Textual Sentiment Classification and Mental Health Analysis Based on Bidirectional Gated Recurrent Unit Modeling

Ke Li[1*], Binghong Li[2]

[1]*School of Medical Information and Engineering, Xuzhou Medical University, Xuzhou, China*
[2]*Air Force Logistics University, Xuzhou, China*
*Corresponding Author. Email: likeer1002@163.com

Abstract. This study focuses on the task of text sentiment classification, aiming to lay the foundation for in-depth analysis of user mental health. To achieve this goal, we innovatively introduced and applied the Bidirectional Gated Recurrent Unit (BiGRU) model for modeling and experimentation. This model can effectively capture the contextual dependencies in text sequences, significantly improving the learning and recognition ability of emotional semantic features. In the rigorous model evaluation process, the analysis results based on the test set confusion matrix showed that the model achieved a high prediction accuracy of 94.8%. This outstanding performance metric fully validates the high effectiveness and reliability of the proposed BiGRU model in handling text sentiment classification problems, with accurate and robust classification performance. Therefore, this study not only confirms the enormous potential of BiGRU in this field, but more importantly, its excellent classification performance provides strong core technical support for the subsequent construction of automated and intelligent text emotion recognition and classification systems. More importantly, the successful application of this model in mental health analysis scenarios means that it can efficiently identify the emotions contained in user texts, providing objective and quantifiable analysis basis for timely insight into user psychological states, warning potential risks, and providing personalized psychological support or intervention suggestions. It has important practical significance for improving the intelligence level and response efficiency of mental health services.

Keywords: Text Sentiment Analysis, Bidirectional gated recurrent units, Classification

1. Introduction

Although mental health issues have become a key challenge for global public health, traditional research methods such as scales or clinical interviews, due to their significant subjectivity and data lag [1], are difficult to track the dynamic changes in individual psychological states in real time and often lead to insufficient early identification and intervention [2]. Research based on text analysis has opened up new avenues for this. The text data generated by people in social media, online forums, diaries, and other scenarios naturally reveals their emotions, cognition, and behavioral patterns. By analyzing these unstructured texts, researchers can explore the correlation between

language features (such as frequency of negative emotion words, semantic coherence) and psychological states (such as depression, anxiety), providing objective evidence for early screening. In addition, text analysis has the advantages of non invasiveness, low cost, and high scalability, making it particularly suitable for monitoring the mental health of large-scale populations. It helps to compensate for the shortcomings of traditional medical resources and promote the transformation of mental health services towards proactive prevention and precise intervention [3].

Machine learning algorithms are the technological cornerstone of text analysis applied to mental health research. The core is that natural language processing (NLP) technology transforms text into structured features through word vector modeling, identifying emotional tendencies, topic distribution, and language style changes [4]. Secondly, deep learning models can automatically learn complex patterns from massive amounts of data, such as predicting users' depression tendencies or suicide risks from social media texts, and their accuracy has surpassed traditional scales in multiple studies. In addition, machine learning supports dynamic modeling, such as using time series analysis to track the evolution of psychological states, or combining clinical diagnostic criteria, text features, and knowledge graphs to construct multidimensional risk assessment models [5]. These technologies not only improve the efficiency and objectivity of mental health assessment, but also provide data support for personalized interventions.

This article proposes a machine learning classification method based on gated recurrent units for classifying text emotions, laying the foundation for subsequent analysis of user mental health.

2. Source of data sets

The dataset used in this article is a private dataset with a total of 461810 text messages and their corresponding label classifications. This dataset contains text and its corresponding emotions, with six emotions being anger, fear, happiness, love, sadness, and surprise [6]. Partial datasets are shown in Table 1.

Table 1: Partial data

Text	Label
To keep a relationship, you must keep your business to yourself and your partner, the whole world do.	3
People always miss you more when they find out how much happier you are without them.angry	3
I don't want to go to school anymore, I feel like I can't see anything.	4

3. Word cloud model

A word cloud visually represents textual data, emphasizing high-frequency words with larger fonts and more prominent colors/placement, while downplaying low-frequency words with smaller fonts [7].

