DOI: 10.2478/ijasitels-2023-0007

Sentiment Analysis Using Bert Model

Dorca Manuel-Ilie[1], Pitic Antoniu Gabriel[1], Crețulescu Radu George[1]
[1]Computer Science and Electrical and Electronics Engineering Department, Faculty of Engineering, "Lucian Blaga" University of Sibiu, Romania
{manuel.dorca, antoniu.pitic, radu.cretulescu} @ulbsibiu.ro

Abstract

The topic of this presentation entails a comprehensive investigation of our sentiment analysis algorithm. The document provides a thorough examination of its theoretical underpinnings, meticulous assessment criteria, consequential findings, and an enlightening comparative analysis. Our system makes a substantial contribution to the field of sentiment analysis by using advanced techniques based on deep learning and state-of-the-art architectures.

Keywords: Sentiment analysis, Sentiment polarity, Emotion prediction, BERT model

1. Introduction

Sentiment analysis, an essential aspect of natural language processing, is becoming increasingly important in understanding human emotions expressed through text data. The increasing amounts of social media content, consumer feedback, and various textual datasets highlight the need for advanced sentiment analysis algorithms. We address this need by presenting a methodology that utilizes advancements in machine learning and deep learning.
Our program aims to analyze feelings with a different approach, beyond traditional methods. The increasing abundance of text data on social platforms and other digital spaces necessitates accurate sentiment interpretation, which is crucial for various applications such as market research and social media monitoring.

2. Main Work

2.1. Theoretical Fundamentals of the Proposed Method for Sentiment Analysis

Our sentiment analysis method was developed by creating a theoretical framework that combines ideas from deep learning with advanced technology approaches. Our process goes beyond just combining existing components; instead, it involves integration of several methods, taking inspiration from the most recent breakthroughs in deep learning technology. Our approach to sentiment analysis includes new elements such as dynamic lexicon development and creative management of subjects and emotions. The

DOI: 10.2478/ijasitels-2023-0007

theoretical foundation is to provide a versatile understanding of emotions in written content, surpassing the constraints of conventional methods.

The effectiveness of our solution is based on the use of a versatile dictionary expansion mechanism, which allows the algorithm to adjust its understanding of emotional subtleties. When combined with the neural network structure explained in the paper, this collaboration creates a strong basis for accurately identifying emotions. Our program goes beyond simply copying current methods; it incorporates insights obtained from the complexities described in the study, thus advancing sentiment analysis into a more practical realm. This approach represents a significant shift in the way sentiment analysis is conducted in the current field of machine learning, utilizing advanced techniques to effectively handle the changing requirements.

2.2. Analysis of the Kaggle Sentiment140 Dataset

2.2.1. Summary of the Dataset

The Kaggle Sentiment140 dataset [7] is a highly utilized resource in sentiment analysis, comprising of tweets sourced from Twitter, with each tweet being categorized as either positive or negative. The binary sentiment classification sets a definitive goal for the development and evaluation of sentiment analysis methods.

2.2.2. Dataset Source and Preprocessing

The dataset was acquired using Twitter's public API, assuring diversity across time periods and user demographics. It was curated by Stanford University. Tweets were recovered by identifying specific emoticons that indicate either good or negative attitudes. The preprocessing stage encompassed the elimination of irrelevant information, the segmentation of the text into individual tokens for further analysis, and the organization of the dataset to facilitate sentiment analysis.

2.2.3. Analysis of Sentiment Distribution and Text Length

Examining the distribution of sentiment allows for a deeper understanding of the dataset's accuracy in representing various viewpoints and the potential difficulties in classifying them. Furthermore, analyzing the lengths of tweets helps to discover limitations that could affect the effectiveness of a model, enabling researchers to adjust their approaches accordingly.

2.3. Algorithm Details

Our sentiment analysis method consists of meticulously developed steps to capture and analyze sentiment information effectively and precisely. The following sections outlines the key steps of our approach.

DOI: 10.2478/ijasitels-2023-0007

2.3.1. Topic Annotation

Our algorithm begins by performing topic annotation, which is a method aimed at identifying and isolating the main subjects or topics present in the input text. By utilizing natural language processing techniques, particularly part-of-speech tagging with the NLTK toolkit [1], our implementation guarantees a focused examination of emotional tone. We exemplify the concept by providing code snippets that showcase the process of tokenizing sentences, assigning parts of speech to words, and extracting the key themes. To address phrases that encompass many themes, we utilize the TF-IDF method, which offers a strong and effective method for identifying the most pertinent subject matter.

```
Sentence: @stellargirl I loooooooovvvvvveee my Kindle2. Not that the DX is cool
________________________________________________________________________
| | | | | | | Topic: Kindle2
________________________________________________________________________

Sentence: Reading my kindle2... Love it... Lee childs is good read.
________________________________________________________________________
| | | | | | | Topic: kindle2, Lee childs
________________________________________________________________________

Sentence: "Ok, first assessment of the #kindle2 ...it fucking rocks!!!"

________________________________________________________________________
| | | | | | | Topic:kindle2
________________________________________________________________________
```

