Early Warning of Negative Emotions on social media with Deep Learning: A Twitter Sentiment Analysis Study

***Abstract*—This study compares three deep learning-based sentiment analysis methods - RNN, LSTM, and BERT - to determine the best model for early detection of negative emotions on social media. This article uses Twitter data, preprocesses the data through natural language processing (NLP) technology, applies three deep learning models to sentiment analysis, and uses accuracy and F1 score as evaluation indicators. The results show that BERT outperforms RNN and LSTM, achieving an accuracy and F1 score of more than 89%. This result highlights the effectiveness of BERT in sentiment analysis and provides important support for enhancing real-time monitoring of online platforms, timely response to negative emotions, and sentiment management and early warning systems.**

Keywords: Sentiment analysis; negative emotions; Recurrent Neural Network (RNN); Long Short-Term Memory (LSTM); Bidirectional Encoder Representations from Transformers (BERT); Natural Language Processing (NLP).

I. Introduction

With the widespread popularity of social media, the emotions expressed by users on the platform have a profound impact on the real world, especially the spread of negative emotions may bring a series of social risks, such as youth suicide, violent incidents, hate crimes, and even recruitment of terrorist organizations. Therefore, the application of sentiment analysis technology to detect negative emotions in social media data is of great practical significance. This technology can not only help enterprises grasp the brand image in time, find and respond to negative information, but also provide data support for sociological research on the public's emotional tendency to social issues, which is conducive to the scientific nature of policy making. In addition, sentiment analysis can identify bad information in advance in monitoring online public opinion and maintain social harmony and stability. In the field of mental health, it provides strong support for early warning of suicide risk and lays the foundation for psychological intervention.

In the development of sentiment analysis, early methods mainly relied on rule-based models and traditional machine learning algorithms. Rule-based approaches used predefined sentiment lexicons to determine text polarity; these methods are simple and straightforward but lack effectiveness in capturing complex contexts (Taboada et al., 2011). Traditional machine learning methods like Naïve Bayes and Support Vector Machines (SVM) classify sentiment by utilizing labeled datasets. Although these methods can deliver higher accuracy, they require the manual construction of features, which makes the training process more intricate and reliant on large amounts of data (Bengio et al., 2013). In recent years, deep learning models have shown remarkable progress in sentiment analysis. The RNN model, which relies on its recurrent structure, is effective in capturing context in sequential data. However, it struggles with the vanishing gradient problem when dealing with longer sequences (Bengio et al., 1994). To address this, LSTM introduced a gating mechanism that better maintains long-term dependencies, making it more effective than traditional RNNs for tasks involving long sequences (Hochreiter & Schmidhuber, 1997). BERT has taken sentiment analysis to the next level by improving both accuracy and contextual understanding. Built on the Transformer framework, BERT uses bidirectional encoding and a self-attention mechanism, allowing it to capture detailed semantic information, especially when pre-trained on large datasets.

This study conducts sentiment analysis on a Twitter sentiment dataset using natural language processing (NLP) techniques and deep learning models. The data preprocessing includes steps such as tokenization, part-of-speech tagging, syntactic parsing, and named entity recognition (NER) to organize the text. We then applied RNN, LSTM, and BERT models to categorize sentiment in the dataset, assessing model performance through cross-validation and confusion matrix analysis, focusing on accuracy, recall, and F1 score. Furthermore, we compared the effectiveness of each model in sentiment classification, particularly regarding their text preprocessing and feature extraction abilities.

Section 2 reviews the applications of sentiment analysis and its significance on social media, highlighting key studies on deep learning-based sentiment analysis models like RNN, LSTM, and BERT. Section 3 outlines the research methodology, covering NLP techniques for data preprocessing, model development, comparison of various models, and evaluation strategies. Section 4 presents an analysis of the experimental results and discusses the limitations of deep learning approaches in sentiment analysis..

