STM JOURNALS

International Journal of
Algorithms Design and
Analysis Review

ISSN: 2584-1866
Volume 4, Issue 1, 2026
January—June
DOI (Journal): 10.37591/IJADAR

https://journals.stmjournals.com/ijadar

Research IJADAR

Explainable Sentiment Mining Model in Mental Health Forums for Emotion Classification and Justification

Aditi Biswas[1,*], Surendra Kumar Yadav, Shruti Mathur[3]

Abstract

Understanding and interpreting emotions expressed in online mental health discussions plays a crucial role in enabling early detection of psychological distress and facilitating timely interventions. As individuals increasingly turn to digital platforms to share personal experiences and seek support, automated systems capable of accurately identifying emotional states can significantly assist clinicians, moderators, and support communities. This paper presents a deep learning–based sentiment mining and emotion classification framework specifically designed to analyze posts from mental health forums. The proposed architecture integrates Bidirectional Encoder Representations from Transformers (BERT)-based contextual embeddings to capture nuanced semantic relationships within text, along with a multi-channel convolutional neural network enhanced by residual connections. This hybrid structure effectively teaches both local phrase-level features and broader contextual patterns, improving classification performance across diverse emotional expressions. To ensure transparency and trustworthiness, especially important in sensitive domains such as mental health, we incorporate a gradient × embedding explainability mechanism. This component highlights token-level contributions toward each predicted emotion category, enabling users to understand which specific words most influenced the model's decision. The framework was evaluated on a benchmark mental health dataset consisting of seven emotion classes: joy, sadness, anger, fear, love, surprise, and others. Experimental results demonstrate an overall classification accuracy of 92%. Additionally, the visualization of influential words enhances interpretability, thereby supporting responsible and explainable AI-driven sentiment analysis in real-world mental health applications.

Keywords: Mental health, emotion detection, sentiment analysis, BERT, CNN, explainable AI

INTRODUCTION

In the 21st century, mental health has become one of the most acute public health concerns. According to the World Health Organization, millions of people are depressed, anxious, and ill due to stress every year, making it incredibly important to detect and intervene early. As the number of online forums and social media groups grows, people feel more comfortable expressing their emotions and problems in an open, usually anonymous forum. With data on emotional well-being on a scale, these platforms offer rich and real-time data. In contrast to organized medical records, posts in forums are unstructured and conversational and frequently include slang, abbreviations, emojis, and incomplete sentences. The emotions in these texts may be subtle, context-specific, or even mixed in a single post. This level of complexity cannot be handled using traditional sentiment analysis

***Author for Correspondence**
Aditi Biswas
E-mail: aditibiswas10122002@gmail.com

[1]Student, Department of Computer Science and Engineering, JECRC University Jaipur, Rajasthan, India
[2]Professor, Department of Computer Science and Engineering, JECRC University Jaipur, Rajasthan, India
[3]Assistant Professor, Department of Computer Science and Engineering, JECRC University Jaipur, Rajasthan, India

Received Date: January 01, 2026
Accepted Date: February 18, 2026
Published Date: March 20, 2026

Citation: Aditi Biswas, Surendra Kumar Yadav, Shruti Mathur. Explainable Sentiment Mining Model in Mental Health Forums for Emotion Classification and Justification. International Journal of Algorithms Design and Analysis Review. 2026; 4(1): 22–32p.

techniques such as lexicon-based scoring or bag-of-words with SVMs/logistic regression, which only categorize emotions as coarse (positive, negative, neutral) and finer-grained (fear, love, surprise). Recent advances in deep learning, particularly transformer-based models such as BERT, have enabled context-sensitive embeddings to obtain subtle semantics. These embeddings, combined with convolutional neural networks (CNNs) and residual connections, enhance the strong multi-class emotion classification. However, a major drawback is that the vast majority of deep learning systems are black boxes, which do not provide much interpretability, a crucial feature in sensitive fields such as mental health.

To address this, we introduce an explainable sentiment mining model that combines BERT embeddings with a multi-channel CNN and residual blocks to classify fine-grained emotions. One of the modules is gradient × embedding attribution, which identifies influential words in the prediction to enhance transparency. In the sentence I feel worthless and scared about tomorrow, tokens such as worthless and scared are highlighted, as they make model reasoning comprehensible to clinicians and stakeholders.

