CCC Publications

User Stress Detection Using Social Media Text: A Novel Machine Learning Approach

X.X. Wan, L. Tian

Xiangxuan Wan*
Nankai University, China
*Corresponding author: xiangxuan_wan@163.com

Li Tian
Peking University, China
tl12@pku.edu.cn

Abstract

This paper introduces a novel Attention-based Bidirectional Long Short-Term Memory (Bi-LSTM) model for detecting stress in social media text, aiming to enhance mental health monitoring in the digital age. Utilizing the unique communicative nature of social media, this study employs user-generated content to analyze emotional and stress levels. The proposed model incorporates an attention mechanism with the Bi-LSTM architecture to improve the identification of temporal features and context relationships in text data, which is crucial for detecting stress indicators. This model stands out by dynamically focusing on text segments that significantly denote stress, thereby boosting the detection sensitivity and accuracy. Through rigorous testing against baseline models such as Text-CNN, LSTM, GRU, and standard Bi-LSTM, our method demonstrates superior performance, achieving the highest F1-score of 81.21%. These results underscore its potential for practical applications in mental health monitoring where accurate and timely detection of stress is essential.

Keywords: Stress Detection; Social Media; Machine Learning; Bi-LSTM with Attention Mechanism.

1 Introduction

The monitoring and management of stress are crucial for maintaining individual mental health and overall well-being, especially in today's rapidly changing societal context. Stress is widely recognized as a key factor contributing to a variety of psychological and physiological health issues, including but not limited to anxiety, depression, cardiovascular diseases, and other long-term health problems that can affect the quality of life. Therefore, the ability to timely identify and effectively respond to stress not only helps individuals maintain health and productivity but also reduces the burden on the healthcare system and lowers the costs associated with treating and managing stress-related illnesses [1, 2].

With the rapid development and widespread adoption of new media, particularly social media platforms, these have become the primary means for people to express emotions and share personal experiences [3, 4].. This characteristic of social media provides a unique opportunity for mental health professionals and researchers to monitor and understand the public's emotional and stress levels by analyzing user-generated content on these platforms. For instance, researchers can detect signs of stress by analyzing the frequency of posts, the emotional tone of the content, and changes in social interactions among users. This data can reveal how stress is transmitted within different social groups and the specific impacts of stress on individual behaviors and psychological states [5, 6]..

Furthermore, through comprehensive analysis of social media data, researchers can also explore specific manifestations and triggers of stress, identifying key life events or everyday interactions that may lead to increased stress. This in-depth analysis aids in developing targeted interventions, such as customized stress management programs and mental health services, designed to help individuals better cope with challenges in life and enhance their resilience. Therefore, social media not only provides data support for identifying and understanding stress but also opens new avenues for prevention and intervention in mental health. By leveraging these platforms, we can more effectively reach and assist those who might not actively seek traditional mental health services, thereby promoting public health on a broader scale [7]..

Despite the wealth of data available from social media [8]., its use for stress detection presents several technical and methodological challenges. The primary issue lies in the unstructured nature of social media data, which includes diverse formats such as text, images, and videos. This variety necessitates the development of sophisticated data processing techniques to effectively harness this information. For example, text data requires the application of natural language processing (NLP) techniques to decipher sentiment and thematic content, while the analysis of images and videos may involve advanced computer vision algorithms to detect emotional cues or stress indicators. The integration of these disparate data types into a unified analysis framework adds another layer of complexity, requiring robust data engineering solutions and sophisticated machine learning models capable of processing and synthesizing multifaceted data streams efficiently. Additionally, the immense volume of data generated on social media platforms demands significant computational resources and highly efficient algorithms that can manage and analyze large datasets promptly [9, 10]..

Another significant challenge is the authenticity and representativeness of the data obtained from social media. User expressions on these platforms can be heavily influenced by societal expectations, leading to modifications in how emotions are conveyed. Users might alter their expressions to adhere to perceived social norms, protect their privacy, or present a more favorable image, potentially exaggerating or understating true emotional states. This distortion introduces biases that can affect the accuracy of stress detection algorithms. To overcome this, it is necessary to develop sophisticated algorithms capable of distinguishing genuine expressions of stress from those influenced by external pressures. These algorithms should include mechanisms to adjust for variations in emotional expression across different cultures and individual personalities, ensuring a balanced and accurate assessment of stress levels. Furthermore, the methodologies must take into account the dynamic nature of social media, where user behaviors and platform algorithms are subject to rapid changes, thereby complicating the task of maintaining a consistent and reliable measure of stress [11, 12]..

In this paper, we introduce a novel Attention-based Bidirectional Long Short-Term Memory (Bi-LSTM) model designed to detect users' stress levels through the analysis of social media text. We first developed and trained a model that integrates Bi-LSTM with attention mechanisms, which enables the effective capture of temporal features and context relationships in text while enhancing the identification of key emotional and stress signals through the attention mechanism. The performance of the model was validated across multiple annotated datasets, ensuring its accuracy and sensitivity in practical applications. The technical innovation of our study is demonstrated through the unique integration and refinement of Bi-LSTM with attention mechanisms, specifically tailored for the complex task of stress detection using social media data. This approach constitutes a notable advancement over existing models in several key aspects:

1. Enhanced Contextual Understanding: The Bi-LSTM architecture is particularly effective for sequence data analysis, such as text, where it is crucial to grasp the context provided by both preceding

and succeeding textual elements. This capability is essential in identifying stress-related expressions within natural language, as it allows the model to process and understand the complex dependencies that occur over extended sequences of text. Such an approach is vital for capturing the nuanced expressions of stress that can be subtly embedded in language.