II. Literature Review

Sentiment analysis plays an important role and is increasingly applied in various real-world scenarios. On social media, it has proven valuable in a wide range of fields. For example, Widodo and Riyanto (2019) analyzed public sentiment on Indonesia's Twitter platform regarding the anti-LGBT movement, finding that most comments were neutral. This highlights the importance of gaining a deeper understanding of public views on sensitive topics. Similarly, Desmet and Hoste (2017) used social media data to predict suicidal tendencies, providing new tools and methods to support suicide prevention efforts. In another study, Ismail et al. (2020) performed sentiment analysis on tweets from cancer survivors, successfully identifying individuals suffering from post-traumatic stress disorder (PTSD) and supporting mental health interventions. Additionally, Singh et al. (2018) applied sentiment analysis to predict disease outbreaks and epidemics, underscoring its vital role in public health monitoring and response.

In recent years, sentiment analysis has shifted from traditional machine learning models (such as Support Vector Machines (SVM) and Naïve Bayes) to more advanced deep learning models including RNN, LSTM, and BERT, significantly improving classification accuracy and automated feature extraction. For instance, RNN was used to classify sentiment in Twitter posts during the COVID-19 pandemic, where a refined analysis of sentiment intensity categories revealed that, despite the pandemic's negative impact, overall sentiment remained positive and even showed a gradual increase in positivity over time, achieving an accuracy of 80% (Nemes & Kiss, 2020). Similarly, LSTM was applied to cryptocurrency-related data from Sina Weibo in China, enabling the prediction of price fluctuations for digital currencies like Bitcoin, ETH, and XRP, with an accuracy of 87%, thus supporting investment decisions (Huang et al., 2021). BERT was also employed to analyze COVID-19-related sentiment on Twitter among users in India and globally, showing that Indian users tended toward more positive sentiment and exhibited less negative dissemination. This analysis, achieving a validation accuracy of 94%, provides valuable insights for evaluating the effectiveness of government pandemic measures (Singh et al., 2021).

In summary, the research shows how sentiment analysis can be applied to social media data across various fields, such as social issues, mental health, public health, and investment decisions. It also highlights the value of deep learning models like RNN, LSTM, and BERT, which significantly improve accuracy. Building on this, the next section will use the Twitter database to perform sentiment analysis with RNN, LSTM, and BERT, aiming to find the best models for detecting negative sentiment. These models can be useful for real-time social media monitoring, particularly in areas like public health crisis management and mental health support.

III. MATHOLOGY

3.1 Data Processing

3.1.1 Data Collection

This study utilizes a Twitter dataset sourced from public platforms like Kaggle. Alternatively, data can be dynamically collected using the Tweepy library via the Twitter API. The collected data is saved in JSON or CSV format for subsequent processing. For simplicity, this study employs a publicly available labeled dataset, containing tweets tagged as positive, negative, or neutral.

3.1.2 Data Preprocessing

The goal of data preprocessing is to convert raw data into a format understandable by the model while reducing noise. Common text preprocessing steps include tokenization, stemming, lemmatization, part-of-speech tagging, named entity recognition, and dependency parsing. In this study, Python's NLTK and spaCy libraries are used to implement these steps, including tokenization, stemming, lemmatization, stop word removal, named entity recognition (NER), and dependency parsing

3.1.3 Feature Extraction

The aim of feature extraction is to transform text into numerical features for improved sentiment prediction. Common feature extraction methods include Unigram and N-gram features, as well as Term Frequency-Inverse Document Frequency (TF-IDF). In this study, the TF-IDF method is used to measure the importance of terms within the text. By using the TfidfVectorizer function, the text is transformed into TF-IDF feature vectors.

3.2 Model Comparison

3.2.1 Dataset Splitting

The dataset is split with 80% used for training and 20% for testing. This standard split ensures that the training set has sufficient data for model training, while the test set is used to evaluate the model's generalizability.

3.2.2 Model Experiment

Three deep learning models were selected for this study: RNN, LSTM, and BERT. These models were implemented using Keras and TensorFlow and trained on the same dataset. Each model was trained for 15 epochs to allow sufficient time for learning while preventing overfitting.