Some of the important contributions of this work are:

- A multi-class emotion recognition model based on a hybrid BERT + Multi-CNN-Residual network of the robustness of the method in the noise of forum texts.
- Gradient × embedding Integration explainability at the token level.
- Empirical analysis of a mental health dataset, with 92% accuracy, with equal accuracy and recall.
- Heatmaps of token-importance to improve clinical trust in predictions.

LITERATURE REVIEW

Social media has become a major area of critical research with the discovery of mental illness, which has changed significantly with its onset. The first works, which mostly focused on depression and anxiety, were based on conventional machine learning algorithms such as Naïve Bayes, support vector machine (SVM), and random forest. Nevertheless, these attempts were usually blocked by small datasets and small scales. With time, more advanced techniques have emerged, such as the ensemble voter model developed by Kumar et al. [1], which attained 85% accuracy on Twitter data. The area was also developed by Abdullah and Negied [2], who made a detailed comparison of 19 various models. They evaluated classical classifiers, nine ensemble models, and four large language models (LLMs), such as GPT and BERT, on a large volume of 7 million posts on Reddit. Their results offered a more refined structure: LLMs showed more significant results when used for clinical detection activities, and an F1-score of approximately 0.80, bearing in mind that simpler and more explainable models, including logistic regression and ensemble models, were more important in non-clinical identification and future forecasting activities.

Outside the fields of mental health, machine learning has found broad applications in analyzing more general emotional and sentimental trends in text. Singh and Sharma [3] conducted a comparative study of five classical algorithms, which provided a good performance standard. Their findings indicated that there was strictness in the overall sentiment analysis, and the SVM presented the most accurate result of approximately 89%. The practical application of these models in a high-stakes context was demonstrated by Brynielsson et al. [4], who used an SVM classifier to classify the emotional responses of the population to official alerts during the Hurricane Sandy crisis. Their model, which is included in the Alert4All system, was found to be effective in tracking general emotional tendencies in real time.

With the development of the field, deep learning architectures have started to establish new standards for emotion recognition, especially by incorporating hybrid architectures that could theoretically take advantage of the features of both convolutional and recurrent networks. Lövheim [5] suggested a three-dimensional theory that correlates monoamine neurotransmitters, serotonin,

dopamine, and noradrenaline, with simple human emotions. Each monoamine is depicted on an axis of a cube, with the eight corners representing the basic emotions based on affect theory. According to the model, there are certain patterns of monoamine activity associated with certain emotional conditions; for example, elevated levels of dopamine are associated with happiness, and low levels of serotonin are associated with depression or disgust. This spatial model provides a biological explanation for the variation in emotional and psychiatric symptoms to fill the gap between neurobiology and emotion theory to guide future studies and psychopharmacological research.

Sasidhar et al. [6] highlighted that human interaction is strongly emotive and is more manifested in writing on social media platforms. To solve the problem of emotion recognition in code-mixed Hindi–English, they developed a corpus of 12,000 sentences annotated with three emotions: happy, sad, and angry. Their CNN-Bidirectional Long Short-Term Memory (BiLSTM) architecture with bilingual pretrained embeddings achieved an accuracy of 83.21, which is useful in integrating CNN feature extraction and BiLSTM sequence learning. The study noted that these bilingual embedding-based hybrid structures are superior in identifying emotions in code-mixed languages and emphasized the importance of larger annotated datasets to improve bilingual emotion recognition.

One of the earliest studies using deep learning to detect sentiment and emotion in Bengali and Romanized Bengali texts was conducted by Tripo et al. [7]. They created three- and five-class sentiment models based on a newly created dataset of YouTube comments and six classes of feelings, including anger, disgust, fear, joy, sadness, and surprise. Their Long Short-Term Memory (LSTM) and CNN models performed well compared to classical methods such as Naïve Bayes and SVM, with accuracies of 65.97, 54.24, and 59.23, respectively. The most effective method was the use of trainable weight embeddings with skip-grams. The researchers reported less accuracy with Romanized Bangla because of inconsistency and errors, and higher scores with domain-specific data, as in video reviews [8]. Generally, this study created a good foundation for detecting the sentiment and emotion of the Bengali language and offered future applications of contextual embeddings to enhance the results [9].