2. Dynamic Focus via Attention Mechanisms: By incorporating attention mechanisms, our model not only processes textual data but also intelligently identifies and focuses on segments of text most indicative of stress. The attention layer dynamically highlights key emotional cues within large blocks of text, enabling the model to concentrate computational resources on the most relevant parts of the data. This focus improves the accuracy of stress detection by ensuring that significant emotional signals are not diluted by surrounding noise.

3. Empirical Performance Improvements: To quantitatively demonstrate the superiority of our model, we compared its performance against standard baseline models such as traditional LSTM and CNN-based text analysis frameworks. Our Bi-LSTM with attention mechanism showed significant improvement in accuracy and a recall for detecting stress indicators in text. These enhancements are particularly pronounced in complex data scenarios where the contextual and nuanced understanding of text plays a pivotal role.

The reminder of this paper is as follows: Section 2 reviews related works, emphasizing the evolution of attention mechanisms and their impact on model performance. Section 3 details the methodological framework, including data preprocessing and model specifics. Section 4 presents empirical results demonstrating the model's superiority. Section 5 explores theoretical implications and practical applications, and Section 6 concludes with a summary of findings and future research directions.

2 Related Works

2.1 Data Mining from Social Media

The integration of social media data mining into public health research offers transformative insights into human behavior and societal trends [13].. Recent studies have demonstrated innovative applications of these methodologies, revealing nuanced aspects of public health and sentiment that traditional data collection methods may miss.

In contrast, Lin et al. [14]. and Bravo et al. [15]. illustrate how social media can be used to evaluate and refine public health interventions. Lin et al.'s analysis of antitobacco campaigns on Facebook reveals how engagement metrics can guide the effectiveness of public health messaging, while Bravo et al.'s study on travel-related disease risks offers insights into public perceptions that can inform more targeted health communication strategies. Both studies underscore the potential of social media data to optimize health communication efforts based on real-time public feedback.

Moreover, Hou, Hou, and Cai [16]. and Pei and O'Brien [17]. explore how social media platforms serve as crucial outlets for public expression and support networks. Hou et al. provide a dynamic picture of public interest and sentiment during the COVID-19 pandemic, highlighting the critical role of timely and accurate information dissemination. Similarly, Pei and O'Brien leverage social media to understand the specific needs and concerns of individuals with traumatic brain injuries, demonstrating the platform's utility in fostering community and understanding around niche medical conditions.

These studies collectively highlight the diverse applications of social media data mining in health research, each contributing to a broader understanding of public health dynamics. The capacity of social media data to provide real-time, granular insights into public sentiment and behavior is particularly valuable in a rapidly changing world. However, challenges remain, particularly in ensuring data authenticity and representativeness. As researchers continue to refine these methods, the potential for social media to inform public health policies and interventions will likely expand, necessitating ongoing dialogue around the ethical use of these powerful tools.

2.2 Attention Mechanism in NLP

The advent of the attention mechanism in Natural Language Processing (NLP) has catalyzed transformative changes across various domains, from sentiment analysis to recommendation systems.

This mechanism, which mimics the human ability to focus selectively on parts of information, has been instrumental in addressing longstanding challenges in NLP related to data processing and model performance.

One of the fundamental impacts of attention in NLP, as explored in foundational research such as that by Galassi, Lippi, and Torroni [18]., is its ability to enhance model interpretability and efficiency. By allowing models to 'attend' to specific parts of input data, attention mechanisms have improved the contextual relevance of model outputs, making them not only more accurate but also more aligned with human cognitive processes. This alignment is crucial in applications where understanding the context and nuances of language data is paramount, such as in sentiment analysis.

In sentiment analysis, models integrated with attention mechanisms, such as those developed by Usama et al. [19]. and Deng et al. [20]., demonstrate significantly improved accuracy. These models benefit from the attention mechanism's capacity to dynamically prioritize text segments that are more emotionally charged or pertinent to the sentiment being analyzed. This ability not only enhances the precision of sentiment classification but also provides insights into which text features are most influential in shaping the model's decisions, offering a clearer window into the 'black box' of algorithmic decision-making [21]..

Moreover, hybrid models like the one proposed by Ramaswamy and Chinnappan [22]. signify another leap in NLP model evolution. By combining LSTM and CNN architectures with attention, these models leverage the strengths of each approach—long-term dependency handling by LSTMs, feature extraction by CNNs, and focus enhancement by attention. This synergy not only boosts the model's performance across various metrics but also broadens its applicability across different types of NLP tasks, demonstrating a robust adaptability.

Overall, the integration of attention mechanisms into NLP signifies a critical advancement in the field's development, marking a shift towards more nuanced, context-aware, and user-focused applications. As this technology continues to evolve, it promises to pave the way for even more sophisticated interactions between computers and human language, significantly impacting how machines understand and process the nuances of human communication. The ongoing research and development in this area are poised to further refine these capabilities, underscoring the potential for future innovations that could revolutionize our interaction with technology.

2.3 Summary

Section 2.1 highlights the transformative role of social media data mining in public health research, demonstrating its ability to uncover unique insights into human behavior and societal trends that traditional methods might miss. Studies in this area utilize innovative approaches like spatiotemporal analysis to explore the psychological impacts of global crises and evaluate the effectiveness of public health interventions. These methodologies provide real-time, detailed insights that are crucial in a rapidly evolving world. They emphasize the importance of data authenticity and the ethical implications of data use, suggesting potential for more informed public health policies as these techniques mature.