Figure 1 Output of topic anotation step

2.3.2. Emotion Annotation with Dynamic Lexicon Expansion

The second stage centers on emotion annotation, augmenting the comprehension of natural language and sentiment analysis. We apply a predetermined lexicon to assign emotional scores to words, utilizing the Levenshtein distance [8] technique for word comparison. The algorithm progressively enlarges the lexicon as time goes on, consistently enhancing the process of emotion annotation. We explore the intricacies of lexicon-based emotion annotation, Levenshtein distance word matching, sentiment score calculation, and sentiment mapping. The method of dynamic lexicon expansion guarantees the ability to adapt to changing language patterns, which is crucial for appropriately assigning emotion scores to words that have not been encountered before. After the algorithm has determined the word that closely corresponds, it next proceeds to allocate the emotion score linked to that word. This approach ensures that even if a term is not found in the lexicon, it can still be assigned an emotional value by evaluating its similarity to existing words. [1]

2.3.3. Emotion Categorization

Utilizing the sentiment scores calculated in the preceding step, the third phase categorizes each sentence into positive, negative, or neutral sentiment groups according to the dataset. As our objective is to predict sentiment, the general emotion of the Kaggle dataset [7] has been standardized to a scale of -1 for negative, 0 for neutral, and 1 for positive. The algorithm discerns the most pertinent emotions associated with

DOI: 10.2478/ijasitels-2023-0007

negative, neutral, and positive phrases, so offering a nuanced comprehension of the emotional context. We analyze the approach used for classifying sentiments, identifying emotions, and managing discrepancies between given categories and sentiment scores. This phase guarantees a thorough sentiment analysis, capturing the entire emotional tone of every line.

The algorithm begins by annotating the topic and then proceeds to detect and categorize the main subjects present in the textual data. This establishes a fundamental comprehension of the surrounding circumstances, setting the basis for further examination of emotions.

By annotating the subjects and emotions, the computer can ascertain the overall sentiment of the text. This stage enables a thorough sentiment analysis by combining specified subjects and emotions. The program considers the interaction between themes and emotions, guaranteeing a sophisticated evaluation.
Our method uses a sentiment mapping dictionary to establish a connection between sentiment scores and distinct emotion categories. The purpose of this dictionary is to provide clear definitions of different types of emotions and the specific ranges of scores associated with each category. For example, the sentiment category labeled as "Positive" encompasses feelings such as "*Happy*," "*Excited*," and "*Amazed*," each of which is linked to specific score ranges.

The system categorizes the emotion of a given sentence by comparing its sentiment score to predetermined ranges. When the score is within a defined range, the statement is associated with the corresponding emotion category.

2.4. Neural Network Configuration

During the fourth stage of our method, we concentrate on customizing the neural network by utilizing the BERT architecture specifically designed for sentiment analysis. The method requires evaluating several hyperparameter configurations, such as training epochs, learning rate, and batch size, in order to improve model performance. The BERT model, which is built on transformers, effectively collects contextual information in both directions and has demonstrated good performance in tasks related to natural language processing.[3][4]

The training epochs significantly impact the model's exposure to the dataset, striking a balance between acquiring knowledge and achieving generalization. Similarly, the learning rate, which determines how the model's parameters are adjusted during optimization, affects how quickly the model converges and the likelihood of overfitting. The batch size, which refers to the number of training instances processed at once, has an effect on both memory usage and computational efficiency.

Hyperparameter tuning entails a methodical examination utilizing techniques such as grid search or random search to determine the ideal configuration. The utilization of regularization strategies, which aim to mitigate overfitting, is explored. This includes the incorporation of dropout rates and weight decay. Model evaluation is based on parameters like as accuracy, precision, recall, F1 score, and AUC-ROC[5]. In order to

DOI: 10.2478/ijasitels-2023-0007

guarantee impartial assessment, the dataset is partitioned into training, validation, and test sets.

Configuring the neural network is crucial for utilizing BERT's capabilities to provide successful sentiment analysis. This involves capturing contextual nuances and accurately classifying sentiment in textual input. The method entails a thorough and systematic approach of conducting experiments, validating results, and fine-tuning in order to achieve an optimal configuration that improves the performance of the model.