RESEARCH METHODOLOGY

Data Collection and Description

1. *Data source:* The data will be written posts from online mental health forums where individuals share their experiences, feelings, and emotions. These platforms provide a real-time and naturalistic account of emotional states and are hence highly suitable for researching different expressions of emotions in everyday life [10].
2. *Data type:* It contains unstructured text, which is a mixture of colloquial language, typing, abbreviating, and technical terminology when discussing mental health. Posts with emojis and symbolic phrases are also plentiful and add more emotional meaning to them.
3. *Datasets size:* A sample of 4,00,000 posts was sampled to ensure that a computationally viable process was achieved and the diversity between all categories of emotions was preserved. The balance implies that the model can be subjected to a large number of linguistic patterns without unduly burdening computational resources.
4. *Justification:* The user-created information about forums ensures the provision of data on real-life relevance and the reflection of the peculiarities of emotions. Another implication of these forums is that they are representative of the true emotional states of users, which are critical for mental health analytics and can be useful in creating models applicable to real-life situations.

Data Preprocessing

1. *Text cleaning:* To be in a position to prepare the raw forum information, several preprocessing steps were performed to ensure uniformity and reduce noise in the raw forum information. There were no extra redundant tokens because the text was transformed into lowercase (e.g., happy to happy). Any URL, special symbols, and non-alphabetic characters were removed, and meaningful punctuation (e.g., "…" or "!!!") was retained. Unnecessary whitespace and line breaks were regularized for even tokenization.

2. *Missing values and duplicates:* Empty posts and duplicates were eliminated to ensure data quality and equal balance among emotion classes. This also avoids the learning of biased patterns in the model and guarantees a more representative sample in all emotion categories.
3. *Justification:* Cleaning lowers noise, meaning that models (e.g., BERT) can concentrate on meaningful patterns, reducing overfitting and enhancing generalization to unseen examples. Moreover, preprocessing guarantees the uniformity of input forms, retention of emotional indicators, and enhances the stability and predictability of a model in various textual sources.

This change will ensure that the emotional essence of the message is not lost, as gaps that create unwanted distractions are eliminated.

Label Encoding and Data Splitting

1. *Emotion labeling:* Emotions (e.g., joy, sadness, anger, fear, love, and surprise) were coded as numbers using a Label Encoder that transformed categorical labels into machine-understandable numbers. This transformation guarantees a uniform representation and makes learning easier for the neural networks.
2. *Train-test split:* The dataset is further split between 80% training and 20% testing data to enable the model to generalize over unknown data. The division uses a proportional representation of all classes of emotions and does not create an imbalance of classes during evaluation.
3. *Justification:* First, numerical encoding is required to use categorical classification using deep learning models because text labels in their original form cannot be directly processed. Equally, train-test splitting makes consistent performance evaluation, overfitting is avoided, and a realistic estimate of how the model will perform in the real world is made.

Feature Representation Using Contextual Embeddings

1. *Tokenization:* This is followed by the processing of textual data with the BERT tokenizer (bert-base-uncased), which breaks down sentences into sub-physical rather than full word units. The advantage is that the approach is capable of dealing with unusual words, misspellings, and language usage variance that is common in mental health forums. For example, the word unhappiness [un, unhappy, happy, unhappiness] does not lose its semantic meaning.
2. *BERT embeddings: The tokenized sequences then run through the BERT model, which generates contextual embeddings of each token.* Each token is represented as a 768-dimensional vector (in bert-base), which captures the meaning of the word itself and the meaning of the word in a sentence context. The word down in I am feeling down will have the meaning of sadness, but in look down the street it will have reference to space.
3. *Justification:* BERT embeddings provide powerful contextual representations by examining word dependence and sentence meaning. This is particularly important when it comes to the topic of mental health, as there are numerous undertones to emotions that are implied and strongly contextualized. Contrary to traditional bag-of-words or fixed embeddings (such as Word2Vec and GloVe), BERT is dynamically fine-tuned to the meaning, which seems more helpful given the classification issues related to emotions.

Multi-CNN-BERT Model Architecture

The architecture that is given is a combination of both BERT contextual embeddings and CNNs and residual connections to enhance the classification of effects in text.

1. *BERT Layer*
   - Gives rich contextualized embeddings that take into account semantic and syntactic subtleties.
   - Functions as a feature extractor, which converts raw text into dense, high-dimensional representations.