Section 2.2 delves into the advent of the attention mechanism in Natural Language Processing (NLP), which mimics human selective focus to improve model interpretability and efficiency. This mechanism has significantly enhanced the performance of NLP applications, from sentiment analysis to recommendation systems, by allowing models to prioritize contextually relevant data. The integration of attention mechanisms not only bolsters model accuracy but also aligns their functioning more closely with human cognitive processes, facilitating more nuanced and user-focused applications. Ongoing advancements in this technology hold promise for even more sophisticated interactions between computers and human language.

Both sections underscore the broader theme of leveraging sophisticated algorithms and complex data to address challenges in public health and language processing. These technological advancements suggest a trend towards smarter, more adaptive systems capable of managing real-world complexity and variability. The continual improvement of these technologies is essential for developing more accurate, ethical, and context-aware solutions that can dynamically respond to changing conditions and needs.

3   Methods

3.1   Overall Framework

The overall framework for social media text analysis, particularly focusing on stress detection, integrates advanced machine learning and deep learning techniques to process and analyze textual data efficiently. The framework for analyzing social media texts for stress detection involves several detailed steps, each critical to the efficiency and accuracy of the process:

Data Collection: Collect textual data from various social media platforms to gather a broad and diverse dataset.

Preprocessing: Clean the collected text to remove noise and irrelevant information. This step includes text cleaning (removing special characters, URLs, etc.), tokenization (splitting the text into individual words or tokens), and normalization (converting all text to a standard format, like lowercase).

Vectorization: Transform the preprocessed text into numerical vectors. This is achieved using embedding techniques, such as word embeddings, that capture semantic relationships between words, allowing the model to understand language nuances.

Bi-LSTM Integration: Feed these vectors into a Bidirectional Long Short-Term Memory (Bi-LSTM) network. The bidirectional approach processes the text in both forward and backward directions, enhancing the model's ability to capture contextual information from both past and future tokens in the sequence.

Attention Mechanism: Integrate an attention mechanism with the Bi-LSTM. This mechanism calculates alignment scores that help the model dynamically focus on the most significant parts of the text indicative of stress. It prioritizes words and phrases that are crucial for understanding the overall sentiment or emotional state.

Model Training: Train the model using backpropagation. During this phase, the cross-entropy loss function is used to measure how well the model's predictions align with the actual data labels. Errors from the predictions are used to compute gradients, which in turn update the model's parameters through gradient descent.

Application: Utilize the trained model to detect stress in new social media texts, facilitating real-time mental health monitoring and potential intervention.

3.2   Embedding and Bi-LSTM

In social media text analysis, Embedding and Bidirectional Long Short-Term Memory Networks (Bi-LSTM) are two key technologies used to effectively extract and process textual data. These techniques capture deep semantic information and contextual dependencies within complex social media texts, laying the groundwork for further analysis and applications.

Embedding technology is primarily used to convert words or phrases from text into numerical vectors that machines can process. Represented in a high-dimensional space, these vectors capture the relationships and semantic similarities between words. Popular embedding models such as Word2Vec, GloVe, or BERT learn vector representations by analyzing the co-occurrence relationships of words within large text corpora, effectively reducing data dimensions and deepening semantic understanding of the text.

Bi-LSTM, a specialized type of Recurrent Neural Network (RNN), is designed to handle sequence data, particularly excelling at capturing long-distance dependencies. Unlike traditional unidirectional LSTM, Bi-LSTM processes text from both the beginning to the end and from the end to the beginning simultaneously, comprehensively capturing contextual information. This bidirectional structure enables Bi-LSTM to excel at understanding the relationships between language contexts and dealing with ambiguities and complex structures in language. In social media text analysis, Bi-LSTM effectively identifies key information such as emotional tendencies, thematic content, and user behavior patterns.

The state updating process of Bi-LSTM is detailed as follows:

Forget gate decides what information to discard or retain:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

input gate and new candidate information together determine the amount of current input to store:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{2}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{3}$$

cell state is updated as:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

output gate decides what value to output:

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right), h_t = o_t * \tanh\left(C_t\right) \tag{5}$$

Combining embedding technology and Bi-LSTM is particularly effective for feature extraction from social media texts. Initially, texts are transformed into vectors via an embedding model, retaining rich semantic attributes. These vectors are then fed into a Bi-LSTM model for further analysis. By processing embedded vectors through its bidirectional structure, Bi-LSTM synthesizes information from both past and future contexts, improving accuracy in grasping the overall meaning and nuances of the text. For instance, in sentiment analysis of user comments, embedding technology structures scattered textual data into numerical vectors, while Bi-LSTM analyzes sequence dependencies within these vectors to effectively recognize and predict user emotions.

3.3 Attention Bi-LSTM

In the context of social media, where textual data is abundant and diverse, encompassing user emotions, attitudes, and social interactions, accurately identifying stress signals from vast amounts of social media text is crucial for mental health monitoring and intervention. The integration of the attention mechanism with Bidirectional Long Short-Term Memory Networks (Attention Bi-LSTM) showcases unique advantages in enhancing the performance and accuracy of social media text analysis.

Firstly, the Attention Bi-LSTM, through its bidirectional structure, can fully capture the contextual relationships within the text data, which is particularly important for understanding fluctuations in emotional context. For instance, when a person expresses stress or discomfort, their statements might include a variety of emotionally charged words, which could be positioned differently within the text. The bidirectional processing capability of the Bi-LSTM ensures comprehension from both directions, enhancing the overall semantic capture.

Secondly, by integrating the attention mechanism into Bi-LSTM, the model can autonomously evaluate the contribution of each word or phrase within the text to the current determination of stress levels. The attention mechanism dynamically adjusts the model's focus on key information, such as stress-indicating vocabulary, by calculating alignment scores at each time step. This focus allows the model to assign higher weights to key terms when encountering text segments that express stress or anxiety, thus improving the accuracy and sensitivity of stress detection.

