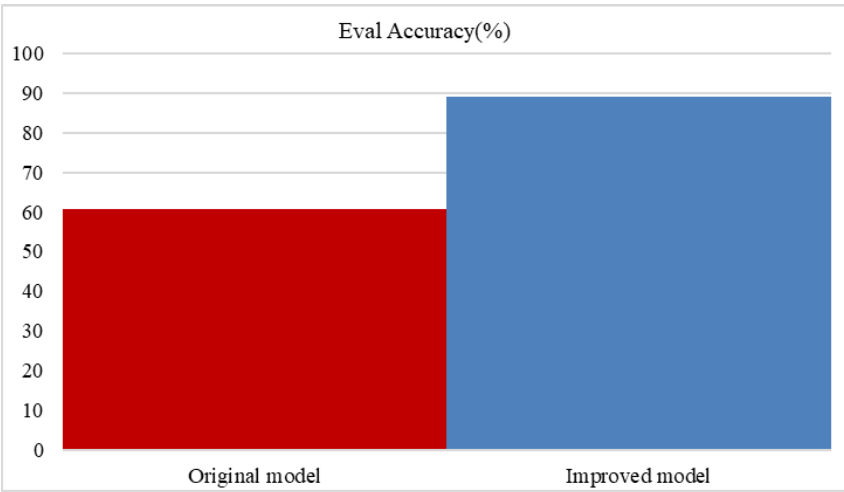

Figure 3. The comparison of accuracy between the improved model and the original model (Photo/Picture credit: Original).

4.2. Results

Ultimately, the model demonstrates satisfactory performance on the test set. Specifically, it meets the benchmarks set for accuracy and F1 scores, thereby validating the approach's effectiveness in the sentiment classification task. A confusion matrix analysis of the model's predictions further elucidates its performance across different sentiment categories. While the model provides reasonable accuracy for most sentiment categories, some misclassifications in certain categories do occur. The approach taken in this research incorporates sentiment analysis based on the BERT model, with the introduction of a novel hybrid loss function, CombinedLoss. This function merges weighted cross entropy with standard cross entropy loss, and includes an L2 regularization term to thwart overfitting. Differing weights in the model training help balance the data imbalance between various sentiment categories, thereby enhancing the model's performance across all categories. Post training and testing, the model exhibits commendable results on the sentiment analysis task.

This study reveals that using a hybrid loss function with weights and L2 regularization can effectively fine-tune the BERT model to improve sentiment analysis performance. In situations of category imbalance, this method notably enhances the model’s ability to recognize minority classes. The following segment presents the experimental results. Figure 3 illustrates a comparison of accuracy between the enhanced model and the original model. Figure 4 showcases the alterations in precision, loss function, and F1 value during the model's improvement process. As shown in Table 2.

Table 2. Performance Metrics Across Iterations for the Enhanced BERT Model in Sentiment Analysis.

iteration	1	2	3	4	5	6	7
Accuracy	41%	43%	53%	59%	61%	79%	89%
Loss Decrease	1.66	1.49	1.35	1.30	1.17	1.07	0.57
F1 Score	0.15	0.18	0.24	0.42	0.60	0.77	0.82