Data preparation and normalization play a crucial role in preparing data for sentiment analysis models, and their importance should not be underestimated. During this pivotal phase, the algorithm converts annotated data into tensors that are compatible with the BERT neural network design. The procedure entails tokenization via the DistilBertTokenizer, encoding with padding and truncation to ensure consistent sequence length, and converting sentiment labels and subcategories into numerical representations.

Subsentiment analysis offers an extra training feature that enables the replacement of labels with appropriate subcategories, enhancing the evaluation of sentiment with more nuanced precision. The processed data is structured as dictionaries nested within a list. Each dictionary contains 'input_ids' and 'attention_mask' tensors, which represent tokenized and encoded texts. Additionally, the dictionaries include 'labels' for sentiment labels and 'topics' for supplementary topic information. The utilization of this organized structure guarantees the BERT model's ability to process and comprehend the data during both training and evaluation.

In order to enhance efficiency, the code utilizes caching mechanisms by using the *functools* library, specifically the *lru_cache* decorator. This improves the computing efficiency for repetitive tasks, particularly when dealing with large datasets.

The technique systematically investigates various configurations for the BERT neural network, simulating a wide range of parameters such as the number of epochs, learning rates, and batch sizes. The number of epochs ranges from 4 to 16, the learning rates consist of values such as 0.0002, 0.00002, 0.000002, and 0.0000002, and the batch sizes span from 8 to 256 in powers of 2. The purpose of this systematic experimentation is to determine the most effective hyperparameter settings, obtaining a good combination of training efficiency and model performance.

3. Results

3.1. Evaluation Metrics

For evaluating our sentiment analysis algorithm we used as metrics the accuracy, precision, recall, and F1 score. These metrics served as crucial indicators of the algorithm's performance in sentiment classification across diverse datasets and domains.

DOI: 10.2478/ijasitels-2023-0007

3.2. Performance Analysis

The algorithm demonstrated robustness in handling challenges such as sarcasm and sentiment shifts, showcasing its adaptability to complex sentiment analysis tasks. This performance analysis provided valuable insights into the algorithm's overall efficiency.

3.3. Comparative Analysis

Comparing our algorithm to prior research revealed accuracies ranging from 58% to 82%. While our accuracy aligns with [3], the inclusion of additional metrics such as precision and recall offered a more comprehensive evaluation.

The sentiment polarity prediction results (Table 1) show an overall mean accuracy of 67.4%, with a standard deviation of 7.4%. This indicates a moderate level of consistency in sentiment prediction. The F1 score, which balances precision and recall, is consistently high across epochs, suggesting a well-rounded performance in sentiment classification.

Table 1 Sentiment Polarity Prediction Results

	Accuracy	Precision	Recall	F1 Score
Epoch count	16	16	16	16
mean	0.67389	0.81571	0.82534	0.81833
std	0.07396	0.07346	0.05935	0.0505
min	0.57137	0.7088	0.7434	0.75164
25%	0.60813	0.75752	0.7752	0.7691
50%	0.66926	0.81826	0.81433	0.81834
75%	0.73423	0.87628	0.87819	0.86176
max	0.80001	0.91967	0.93312	0.8956

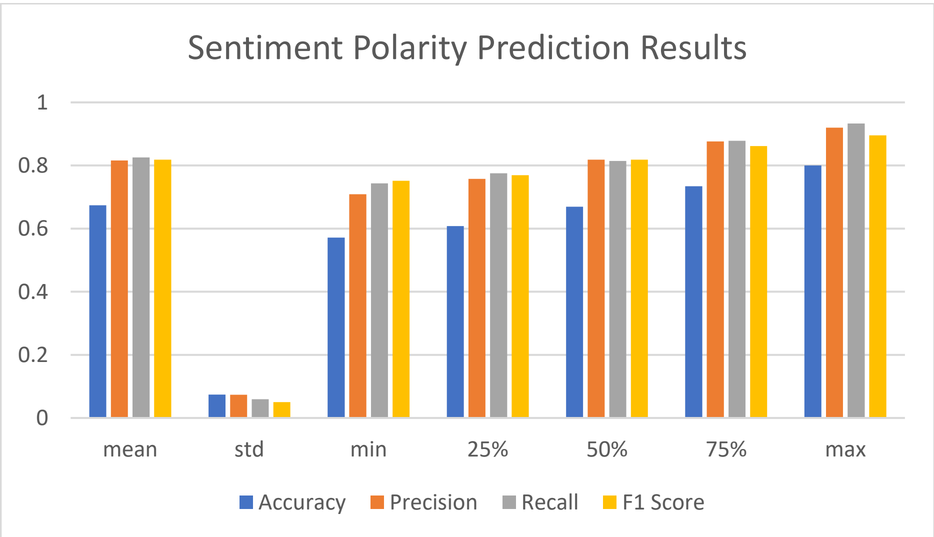

Figure 2 Sentiment Polarity Prediction Results

DOI: 10.2478/ijasitels-2023-0007

In the base emotion prediction results (Table 2), the mean accuracy of 70.2% with a standard deviation of 6.6% indicates a stable performance. The F1 score, emphasizing a balance between precision and recall, maintains a consistently high level, further reinforcing the algorithm's proficiency in emotion prediction.