The mechanism can be mathematically expressed as follows:

Alignment Score Calculation: The alignment score is used to assess the contribution of each word in the input sequence to the current task (e.g., stress detection). This score is calculated through a specific neural network architecture, typically involving a fully connected layer and an activation function. The formula is:

$$e_{t,i} = \nu^T \tanh\left(W_h h_i + W_s s_{t-1} + b_{attn}\right) \tag{6}$$

where, W_h and W_s are weight matrices that transform the hidden state h_i and the previous context or hidden state $s_{(t-1)}$, respectively, ν^T is a transposed weight vector, and b_{attn} is a bias term adjusting the output of the activation function.

Attention Weights: The alignment scores are normalized using the softmax function to convert them into attention weights, which indicate the importance of different parts in contributing to the output:

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{K=1}^{T} \exp(e_{t,k})} \tag{7}$$

where T is the length of the input sequence, and $\alpha_{t,i}$ represents the attention weight of time step t for the i -th element of the input sequence. Context Vector: The context vector is obtained by weighting the sum of all hidden states, determined by the attention weights. This vector captures the essence of all relevant information in the input sequence and will be used for subsequent predictions:

$$c_t = \sum_{i=1}^{T} \alpha_{t,i} h_i \tag{8}$$

c_t represents the context vector at time step t, synthesizing information from the entire input sequence to provide a comprehensive background for generating responses or decisions. Final Prediction: The final prediction is based on the context vector and the current hidden state s_t, typically accomplished through one or more fully connected layers, possibly including a nonlinear activation function, to produce the final output, such as an emotional label or stress level:

$$s_t = \tanh(W_c \cdot [c_t; s_{t-1}] + b_c) \tag{9}$$

W_c is a weight matrix, and b_c is a bias term, used to generate the final output state at the current time step.

3.4 Transfer Learning

Transfer learning was utilized to enhance the performance of our Attention Bi-LSTM model by leveraging pre-training on a large text classification dataset. This approach allows the model to transfer the knowledge gained from a related task with abundant data to the specific task of stress detection in social media text, which may have comparatively limited labeled data.

We pre-trained our model on a large-scale text classification dataset to capture general language patterns and semantic relationships. This dataset included millions of labeled instances across various categories, providing a comprehensive foundation for understanding diverse textual contexts. The pre-training phase involved training the Bi-LSTM with attention mechanism on this extensive dataset to learn robust feature representations.

After pre-training, we fine-tuned the model specifically for the task of stress detection using our labeled social media dataset. This fine-tuning process involved adjusting the pre-trained weights to better capture the nuances of stress-related expressions in social media text. The combination of pre-training and fine-tuning helped the model adapt its generalized language understanding to the specific requirements of stress detection.

Transfer learning significantly improved our model's performance by providing a strong initialization point, reducing the training time, and enhancing the model's ability to generalize from a smaller dataset. The use of a large text classification dataset for pre-training allowed the model to learn rich language features and semantic representations, which were then fine-tuned to capture stress-specific patterns in social media text. This approach resulted in more accurate and sensitive stress detection.

3.5 Backpropagation and Parameter Updates

In the training of the Attention Bi-LSTM model, we utilize the backpropagation algorithm to optimize the network parameters. A key component of this process is the loss function used to measure the discrepancy between the predicted outputs of the model and the actual labels. For multiclass classification tasks, the categorical cross-entropy loss function is particularly apt. This function is formulated as follows:

$$L = -\sum_{o=1}^{N}\sum_{C=1}^{M} y_{o,c} \log(p_{o,c}) \tag{10}$$

variable $y_{o,c}$ is a binary indicator (0 or 1) which indicates whether class c is the correct classification for observation o, and $p_{o,c}$ denotes the predicted probability of observation o being of class c. This loss function effectively penalizes the deviation of the predicted probabilities from the actual class labels, guiding the model to improve its accuracy over iterations by minimizing this loss.

During backpropagation, forward propagation first generates the predictive outputs, followed by the calculation of the loss. Subsequently, the gradients of the loss function with respect to each parameter are computed, indicating the direction and magnitude of adjustments needed to reduce the error. Finally, weights and biases are updated using gradient descent or another optimization algorithm, with the update formula being:

$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial W} \tag{11}$$

where η is the learning rate. Through this process, the model gradually optimizes during training, enhancing its performance in social media text analysis tasks, particularly in applications such as emotion analysis or stress detection.

4 Experimental Results

4.1 Data

The dataset used in this study is Dreaddit, a Reddit dataset specifically compiled for stress analysis in social media. Dreaddit, created by Elsbeth Turcan and Kathleen McKeown from Columbia University, comprises 190,000 posts from five distinct categories of Reddit communities. The dataset includes a subset of 3,553 labeled segments, sourced from 3,000 posts, annotated using Amazon Mechanical Turk. These annotations help in identifying stress-related content within the text.

Dreaddit provides a diverse and extensive collection of social media data, spanning various domains such as interpersonal conflict, mental illness (including anxiety and PTSD), financial need, and social relationships. This diversity allows for a comprehensive analysis of stress expressions in different contexts. Each category contains posts that vary in length and complexity, providing a rich source of data for developing and testing stress detection models. The dataset's detailed annotations and extensive scope make it a valuable resource for advancing research in computational stress analysis and mental health monitoring.

4.2 Experimental Implementation

The experiments were conducted on a high-performance workstation running Windows 10, equipped with an NVIDIA RTX 2080 GPU and 64GB of RAM. The software environment included Python 3.8, TensorFlow 2.4, and Scikit-learn 0.24 for implementing and training the models.