3.2.3 Experiment Evaluation

To evaluate the performance of each model, two key metrics were chosen: accuracy and F1 score.

Accuracy is the proportion of correctly predicted labels out of the total number of predictions.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + FN}$$

F1 score is the harmonic mean of precision and recall, offering a better balance between precision and recall when the data is imbalanced:

$$\text{F1} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

3.2.4 Experiment Results

RNN: Accuracy = 0.8017, F1 Score = 0.8002

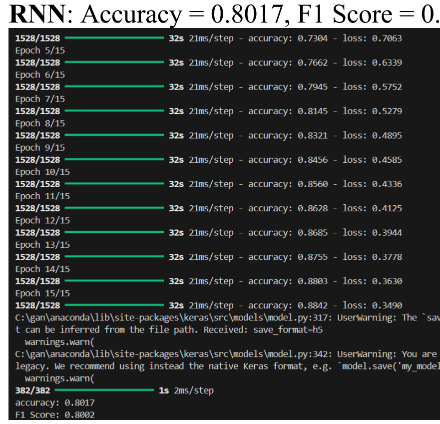

Figure 1 result of RNN

LSTM: Accuracy = 0.7704, F1 Score = 0.7732

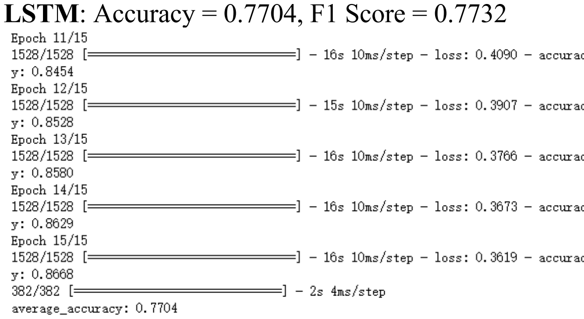

Figure 2 result of LSTM

BERT: **A**ccuracy = 0.8988, F1 Score = 0.8980

```
Epoch 1/3
382/382 [==============================] - 132s 327ms/step - loss: 0.6741 - accuracy: 0.714
0
Epoch 2/3
382/382 [==============================] - 125s 328ms/step - loss: 0.4180 - accuracy: 0.832
9
Epoch 3/3
382/382 [==============================] - 125s 327ms/step - loss: 0.2337 - accuracy: 0.909
5
96/96 [==============================] - 12s 114ms/step - loss: 0.2826 - accuracy: 0.8989
Test Loss: 0.2825864553451538, Test Accuracy: 0.8988874554634094
96/96 [==============================] - 12s 111ms/step
F1 Score: 0.8980906188767069
```

Figure 3 result of BERT

Compared three model by bar chart

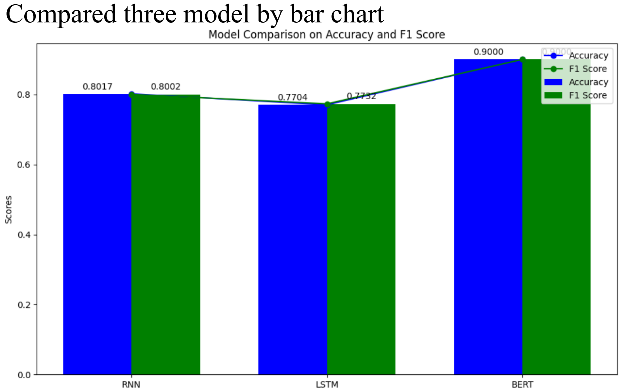

Figure 4 compare the models

IV. Discussion

RNN and LSTM, as variants of recurrent neural networks, are good at processing sequence data and long-term dependencies. However, in the data set of this study, because sequence relationships do not require long-term dependence, simple RNNS may be more suitable, and thus their performance is better than LSTM. In contrast, BERT, based on Transformer architecture, can better understand semantics and context through pre-training of large-scale corpus, and is significantly superior to RNN and LSTM in sentiment analysis tasks, showing higher accuracy and F1 scores. Overall, BERT performed best on the Twitter sentiment analysis task, but different tasks and data sets may require selecting the most appropriate model for the specific situation.