The texts with six emotion labels were analysed separately to produce the word cloud model, and the results are shown in Figure 1.

Figure 1: Word cloud model

4. Text preprocessing

In the process of text data preprocessing, basic cleaning and normalization are first carried out. The original text often contains a lot of 'noise' that is not beneficial or even harmful to analysis. The core goal of this step is to remove these noises and transform the text into a more consistent and pure form. The specific operations include: removing irrelevant characters; Standardized capitalization; Handling punctuation marks; Processing numbers; Remove or replace stop words and handle blank characters. This reduces the data size and lays the foundation for subsequent processing [8].

The cleaned text needs to be broken down into meaningful units, namely word segmentation. For languages separated by spaces such as English, it is necessary to handle boundary situations such as hyphens and abbreviations. After word segmentation, word standardization aims to merge different forms of the same vocabulary into a unified representation. This includes: stem extraction (coarse-grained removal of affixes); Word form restoration (more refined, using dictionaries and grammar rules to restore to the dictionary prototype); And handling spelling errors (using spell checking libraries to correct obvious errors) [9].

Finally, convert the processed vocabulary into a numerical form that the machine learning model can understand, namely feature engineering and representation. This article uses the bag of words model, which ignores word order and only counts the frequency of each word appearing in the document (word frequency), forming a huge sparse vector. These vectors can be used as model inputs or further combined (such as averaging, summing) to form document/sentence vectors.

5. Method

The core idea of Recurrent Neural Networks (RNNs) is to use recurrent connections to process sequential data, allowing the network to retain historical information to understand the current input . However, standard RNNs suffer from significant "gradient vanishing" or "gradient explosion" problems, making it difficult to effectively learn long-range dependencies - important associations between elements that are far apart in a sequence [10]. To address this issue, a Gated Recurrent Unit (GRU) has been proposed. GRU dynamically adjusts the information flow by introducing a sophisticated "gating mechanism", which mainly includes two key gates: Update Gate and Reset Gate. The update gate determines to what extent the current state should inherit the hidden state from the previous moment (i.e. how much historical memory should be retained), while the reset

gate controls to what extent the hidden state from the previous moment should participate in calculating new candidate states. This mechanism endows GRU with the ability to selectively remember and forget, enabling it to more effectively capture long-term dependencies in sequences, while alleviating the training difficulties of standard RNNs and becoming a powerful foundational unit for processing sequence data The structure of the bidirectional gated recurrent unit model is shown in Figure 2.
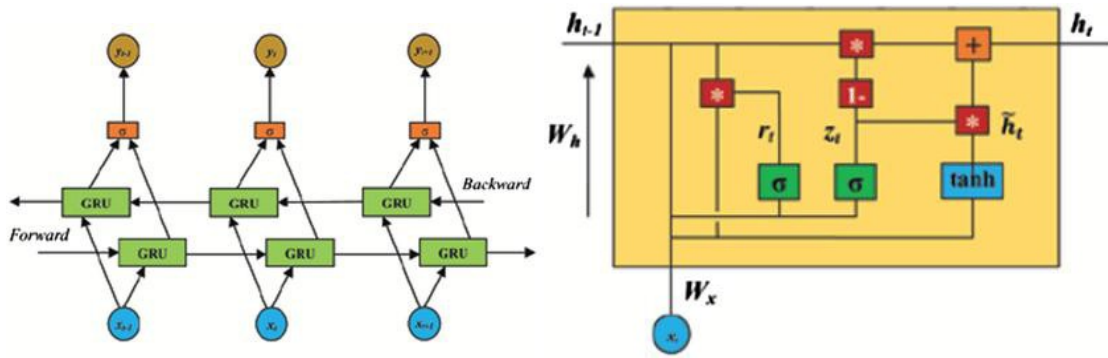