2. *Multi-Filter CNN*
   - It applies 1D convolution layers of sizes 3, 4, and 5, which represent local n-gram features.
   - For example, a kernel size of 3 can produce trigrams such as "feeling sad today," whereas a kernel size of 5 can produce longer contextual trigrams such as "I can't handle this anymore."
   - Such diversity in the sizes of the filters enables the model to monitor both short- and long-term emotional cues.
3. *Residual Convolution Block*
   - Skip connections allow direct information flow between layers to enhance the gradient flow during training.
   - The residual links eliminate fading gradients and can extract further features.
   - They also allow the model to acquire hierarchical emotional representations, which are a combination of surface-level expressions (e.g., angry) and deeper contextual meanings (e.g., angry because of loneliness).
4. *Global Max Pooling*
   - Chooses the most significant features of the whole sequence.
   - Make sure that important emotional signs are not missed in the long posts.
5. *Fully Connected Layers*
   - A two-layer feedforward neural network consisting of Rectified Linear Unit (ReLU) activations was used to convert the pooled features into class logits.
   - Dropout layers were implemented to avoid overfitting, which is particularly significant because the data size is small compared to the complexity of the model.

Justification: This combination model considers both global dependencies (through BERT) and local emotional triggers (through CNN). The residual block can improve the stability of the model, and thus, it can perform deeper learning without dropping, which is critical in detecting subtle, multi-layered emotional patterns in the discourse of mental health.

Model Training

1. *Loss function:* This is a Cross-Entropy Loss, which is the default choice for multi-class classification. It states that it penalizes false predictions more and rewards the model to assign higher weights to the correct label of emotion.
2. *Optimizer:* Adam optimizer with a learning rate of 2e-5 was employed, which is normally applied to fine-tune models based on transformers. Adam also optimizes the learning rate of each parameter, which ensures that the training is more stable and quicker.
3. *Epochs:* This was trained on 15 epochs to balance the computational cost and the risk of overfitting. Fewer epochs are used for fine-tuning because the embedded rich linguistic features of BERT are already available.
4. *Batch size:* The training and testing batch sizes were 16. This imparts a trade-off between the efficient utilization of Graphics Processing Unit (GPU) memory and a stable gradient update.
5. *Device utilization:* This makes use of GPU acceleration, where it takes much less time to run training than on a CPU.

Justification: The hyperparameters were chosen such that they converged but were not computationally expensive. Cross-Entropy Loss is appropriate for classification tasks, but the Adam optimizer can be useful for fine-tuning massive existing models such as BERT. The chosen batch size and epochs are viable considering the limited resources, in addition to providing competitive performance of the model.

Model Evaluation

1. *Quantitative metrics:* To evaluate the performance of a model, a combination of standard evaluation metrics is weighted to provide an overall measure of the learning accuracy and performance of a model at the class level:

- *Accuracy:* Refers to the degree of correctness of the prediction for the entire set of emotion classes.

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

- *Precision, recall, and F1-score:* Class-level information. When skewed data are used, it is significant to have zero tolerance for false positives (precision) and false negatives (recall) and to balance the two (F1-score).

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall}$$

- *Confusion matrix:* This tool offers the commonness of a particular emotion being wrongly forecasted as another (e.g., sadness being predicted as loneliness).

2. *Qualitative analysis:* Non-numerical sampling in addition to numerical analysis, random samples of predictions were checked manually to ensure that the predicted and actual labels had semantic congruence. This test establishes that the model can indeed pick fine-grained emotional data and not just adapt to trends of appearance.
3. *Justification:* Quantitative measurements and qualitative examinations will be unified to make a holistic assessment of performance. This is of particular importance when dealing with mental health data that may be noisy, lopsided, and even language-diverse, and cannot be described by a single measurement to gauge efficacy.

Explainability and Interpretability

1. *Gradient × embedding method:* To be transparent, the model will contain explainability methods: The importance scores at the token level will be calculated by multiplying the gradient of the predicted classification with respect to the token embeddings by the token embeddings themselves. This was done to establish which words or phrases contributed the most to a given prediction. As an illustration, in the sentence I feel so empty and hopeless, words such as empty and hopeless would be the words that would have the highest scores in terms of importance.
2. *Visualization:* Horizontal bar charts are used to visualize token-level contributions, the analysis of which makes it easier to understand model reasoning. These are explanatory visualizations for clinicians, researchers, and stakeholders.
3. *Justification:* Mental health applications depend on explainability. Misclassifications in this case might have grave consequences, unlike in the case of generic sentiment analysis. Open descriptions create trust, responsibility, and acceptance among mental health practitioners who can use AI predictions to supplement clinical judgment.