Table 2 Base Emotion Prediction Results

	Accuracy	Precision	Recall	F1 Score
Epoch count	128	128	128	128
mean	0.70186	0.82147	0.84365	0.83037
std	0.06584	0.05873	0.0552	0.03977
min	0.57333	0.68264	0.72509	0.71752
25%	0.64774	0.77471	0.79864	0.80316
50%	0.70532	0.82421	0.8469	0.83147
75%	0.75617	0.87068	0.88863	0.85867
max	0.81999	0.9199	0.93967	0.92575

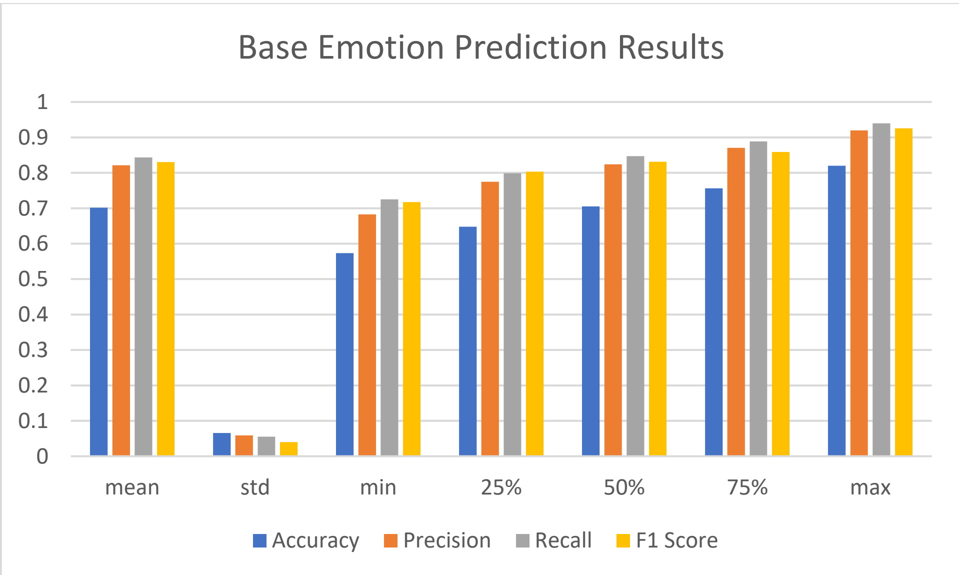

Figure 3 Base Emotion Prediction Results

3.4. Discussion of Results

Our sentiment analysis algorithm, incorporating topic and emotion annotation with a pre-trained BERT model [6], demonstrated a improved sentiment classification. The results showcase its effectiveness, particularly in capturing nuanced emotional

DOI: 10.2478/ijasitels-2023-0007

subtleties within the text. Areas for improvement were identified, such as refining preprocessing techniques and exploring advanced feature extraction methods.

3.5. Future Research Directions

Future research could prioritize refining preprocessing techniques, exploring sophisticated feature extraction methods, and considering ensemble learning approaches to further enhance performance. Addressing challenges specific to certain domains or languages will be crucial for the algorithm's continued adaptability and efficacy.

3.6 Conclusion

In conclusion, the results and discussion chapter provide a concise yet comprehensive analysis of our sentiment analysis algorithm's outcomes. The findings contribute to the broader understanding of sentiment analysis capabilities and guide future research in this dynamic field.

4. 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 (pp. 4171-4186).

[2] LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep learning. Nature*, 521(7553), 436-444.

[3] Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

[4] Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). *Recursive deep models for semantic compositionality over a sentiment treebank*. In Proceedings of the conference on empirical methods in natural language processing (EMNLP) (Vol. 1631, p. 1642).

[5] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. (2017). *Attention is all you need*. Advances in neural information processing systems, 30.

[6] https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270 (last visit december 2023)

[7] https://www.kaggle.com/datasets/rajugc/kaggle-dataset (last visit december 2023)

[8] Kaufmann, M., Principles of Data Integration, ISBN 9780124160446, 2012