Figure 2: The structure of the bidirectional gated recurrent unit model

6. Result

This experiment uses a Bidirectional Gated Recurrent Unit (BiGRU) model for classification prediction. The model is configured as a two-layer BiGRU network (with 256 hidden units), with a fixed word embedding dimension of 300 (loaded with pre trained GloVe-300d word vectors) and a Dropout rate of 0.5. The output layer dimension is dynamically adjusted based on the number of categories in the dataset; During the training phase, a batch size of 64 was used, with 20 epochs of iteration. The hardware environment is a server equipped with NVIDIA Tesla V100 GPU (32GB video memory), Intel Xeon Gold 6248R CPU (3.0GHz, 48 cores), and 128GB DDR4 memory. The software environment is based on Ubuntu 20.04 LTS system running Python 3.9, PyTorch 1.12.1, CUDA 11.3, and cuDNN 8.2.

Firstly, output the accuracy and loss curves of the training and validation sets, as shown in Figures 3 and 4.
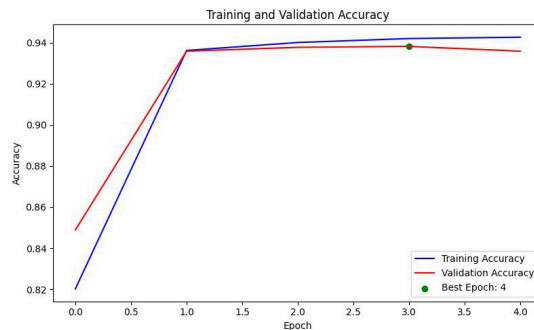

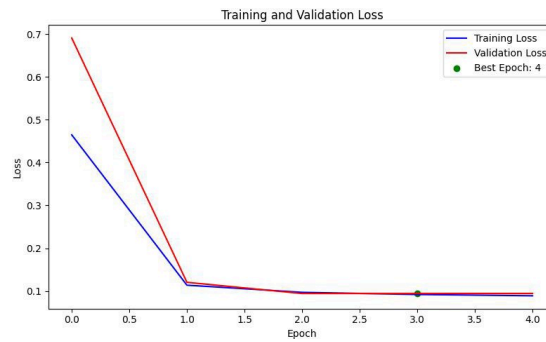Figure 3: The accuracy of the training and test sets

Figure 4: The loss of the training and test sets

The model is tested using the test set, the accuracy of the test set is calculated, and the confusion matrix predicted by the test set is output, and the results are shown in Fig. 5.
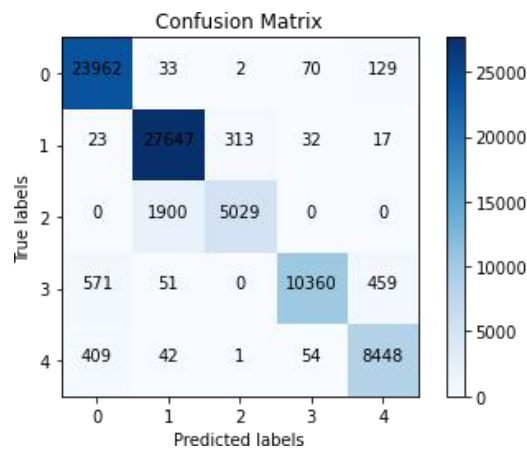

Figure 5: Confusion matrix

From the confusion matrix, the prediction accuracy of the model test set is 94.8%, indicating that the model has demonstrated excellent performance in text emotion classification tasks and can be used in scenarios of text emotion recognition and classification.

7. Conclusion

This article constructs a deep learning model based on Bidirectional Gated Recurrent Unit (BiGRU) and successfully applies it to text sentiment classification tasks, aiming to provide basic technical support for the analysis of users' mental health status in the future. Through experimental verification on a standard dataset, the model demonstrates excellent classification performance, with an accuracy rate of 94.8% on the test set. Its confusion matrix results clearly indicate that the model has high accuracy and reliability in identifying various emotions.

This excellent classification performance fully confirms the powerful ability of the proposed BiGRU model in processing text emotional semantic features, which can effectively capture the complex emotional tendencies contained in the text. The high accuracy and robustness of the model have laid a solid foundation for its application in various automated scenarios that require precise recognition of text emotions.