Deep learning in sentiment analysis faces several technical challenges. Firstly, human emotional expression is complex, often including mixed emotions, sarcasm, and humor, which traditional models struggle to capture fully. Secondly, multilingual support is inconsistent, especially for low-resource languages that lack high-quality annotated data. Additionally, according to Mohammed and Kora (2023), sentiment analysis relies on contextual understanding, but current models still struggle to capture long-range dependencies and complex contexts.

Privacy and ethics are also significant concerns, as sentiment analysis involves extensive personal data. Mann and Matzner (2019) emphasize the importance of ensuring compliance with privacy regulations, such as GDPR, and highlight the need to manage algorithmic bias to avoid perpetuating social inequalities. Furthermore, deep learning models lack transparency, making it difficult to interpret their decisions—particularly problematic in fields like healthcare and law, where high accuracy and accountability are critical.

V. Conclusion

Based on the Twitter sentiment dataset, this study explores the effectiveness of using a deep learning model for early warning of negative emotions on social media. By comparing the performance of RNN, LSTM and BERT models, the results show that BERT performs best in accuracy and F1 scores, and is suitable for capturing complex emotional semantics to detect and alert negative emotions more effectively. However, for different data sets and task scenarios, the choice of model still depends on the situation. While deep learning has shown significant advantages in sentiment analysis, it still faces challenges in areas such as emotional complexity, multilingual support, privacy and ethics. Future research should focus on improving the transparency and interpretability of the model to ensure the reliability and security of sentiment analysis in the field of negative emotion warning.

References

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks, 5(2), 157-166. https://doi.org/10.1109/72.279181

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Desmet, B., & Hoste, V. (2017). Online suicide prevention through optimizing the prediction of suicidal tendencies using Twitter data. Procedia Computer Science, 138, 194-198. https://doi.org/10.1016/j.procs.2018.10.027

Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. Procedia Computer Science, 161, 765-772. https://doi.org/10.1016/j.procs.2019.11.181

Huang, X., Sangroya, A., & Liu, Q. (2021). LSTM based sentiment analysis for cryptocurrency prediction. In C. S. Jensen et al. (Eds.), Database Systems for Advanced Applications. DASFAA 2021. Lecture Notes in Computer Science (Vol. 12683). Springer, Cham. https://doi.org/10.1007/978-3-030-73200-4_47

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Ismail, N. H., Liu, N., Du, M., He, Z., & Hu, X. (2020). A deep learning approach for identifying cancer survivors living with post-traumatic stress disorder on Twitter. BMC Medical Informatics and Decision Making, 20(Suppl 4), 254. https://doi.org/10.1186/s12911-020-01272-1

Mann, M., & Matzner, T. (2019). Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. Big Data & Society, 6(2), 1-14. https://doi.org/10.1177/2053951719895805

Mohammed, A., & Kora, R. (2023). A comprehensive review on ensemble deep learning: Opportunities and challenges. Journal of King Saud University-Computer and Information Sciences, 35(5), 2145-2165. https://doi.org/10.1016/j.jksuci.2023.01.014

Nemes, L., & Kiss, A. (2020). Social media sentiment analysis based on COVID-19. Journal of Information and Telecommunication, 5(1), 1–15. https://doi.org/10.1080/24751839.2020.1790793

Singh, R., Singh, R., & Bhatia, A. (2018). Sentiment analysis using machine learning techniques to predict outbreaks and epidemics. International Journal of Advanced Science and Research, 3(2), 19-24.

Singh, M., Jakhar, A. K., & Pandey, S. (2021). Sentiment analysis on the impact of coronavirus in social life using the BERT model. Social Network Analysis and Mining, 11, 33. https://doi.org/10.1007/s13278-021-00737-z

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational Linguistics, 37(2), 267-307. https://doi.org/10.1162/COLI_a_00049