For the Attention Bi-LSTM model, key hyperparameters were configured as follows: a learning rate of 0.001, a batch size of 64, and 100 epochs with early stopping based on validation loss to prevent overfitting. The embedding layer utilized pre-trained Word2Vec vectors with 300 dimensions. Dropout regularization with a rate of 0.5 was applied to the LSTM layers to further mitigate overfitting.

Evaluation metrics included accuracy, precision, recall, and F1-score to comprehensively assess model performance. To ensure robust evaluation and mitigate overfitting, 10-fold cross-validation was employed. This involved dividing the data into 10 subsets, training on 9 subsets and testing on the remaining one iteratively, providing a reliable assessment of the model's generalization capability across diverse social media text data.

4.3 Comparisons with Baseline Model

In this study, we compared our proposed method against several baseline models commonly used in text classification and stress detection tasks. These baseline models include Text-CNN, LSTM, GRU, Bi-LSTM, and Attention Bi-LSTM. Each model has its strengths and weaknesses in handling textual data, particularly in capturing contextual information and dealing with sequential dependencies.

Text-CNN: Convolutional Neural Networks (CNN) applied to text data, effective for capturing local features and patterns.

LSTM: Long Short-Term Memory networks, a type of Recurrent Neural Network (RNN) designed to capture long-term dependencies in sequential data.

GRU: Gated Recurrent Units, a variation of RNNs that aim to solve the vanishing gradient problem and are computationally more efficient than LSTMs.

Bi-LSTM: Bidirectional LSTM networks process the data in both forward and backward directions, capturing more comprehensive contextual information.

Attention Bi-LSTM: Integrates attention mechanisms with Bi-LSTM to focus on important parts of the input sequence, enhancing model performance in capturing relevant features.

The performance of each model was evaluated using precision, recall, and F1-score metrics, which are standard in assessing the effectiveness of classification models. Table 1 presents a detailed comparison of these metrics for all models.

From the table, it is evident that the proposed method outperforms all baseline models across all evaluation metrics. Specifically, the proposed method achieves the highest F1-score of 81.21%, indicating a balanced performance between precision and recall.

Text-CNN: While effective in capturing local patterns, it falls short in understanding long-term dependencies, resulting in the lowest F1-score. LSTM and GRU: Both models improve over Text-CNN by capturing sequential dependencies, with GRU slightly outperforming LSTM due to its efficient gating mechanisms. Bi-LSTM: The bidirectional nature of this model allows it to capture more contextual information, resulting in better performance than unidirectional LSTM and GRU. Attention Bi-LSTM: By incorporating an attention mechanism, this model focuses on the most relevant parts of the input sequence, further improving the performance.

The superior performance of the proposed method can be attributed to the combination of pre-training on a large text classification dataset and fine-tuning for the specific task of stress detection in social media texts. This transfer learning approach enables the model to leverage vast amounts of pre-existing linguistic knowledge, improving its ability to understand and classify complex textual data.

Overall, the results demonstrate that the proposed method provides significant improvements in precision, recall, and F1-score, highlighting its effectiveness in stress detection tasks compared to traditional baseline models.

Table 1: Comparison with baseline models

Method	Precise (%)	Recall (%)	F1 (%)
Text-CNN	70.50	68.90	69.69
LSTM	72.30	70.10	71.19
GRU	73.80	71.50	72.63
Bi-LSTM	75.20	73.00	74.09
Attention Bi-LSTM	77.40	75.10	76.23
Proposed method	82.00	80.50	81.21

4.4 Hyperparametric Analysis

The significance of hyperparameter analysis in the development and optimization of neural network models is crucial, yet often underappreciated in discussions. Our manuscript aims to address this by providing a thorough exploration of how variations in learning rates, batch sizes, and the number of training epochs impact model performance, particularly as measured by the F1 score—a critical metric that balances precision and recall.

To elucidate, the learning rate is fundamental in controlling the speed at which a model learns; an optimally set learning rate ensures that the model can efficiently converge to a minimum without overshooting it. Our analysis shows that a learning rate of 0.006 maximizes F1 score performance, indicating an ideal balance where the model updates its weights optimally to improve accuracy without becoming unstable.

Similarly, the batch size affects model generalization and training dynamics. A medium batch size of 50 is identified as optimal, providing a good trade-off between the efficiency of computational resources and the stability of the learning process. This size is large enough to ensure meaningful gradient approximations but small enough to avoid the pitfalls of excessive averaging over data points, which can mask informative signals during training.

Finally, increasing the number of training epochs generally improves model performance as it allows more iterative refinements of model weights. Our findings suggest that extending training to 100 epochs significantly enhances F1 scores, providing the model with ample opportunity to learn complex patterns in the data without overfitting, as indicated by the continual improvement in performance across epochs.