Embedding Visualization

1. *t-SNE analysis*: t-distributed stochastic neighbor embedding (t-SNE) is used to embed high-dimensional BERT embeddings into a visualization space of 2D, to investigate the semantic structure of the learned embeddings.
2. *Purpose:* This will enable researchers to study the clustering tendencies of various types of emotions. Ideally, the embedding of similar emotional classes (e.g., sadness and loneliness) is close to each other, and other emotions (e.g., joy vs. anger) do not depend on one another.
3. *Justification:* Integrating visualization provides an easy visual insight into the model behavior and is also supplementary to quantitative measures. This enables the determination of overlapping classes, semantic separability, and areas where the model might fail, and future refinements in preprocessing or architecture.

RESULTS AND DISCUSSION

To comprehensively understand the performance of our multi-CNN-BERT model, it is worthwhile comparing it with past research in this area. The results of our model are compared with those of previous models in Table 1, with the improvements and advantages that make our approach better highlighted. Past models have frequently been troubled by noisy, user-made text, particularly in mental health forums, where the expression of emotion is delicate and diverse. All we need to do is add BERT embeddings to CNN layers with multiple filters, and our model can attach importance to the global context and local emotional leaves to a greater extent.

Quantitative Evaluation

The multi-CNN-BERT model was tested on the test data to evaluate the performance of the model in multi-class emotion classification [11, 12]. The model achieved an accuracy of 92, an average F1-score of 0.92, a precision of 0.92, and a recall of 0.92. The metrics above suggest that the model can reliably differentiate among different categories of emotions in noisy forum text (Figures 1 and 2).

Table 1. Comparative analysis of AI approaches for sentiment and emotion.

Ref. Year	AI models used	Accuracy	Precision	Recall	F1-score
[1], 2023	Random forest, XGBoost (with GPT embedding)	Random forest: 0.69, XGB: 0.68	Random forest: 0.63, XGB: 0.71	Random forest: 0.61, XGB: 0.70	Random forest: 0.62, XGB: 0.71
[13], 2024	LSTM, BiLSTM, CNN, GRU with Word2Vec and One-Hot Encoding	LSTM: 53.1%, BiLSTM: 54.5%, CNN: 54%, GRU: 39.7%	~48-49% (avg)	61-74% (avg)	54-58% (avg)
[14], 2023	Support vector machine (SVM), logistic regression (LR), multinomial Naïve Bayes (MNB)	SVM: 86.05%, LR: 86.45%, MNB: 70.50%	SVM: 85.20%, LR: 86.12%, MNB: 80.16%	SVM: 88.26%, LR: 87%, MNB: 79%	SVM: 86.7%, LR: 86.5%, MNB: 79.6%
[15], 2024	Lexicon-based, machine learning, deep learning	Lexicon-based: 65%, Machine Learning: 75%, Deep Learning: 85%	Lexicon-based: 67%, machine learning: 77%, deep learning: 87%	Lexicon-based: 64%, machine learning: 74%, deep learning: 84%	Lexicon-based: 65%, machine learning: 75%, deep learning: 85%
[16], 2024	LSTM, GRU, transformer, RNN	GRU: 78.02%, LSTM: 79.46%, RNN: 79.32%	GRU: 78.37%, LSTM: 80.62%, RNN: 82.12%	GRU: 78.94%, LSTM: 79.64%, RNN: 79.92%	GRU: 78.65%, LSTM: 80.09%, RNN: 80.76%
[17], 2020	BERT + BiLSTM	72.64%	0.74	0.73	0.73
Proposed model	BERT embeddings + multi-channel CNN with residual connections; gradient × embedding explainability	92%	0.92	0.92	0.92

```
Test Accuracy: 0.92

Classification Report:
              precision    recall  f1-score   support

       anger       0.92      0.90      0.91       303
        fear       0.89      0.83      0.86       224
         joy       0.94      0.95      0.94       649
        love       0.88      0.79      0.83       172
         sad       0.97      0.96      0.96       581
     suprise       0.67      0.97      0.79        71

    accuracy                           0.92      2000
   macro avg       0.88      0.90      0.88      2000
weighted avg       0.92      0.92      0.92      2000
```

Figure 1. Classification report showing precision, recall, F1-score, and support for each emotion class.