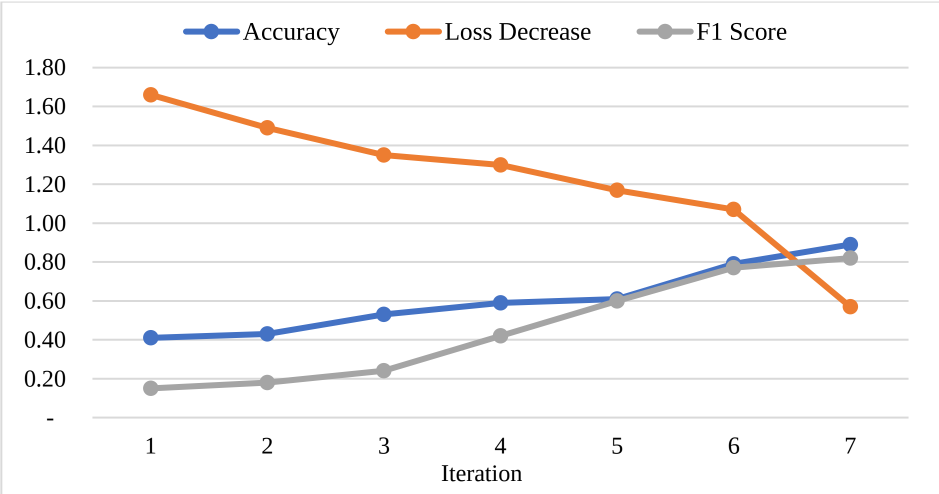

Figure 4. The changes in precision, loss function, and f1 value during the improvement of the model (Photo/Picture credit: Original).

5. Conclusion

This research reveals that the application of a weighted combined loss function and L2 regularization significantly enhances the performance of the BERT model in sentiment analysis tasks. Remarkably, this methodology shows a marked improvement in the recognition of minority classes in instances of class imbalance.

These findings hold substantial practical value for sentiment analysis and natural language processing, offering an effective solution to the common class imbalance issues frequently found in real-world datasets. Moreover, they illustrate how the integration of multiple loss functions and regularization methods can help avoid overfitting and bolster the model's generalization capability. From a theoretical standpoint, this research introduces an innovative model training strategy that optimizes the model through the amalgamation of various loss functions and regularization methods. This novel approach presents a fresh viewpoint on how to address class imbalance problems in natural language processing tasks, and provides valuable insights for the training of deep learning models.

While this study concentrates on text data, future research could extend sentiment analysis to multimodal data, including images, audio, and videos. By integrating sentiment information from diverse data modalities, a more holistic understanding and analysis of emotional expressions can be achieved, thereby broadening the application spectrum of sentiment analysis. In conclusion, this research underscores the practical value of utilizing weighted combined loss functions and L2

DOI: 10.54254/2755-2721/32/20230218

regularization to improve sentiment analysis performance with the BERT model. It also suggests potential future research directions, such as the extension of sentiment analysis to multimodal data, and provides novel insights for tackling class imbalance issues in natural language processing tasks.

Authors contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 4171-4186.

[2] Cao, Y., Sun, Z., Li, L., & Mo, W. (2022). A Study of Sentiment Analysis Algorithms for Agricultural Product Reviews Based on Improved BERT Model. Symmetry, 14(1), 1604.

[3] Birjali, M., Beni-Hssane, A., Erritali, M. (2017). Machine Learning and Semantic Sentiment Analysis based Algorithms for Suicide Sentiment Prediction in Social Networks. Procedia Computer Science, 113, 65-72.

[4] Turney, P.D. (2002). Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), 417-424.

[5] Peñalver-Martinez, I., Garcia-Sanchez, F., Valencia-Garcia, R., Rodríguez-García, M.Á., Moreno, V., Fraga, A., & Sánchez-Cervantes, J.L. (2014). Feature-based opinion mining through ontologies. Expert Systems with Applications, 41, 5995-6008.

[6] Yang, Z., et al. (2016). Hierarchical attention networks for document classification. Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies.

[7] Sun C, Huang L, Qiu X. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence[J]. arXiv preprint arXiv:1903.09588, 2019.

[8] Ray B, Garain A, Sarkar R. An ensemble-based hotel recommender system using sentiment analysis and aspect categorization of hotel reviews[J]. Applied Soft Computing, 2021, 98: 106935.

[9] AlQahtani A S M. Product sentiment analysis for amazon reviews[J]. International Journal of Computer Science & Information Technology (IJCSIT) Vol, 2021, 13.

[10] Sudhir P, Suresh V D. Comparative study of various approaches, applications and classifiers for sentiment analysis[J]. Global Transitions Proceedings, 2021, 2(2): 205-211.