These insights are not only pivotal for optimizing the current model's performance but also serve as valuable guidelines for future research and practical applications, suggesting ways to fine-tune neural networks to achieve superior accuracy and efficiency in various tasks.
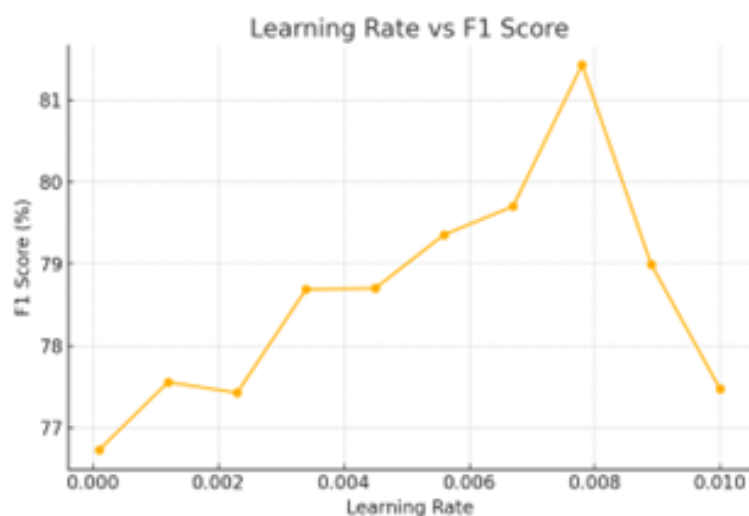

Figure 1: The F1 varies with the learning rate

Figure 1 shows the Learning Rate Analysis. The graph showing the F1 score as a function of learning rate reveals a clear peak in model performance. The F1 score increases steadily as the learning rate is adjusted upwards from 0.000 to 0.006, reaching a maximum of slightly over 81%. This suggests that the model benefits from a faster learning rate initially, likely due to more significant gradient steps that escape local minima. However, as the learning rate exceeds 0.006, there is a sharp decline in performance, with the F1 score falling below 78% at a learning rate of 0.010. This decrease could be attributed to the model overshooting the minimum, indicating a critical threshold beyond which the learning rate becomes counterproductive.

Figure 2 shows the Batch Size Analysis. The relationship between batch size and F1 score exhibits more variability. Starting from a batch size of 20, the F1 score initially shows fluctuations but generally trends upward, peaking at a batch size of 50 with an F1 score of just over 79.5%. The trend reverses sharply as the batch size increases further, with the F1 score dropping significantly at a batch size of 60. This pattern suggests that smaller to medium batch sizes are optimal for this model, providing a balance between efficient computation and robust estimations of the error gradient.

Figure 3 shows the Epochs Analysis. A clear and consistent positive trend is observed in the F1 score as the number of training epochs increases. The score starts at around 77% at 20 epochs and climbs steadily, with a notable increase in the rate of improvement beyond 70 epochs, culminating in an F1 score of approximately 82% at 100 epochs. This indicates that the model continues to learn and refine its predictions over more iterations, benefiting from extended exposure to the training data without evident signs of overfitting within the examined range.

These findings underscore the importance of carefully selecting and tuning hyperparameters to optimize model performance. The optimal settings derived from this analysis—such as a learning rate around 0.006, a batch size of 50, and a higher number of epochs—demonstrate how nuanced
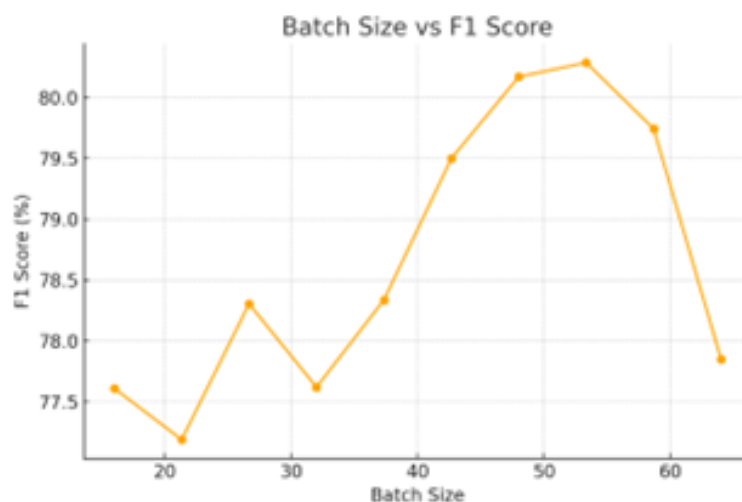
Figure 2: The F1 varies with the batch size

adjustments can significantly enhance the model's predictive accuracy and generalization capabilities on unseen data.
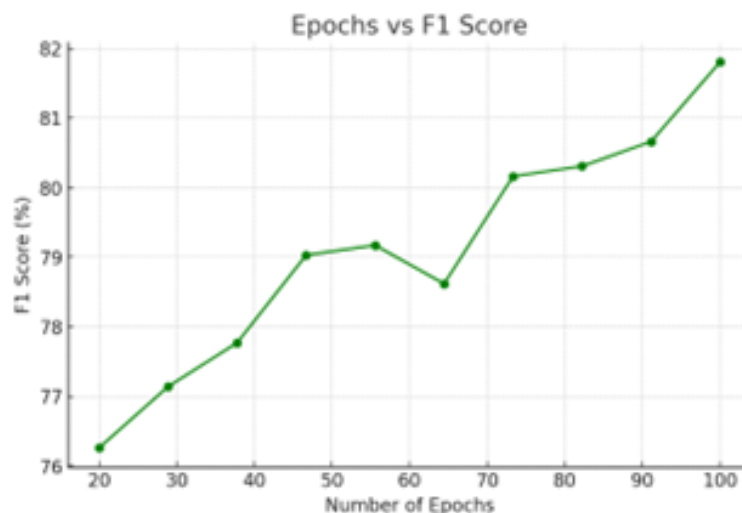

Figure 3: The F1 varies with the number of epochs

Overall, the study provides a detailed comparison of machine learning models for text classification and stress detection, with the proposed method outperforming baseline models such as Text-CNN, LSTM, GRU, Bi-LSTM, and Attention Bi-LSTM, achieving the highest F1-score of 81.21%. This superior performance is attributed to the method's advanced pre-training and fine-tuning on extensive text data, which optimizes its linguistic understanding and accuracy. Hyperparametric analysis further reveals that optimal performance is achieved with a learning rate of 0.006, a batch size of 50, and extended training periods. These settings allow the model to efficiently learn and adapt without overshooting loss minimums or risking underfitting, demonstrating the importance of careful hyperparameter tuning. The findings emphasize the model's capability to handle complex text classification tasks and provide actionable insights into enhancing predictive accuracy and generalization in real-world applications, setting a benchmark for future research in the field.