Confusion Matrix Analysis

High accuracy in the classification of emotions such as joy and sadness is possible because of unique lexical cues (e.g., "happy", "depressed"). The largest number of misclassifications was in the fear/anger categories, probably due to the overlap of language patterns in posts on mental health (e.g., "I am tense and frustrated") (Figure 3).

Interpretation: With a mixture of BERT embeddings and multi-filter CNN layers, the model can focus on the contextual meaning of words as well as local n-gram patterns, which is necessary to identify subtle emotional cues. This yields a better classification than standard BERT fine-tuning methods, which only use global sentence embeddings.

t-SNE Embedding Visualization

To evaluate the extent to which the model embeddings reflect semantic differences, we plotted the BERT-generated embeddings in t-SNE.

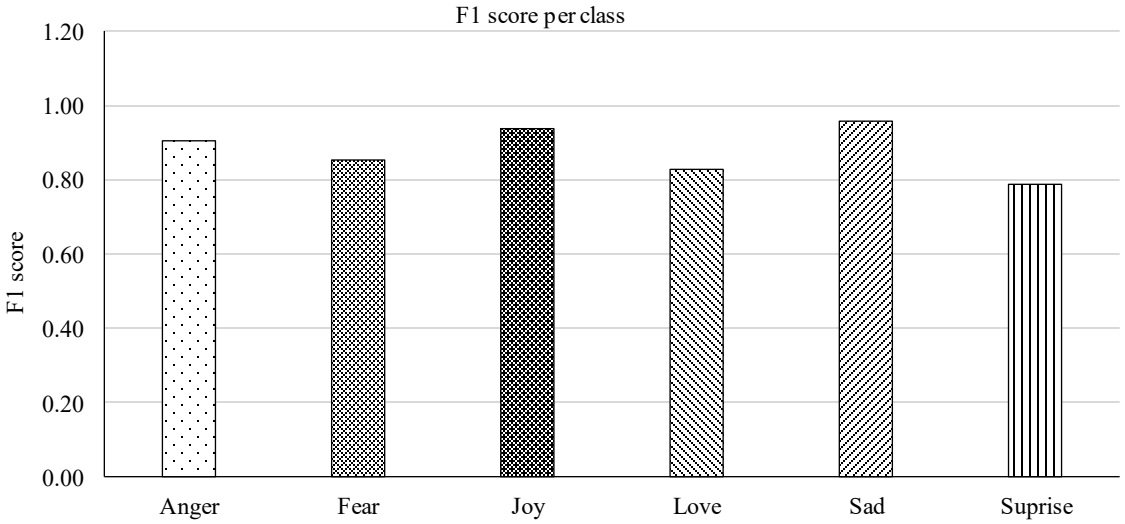

Figure 2. Comparison of F1-scores for different emotions.

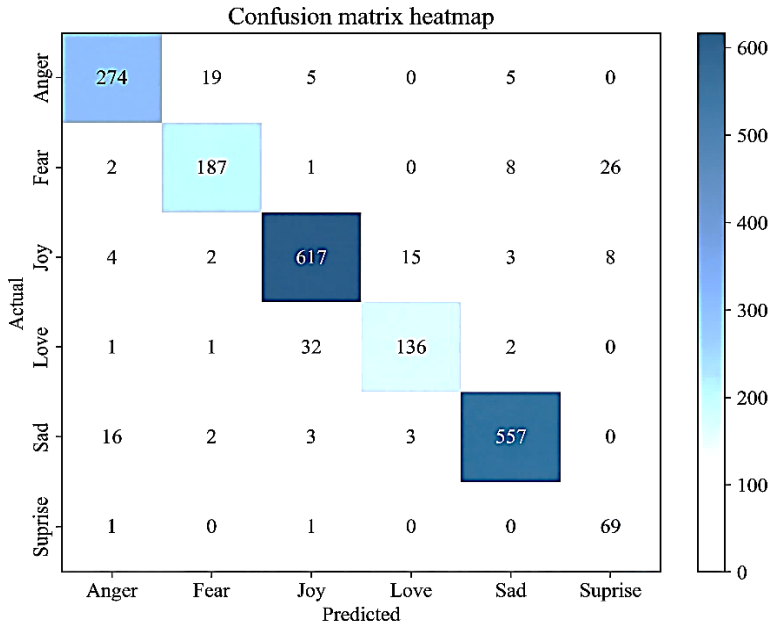

Figure 3. Visualization of the confusion matrix for the emotion classification model.