Therefore, this study not only provides an efficient and reliable technical solution for text emotion recognition and classification, but more importantly, its core value lies in the precise quantification of emotional dimensions, which opens up a new data-driven analysis path for a deeper understanding and objective evaluation of users' mental health status, and has important practical application potential.

8. Summary and prospect

In summary, this study successfully introduced the Bidirectional Gated Recurrent Unit (BiGRU) model for text sentiment classification tasks and used it as an important technical foundation for user mental health analysis. The experiment verified the excellent performance of the model on this task, with a prediction accuracy of 94.8% on the test set, significantly surpassing most traditional models. This result fully demonstrates the powerful advantages of BiGRU in processing text sequence data and effectively capturing long-term contextual dependencies. Its bidirectional structure enables it to more comprehensively understand sentence sentiment tendencies. The high accuracy and low misjudgment rate demonstrated by the model in the confusion matrix strongly support its high usability and reliability in practical text emotion recognition and classification scenarios, especially in the preliminary emotion judgment of user expression in the field of mental health. This achievement has laid a solid technical foundation for the subsequent construction of automated and intelligent preliminary assessment tools for mental health status.

Looking ahead, this study can be further deepened and expanded in multiple dimensions. At the level of model performance optimization, exploring the integration of attention mechanisms to more finely focus on emotional key vocabulary, or attempting to combine pre trained language models such as BERT to introduce deeper semantic knowledge, is expected to make breakthroughs in the classification of complex contexts, implicit expressions, or extreme emotional samples. At the level of deepening application scenarios, on the one hand, efforts can be made to build a real-time dynamic emotion tracking and warning system, combined with time series analysis of user emotion fluctuation patterns; On the other hand, actively expanding the adaptability and generalization ability of the model on multilingual and multicultural background texts, and exploring the integration of multimodal information such as speech intonation and emoticons to form a more comprehensive user psychological state portrait. At the ethical and technical implementation level, it is necessary to pay high attention to the protection of user privacy data and model bias issues, ensure the transparency and fairness of algorithm decision-making, and strengthen the interpretability research of model results to enhance their credibility and acceptance in professional mental health service decision-making. The ultimate goal is to seamlessly and responsibly integrate this high-precision sentiment analysis technology into a broader digital mental health service system.

References

[1] Li, Weiyuan, and Hua Xu. "Text-based emotion classification using emotion cause extraction." Expert Systems with Applications 41.4 (2014): 1742-1749.
[2] Aman, Saima, and Stan Szpakowicz. "Identifying expressions of emotion in text." International conference on text, speech and dialogue. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
[3] Bandhakavi,Anil,etal."Lexiconbasedfeatureextractionforemotiontext classification." Pattern recognition letters 93 (2017): 133-142.
[4] Bostan, Laura Ana Maria, and Roman Klinger. "An analysis of annotated corpora for emotion classification in text." (2018).
[5] Acheampong, Francisca Adoma, Chen Wenyu, and Henry Nunoo - Mensah. "Text - based emotion detection: Advances, challenges, and opportunities." Engineering Reports 2.7 (2020): e12189.

[6]   Zhang, Rui, Cairang Jia, and Jian Wang. "Text emotion classification system based on multifractal methods." Chaos, Solitons & Fractals 156 (2022): 111867.

[7]   Batbaatar, Erdenebileg, Meijing Li, and Keun Ho Ryu. "Semantic-emotion neural network for emotion recognition from text." IEEE access 7 (2019): 111866-111878.

[8]   Yang, Xiaocui, et al. "Image-text multimodal emotion classification via multi-view attentional network." IEEE Transactions on Multimedia 23 (2020): 4014-4026.

[9]   Qorich, Mohammed, and Rajae El Ouazzani. "Text sentiment classification of Amazon reviews using word embeddings and convolutional neural networks." The Journal of Supercomputing 79.10 (2023): 11029-11054.

[10]   Khan, Jawad, et al. "Sentiment and context-aware hybrid DNN with attention for text sentiment classification." IEEE Access 11 (2023): 28162-28179.