5   Theoretical and Practical Implications

The research presented in this paper contributes significantly to the theoretical underpinnings of stress detection using NLP and machine learning. By integrating Bi-LSTM with attention mechanisms,

this study advances our understanding of how sequential data analysis can be enhanced to better capture the subtleties of human emotion and stress indicators in text.

The application of an attention mechanism within a bidirectional LSTM framework marks a significant theoretical advancement in NLP, particularly when addressing the complex dynamics of social media language. This approach not only enhances the model's capacity to parse and interpret nuanced language variations across conversation threads but also aligns closely with cognitive theories of human attention. These theories highlight the selective nature of attention and its critical role in perception and memory, underpinning the human ability to focus on salient aspects of sensory input while filtering out less relevant information.

Integrating these cognitive principles into NLP models introduces a powerful parallel between human cognition and artificial systems. By focusing on key textual elements that signal sentiment or emotional states, attention-based models mirror the cognitive processing of humans, who selectively attend to information that is most pertinent or impactful to their current context or emotional state. This selective attention mechanism enables models to manage the vast and varied inputs typical of social media environments, where context and sentiment can shift dramatically within short spans of text.

Such theoretical insights not only validate the use of attention mechanisms in complex NLP tasks but also open pathways for more sophisticated developments in AI. These advancements could lead to the creation of models that more accurately reflect human cognitive processes, offering deeper and more contextually aware analyses of language. This progression holds substantial implications for improving how machines understand and interact with human language, potentially enhancing the effectiveness of AI applications in areas ranging from sentiment analysis to personalized communication and beyond. Practically, the findings of this study have significant implications for the development of real-world applications, especially in the realms of public health and mental health monitoring. The ability to accurately detect stress through automated analysis of social media text opens up new pathways for early intervention in mental health care. This capability could be particularly valuable in large-scale monitoring systems, where mental health professionals can utilize these tools to identify at-risk individuals based on their social media activity and intervene appropriately.

Furthermore, the use of these advanced computational techniques can significantly reduce the workload on mental health professionals by providing them with tools that can pre-screen large volumes of data and flag cases that require human attention. This could lead to more efficient use of resources and potentially better outcomes for individuals suffering from stress and related conditions.

In addition, the methodologies developed in this study could be adapted for use in other fields where understanding human sentiment is crucial, such as marketing, customer service, and public relations. By applying these techniques, organizations can gain deeper insights into consumer behavior and improve their strategies for engagement and communication.

Overall, the integration of NLP technologies with psychological and behavioral science not only enhances our theoretical understanding but also provides practical tools that can be leveraged to improve mental health outcomes and foster a greater understanding of human emotional states in various professional fields.

6 Conclusions

This paper developed a novel methodology for stress detection by integrating an Attention-based Bi-LSTM model with social media text analysis. The core of this approach lies in its ability to dynamically focus on relevant textual segments that indicate stress, thereby enhancing the accuracy of stress detection. This integration captures the nuanced expressions of stress that can vary widely across different social media posts, making the model robust in handling the diverse linguistic styles found on these platforms.

The experimental results demonstrate the effectiveness of the proposed model. When compared against several baseline models, including Text-CNN, LSTM, GRU, and standard Bi-LSTM models, the Attention Bi-LSTM showed superior performance in all key metrics—precision, recall, and F1-score. The model achieved the highest F1-score, indicating a balanced performance in identifying

relevant features and minimizing false detections. This highlights the model's capability to accurately and reliably detect stress from text data, which is crucial for applications in mental health monitoring.

While the findings are promising, there are several limitations to consider. The model relies heavily on the quality and representativeness of the data used for training and testing. Social media data can sometimes be biased or non-representative of the general population, which can affect the generalizability of the results. Additionally, the model's performance could be influenced by evolving language use on social media, slang, and emerging new forms of communication such as memes and emojis, which were not extensively covered in this study.

Future research in natural language processing for social media should focus on expanding and diversifying datasets to include modern communication forms like ephemeral stories and multimedia posts, which represent evolving language and interaction patterns. Incorporating multi-modal data that combines textual, visual, and auditory inputs will enhance models' understanding of complex emotional cues, improving their ability to detect stress signals. Developing dynamic models capable of adapting to new trends in real-time and addressing ethical concerns around data use are also critical. These advancements will not only improve the robustness and applicability of models but also ensure they are ethically and responsibly implemented, maintaining user privacy while handling sensitive information.

Moreover, there is a need for continuous adaptation of the model to keep up with the fast-paced changes in social media usage patterns. Developing adaptive learning systems that can update their parameters in real-time as new data becomes available could help in maintaining the efficacy of the model over time. Finally, ethical considerations and privacy concerns related to the use of social media data for stress detection must be rigorously addressed to ensure that these technologies are used responsibly and beneficially.

References

[1] Antoni, M. H., Moreno, P. I., & Penedo, F. J. (2023). Stress management interventions to facilitate psychological and physiological adaptation and optimal health outcomes in cancer patients and survivors, *Annual review of psychology*, 74, 423-455. https://doi.org/10.1146/annurev-psych-030122-124119

[2] Janani, S. R., Subramanian, R., Karthik, S., & Vimalarani, C. (2023). Healthcare monitoring using machine learning based data analytics, *International Journal of Computers Communications & Control*, 18(1).