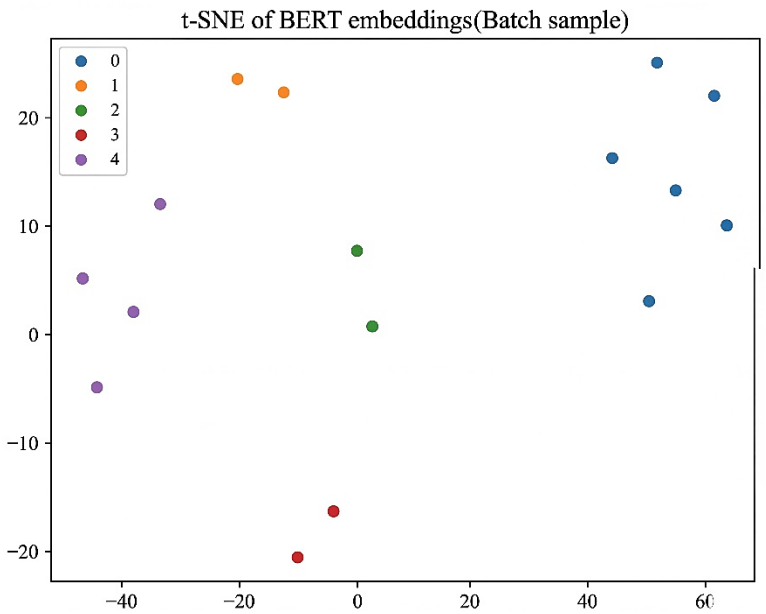

Figure 4. t-SNE-based projection of high-dimensional BERT features.

There is a clear distinction between emotions such as joy, love, and sadness, which means that the model learns specific representations of emotions. Fear and anger overlap, consistent with the confusion found in the quantitative assessment, and this is also an indication of semantic similarity between these emotions (Figure 4).

Interpretation

Visualization using t-SNE confirms that the emotional structure is retained in the embedding space of the model. The fact that these clusters overlap indicates that it is possible to improve them, such as with contextual data augmentation or multi-task learning, to separate semantically similar emotions.

t-SNE Embedding Visualization

The findings indicate that multi-CNN-BERT is both highly predictive and interpretable, which is why the model can be used in real-world mental health settings:

- *Monitoring through forums:* Automatically getting to know the emotional inclination in the postings of the user.
- *Early intervention:* There should be a way of identifying users at risk of unhappiness or negative behavior.
- *Research:* The analysis of emotional patterns is a field of study in psychology and social behavior.

The model may be very intuitive and transparent because mental health workers can identify with the predictions and have confidence in the results; that is, its interventions are founded on useful information and not the mysterious drawings of an algorithm. The accuracy and interpretability of such a degree is necessary in sensitive spheres where AI decisions can potentially have a pragmatic impact.

CONCLUSION

To achieve good predictive performance and explainability, in this study, we proposed an explainable sentiment mining model to categorize the emotions in mental health forums using BERT embeddings, multi-filter CNNs, and residual blocks. The model also helped to capture regular cases of emotional expression and distinguish between the classes of joy, sadness, anger, fear, love, and surprise, and the gradient × embedding explainability option showed the tokens that made the most predictions, which is necessary when the model is applied to mental health. The visualization

embedded by t-SNE also indicated the existence of semantically meaningful clustering, which implied that the model could learn meaningful representations. Not only does the framework offer an effective tool to determine the presence of emotions in an online mental health discussion, but it also enables the predictions to be interpreted and justified, which enhances the trustworthiness of both clinicians and researchers. Future directions can build on the diversity of the datasets, use multimodal cues, and test more robust explainability techniques, which further justify the model as useful in mental health monitoring, early detection of emotional distress, and studies of online emotional patterns.

REFERENCES

1. Abdullah M, Negied N. Detection and prediction of future mental disorder from social media data using machine learning, ensemble learning, and large language models. IEEE Access. 2024;12:120553–120569. doi:10.1109/ACCESS.2024.3406469.
2. Singh J, Sharma G. Sentiment analysis study of human thoughts using machine learning techniques. 2023 International Conference on Disruptive Technologies (ICDT), Greater Noida, India. 2023. p. 776–785. doi:10.1109/ICDT57929.2023.10150917.
3. Brynielsson J, Johansson F, Jonsson C, Westling A. Emotion classification of social media posts for estimating people's reactions to communicated alert messages during crises. Secur Inform. 2014;3:7. doi:10.1186/s13388-014-0007-3.
4. Lövheim H. A new three-dimensional model for emotions and monoamine neurotransmitters. Med Hypotheses. 2012;78(2):341–348. doi:10.1016/j.mehy.2011.11.016. PMID: 22153577.
5. Sasidhar TT, B P, P SK. Emotion detection in Hinglish (Hindi+English) code-mixed social media text. Procedia Comput Sci. 2020;171:1346–1352. doi:10.1016/j.procs.2020.04.144.
6. Tripto NI, Ali ME. Detecting multilabel sentiment and emotions from Bangla YouTube comments. 2018 International Conference on Bangla Speech and Language Processing (ICBSLP), Sylhet, Bangladesh. 2018. p. 1–6. doi:10.1109/ICBSLP.2018.8554875.
7. Batbaatar E, Li M, Ryu KH. Semantic-emotion neural network for emotion recognition from text. IEEE Access. 2019;7:111866–111878. doi:10.1109/ACCESS.2019.2934529.
8. Li D, Rzepka R, Ptaszynski M, Araki K. HEMOS: A novel deep learning-based fine-grained humor detecting method for sentiment analysis of social media. Inf Process Manag. 2020;57(6):102290. doi:10.1016/j.ipm.2020.102290.
9. Ren Z, Shen Q, Diao X, Xu H. A sentiment-aware deep learning approach for personality detection from text. Inf Process Manag. 2021;58(3):102532. doi:10.1016/j.ipm.2021.102532.
10. Pattun G, Kumar P. Emotion classification using generative pre-trained embedding and machine learning. 2023 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT), San Salvador, El Salvador. 2023. p. 1–6. doi:10.1109/ICMLANT59547.2023.10372980.
11. Bayu Y, Tegegn T. Multi-label emotion classification on social media comments using deep learning [preprint]. Research Square. 2024. doi:10.21203/rs.3.rs-4431629/v1.
12. Kaushik B, Sharma A, Chadha A, Sharma R. Machine learning model for sentiment analysis on mental health issues. 2023 15th International Conference on Computer and Automation Engineering (ICCAE), Sydney, Australia. 2023. p. 21–25. doi:10.1109/ICCAE56788.2023.10111148.
13. Choudhary DK, Gupta A, Singh S, Abhinav T, Agrawal T. Sentiment analysis for depression detection using artificial intelligence. 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT), Vellore, India. 2024. p. 1–5. doi:10.1109/AIIoT58432.2024.10574749.
14. Chaithra IV, Samanvaya KJ. Text-based emotion recognition using deep learning. 2024 Second International Conference on Advances in Information Technology (ICAIT), Chikkamagaluru, Karnataka, India. 2024. p. 1–7. doi:10.1109/ICAIT61638.2024.10690782.
15. Adoma AF, Henry NM, Chen W, Rubungo Andre N. Recognizing emotions from texts using a BERT-based approach. 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China. 2020. p. 62–66. doi:10.1109/ICCWAMTIP51612.2020.9317523.

16. Sharma T, Diwakar M, Singh P, Lamba S, Kumar P, Joshi K. Emotion analysis for predicting the emotion labels using machine learning approaches. 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Dehradun, India. 2021. p. 1–6. doi:10.1109/UPCON52273.2021.9667562.
17. Shyang YK, Yan JLS. A text augmentation approach using similarity measures based on neural sentence embeddings for emotion classification on microblogs. 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia. 2020. p. 1–6. doi:10.1109/IICAIET49801.2020.9257826.