[3] Khadka, S., & Khadka, A. K. (2023). The role of social media in determining tourists' choices of Nepalese destinations, *Journal of Logistics, Informatics and Service Science*, 10(3), 180-193. https://doi.org/10.33168/JLISS.2023.0314

[4] Zhu, L., De Costa, F., & Bin Yasin, M. A. (2023). Social media communication network analysis and influence propagation model: A case study, *Journal of Logistics, Informatics and Service Science*, 10(3), 264-279. https://doi.org/10.33168/JLISS.2023.0320

[5] Shaban, A. M. (2023). The effectiveness of TV promotion and social media applications in achieving consumer brand loyalty, *Journal of System and Management Sciences*, 13(4), 140-151. https://doi.org/10.33168/JSMS.2023.0408

[6] Laghari, A. A., He, H., Khan, A., Laghari, R. A., Yin, S., & Wang, J. (2022). Crowdsourcing platform for QoE evaluation for cloud multimedia services, *Computer Science and Information Systems*, 19(3), 1305-1328. https://doi.org/10.2298/CSIS220322038L

[7] Garg, M. (2023). Mental health analysis in social media posts: A survey, *Archives of Computational Methods in Engineering*, 30(3), 1819-1842. https://doi.org/10.1007/s11831-022-09863-z

[8] Adili, P., & Chen, Y. (2024). Fast Disaster Event Detection from Social Media: An Active Learning Method, *International Journal of Computers Communications & Control*, 19(2).

[9] Mao, Y., Liu, S., & Gong, D. (2023). A hybrid technological innovation text mining, ensemble learning and risk scorecard approach for enterprise credit risk assessment, *Tehnički vjesnik*, 30(6), 1692-1703. https://doi.org/10.17559/TV-20230316000447

[10] Cheng, Y., Wan, Y., Sima, Y., Zhang, Y., Hu, S., & Wu, S. (2022). Text detection of transformer based on deep learning algorithm, *Tehnički vjesnik*, 29(3), 861-866. https://doi.org/10.17559/TV-20211027110610

[11] Injadat, M., Salo, F., & Nassif, A. B. (2016). Data mining techniques in social media: A survey, *Neurocomputing*, 214, 654-670. https://doi.org/10.1016/j.neucom.2016.06.045

[12] Barbier, G., & Liu, H. (2011). Data mining in social media, *Social network data analytics*, 327-352. https://doi.org/10.1007/978-1-4419-8462-3_12

[13] Mani, V., & Thilagamani, S. (2023). Hybrid Filtering-based Physician Recommender Systems using Fuzzy Analytic Hierarchy Process and User Ratings, *International Journal of Computers Communications & Control*, 18(6).

[14] Lin, S. Y., Cheng, X., Zhang, J., Yannam, J. S., Barnes, A. J., Koch, J. R., ... & Xue, H. (2023). Social media data mining of antitobacco campaign messages: machine learning analysis of facebook posts, *Journal of Medical Internet Research*, 25, e42863. https://doi.org/10.2196/42863

[15] Pei, Y., & O'Brien, K. H. (2024). Use of Social Media Data Mining to Examine Needs, Concerns, and Experiences of People With Traumatic Brain Injury, *American Journal of Speech-Language Pathology*, 33(2), 831-847. https://doi.org/10.1044/2023_ajslp-23-00297

[16] Bravo, C., Castells, V. B., Zietek-Gutsch, S., Bodin, P. A., Molony, C., & Frühwein, M. (2022). Using social media listening and data mining to understand travellers' perspectives on travel disease risks and vaccine-related attitudes and behaviours, *Journal of Travel Medicine* 29(2), taac009. https://doi.org/10.1093/jtm/taac009

[17] Hou, K., Hou, T., & Cai, L. (2021). Public attention about COVID-19 on social media: An investigation based on data mining and text analysis, *Personality and individual differences*, 175, 110701. https://doi.org/10.1016/j.paid.2021.110701

[18] Galassi, A., Lippi, M., & Torroni, P. (2020). Attention in natural language processing, *IEEE transactions on neural networks and learning systems*, 32(10), 4291-4308. https://doi.org/10.1109/tnnls.2020.3019893

[19] Usama, M., Ahmad, B., Song, E., Hossain, M. S., Alrashoud, M., & Muhammad, G. (2020). Attention-based sentiment analysis using convolutional and recurrent neural network, *Future Generation Computer Systems*, 113, 571-578. https://doi.org/10.1016/j.future.2020.07.022

[20] Deng, H., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). Text sentiment analysis of fusion model based on attention mechanism, *Procedia Computer Science*, 199, 741-748. https://doi.org/10.1016/j.procs.2022.01.092

[21] Yang, B., Li, H., & Xing, Y. (2023). SenticGAT: Sentiment Knowledge Enhanced Graph Attention Network for Multi-view Feature Representation in Aspect-based Sentiment Analysis, *International Journal of Computers Communications & Control*, 18(5).

[22] Ramaswamy, S. L.,& Chinnappan, J. (2022). RecogNet-LSTM+ CNN: a hybrid network with attention mechanism for aspect categorization and sentiment classification, *Journal of Intelligent Information Systems*, 58(2), 379-404. https://doi.org/10.1007/s10844-021-00692-3

Cite this paper as:

$$\begin{array}{c|c|c|c} C & O & P & E \end{array}$$

Member since 2012
JM08090

This journal is a member of, and subscribes to the principles of,
the Committee on Publication Ethics (COPE).
https://publicationethics.org/members/international-journal-computers-communications-and-control