Minimax Theory and its Applications

Volume 11 (2026), No. 5s, 352-374

# Capsule-Attention Enhanced Emotion Recognition for Multimodal Depression Detection

**S Divya Nair[1], M Lalithambigai[2*]**

[1]PhD Scholar, Department of Computer Science, Sri Krishna Adithya College of Arts & Science (Co-Education), Coimbatore, Tamil Nadu, India
[2]Assistant Professor, Department of Computer Science, Sri Krishna Adithya College of Arts & Science (Co-Education), Coimbatore, Tamil Nadu, India
*Corresponding author: Email: divyasukumaran1989@gmail.com

**Abstract**
The problem of depression and mental illnesses has turned into an issue in the entire world, and social media is an excellent source of emotional and psychological information about early diagnosis. Clinical evaluations that are currently in use tend to be subjective, time-consuming and incapable of scaling, which is why automated AI-based solutions are sought. The following research presents a proposed framework of Context-Aware Emotional Risk Identification (CAERI) of depression detection and type categorization that is aimed at extracting hierarchical and temporal emotional trends of textual data. Its main aims are to identify depression accurately, differentiate between different types of depression and have interpretable results to be used in mental health assessment. The proposed methodology consists of the combination of emotion-sensitive pre-processing, tokenization, Identification, and dual-level feature extraction with word-level and sentence-level embeddings. Capsule Networks model the hierarchical patterns of emotions whereas Temporal Convolutional Networks (TCN) learns sequential dependencies so that the system is able to distinguish between temporary emotions and prolonged depressive responses. Large scale experimentation was performed in Python on a social media mental health benchmark dataset, and the CAERI framework performed better than the current approaches in Machine Learning (ML) and Deep Learning (DL). The model has obtained a high accuracy of 95.65 (%) which proves to have better feature representation ability. In general, the findings confirm the usefulness of CAERI in strong and interpretable detection of depression. The framework will be expanded to multimodal and real-time personalized intervention systems in the future.

**Keywords:** Capsule Networks, Depression Detection, Emotional Feature Extraction, Mental Health, Temporal Convolutional Networks, Text Preprocessing

## 1. Introduction

Depression is a prevalent mental illness that has tremendous effects on quality of life, productivity and social well-being. Early prevention and intervention are essential in avoiding extreme consequences, suicidal behavior being one of them. Conventional clinical measurements tend to be subjective, hectic, and not scalable, leading to the researcher to consider automated AI-based methods of monitoring mental health. Recent research has also shown that explainable Machine Learning (ML) can be used to predict depression by giving interpretable insights into the factors that contribute to it and the decision-making process [1]. Social media data can also be processed using Deep Learning (DL) to identify cases of depression early on, identifying trends in user-generated media that can be related to psychological distress [2]. Moreover, multimodal systems, which combine Electroencephalography (EEG) signals with demographic data, have been demonstrated to be effective in strengthening and accuracy of detectors [3], and hybrid DL systems using self-attention have been found to be more effective in processing social networking data to detect depression [4]. Transformer-based models that include explainable AI also make it possible to evaluate severity based on social network interactions with nuance [5], and hybrid DL designs have been shown to be useful in detecting depression in a variety of datasets [6].

Regardless of such progress, there are still difficulties in the development of complex emotional threads, time-dependent issues, and multimodal emotions to conduct a complete evaluation of depression. Recent researches have investigated multimodal data fusion in enhancing predictability [7], multi-instance learning models to optimize clinical interview analysis [8], and ensemble DL in identifying suicidal ideation on social media [9]. Combined ML, DL, language models and explainable AI have also increased the accuracy of detection and explain ability further [10]. Based on these understandings, the current research propose the Context-Aware Emotional Risk Identification(CAERI) framework that generalizes hierarchical and temporal emotional trends

based on the text, allowing one to effectively categorize depression types and be able to evaluate mental health risks through clarity and interpretable results. This approach overcomes the constraints of the previous methods through a combination of pre-processing, dual level feature extraction, hierarchical and capsule modeling, and temporal convolution to the detection of depression in a robust and scalable manner.

- Acquires a CAERI system of true depression recognition in social media text.
- Combines hierarchical and temporal emotional modelling to explicitly classify depression-type.
- Applying emotion sensitive preprocessing and feature selection to enhance performance of model.
- Exhibits high accuracy, precision, recall, and F1-score values than the current ML and DL methods.
- Driven by the necessity to have automated, scalable, and context-conscious mental health monitoring based on AI.

**Organization:** Section 1 is the introduction of the research context, problem definition, motivation, and main contributions of the research. Section 2 is the background research of the work on AI-based depression and mental health detection. Section 3 is the discussion about the proposed CAERI methodology, such as the overall architecture, algorithm workflow, mathematical formulations, and pseudocode. Section 4 is about the experimental results and performance evaluation and comparative analysis. Lastly, the research is concluded in Section 5 and presents future directions of research.

## 2. Background Study

Alsharif et al. (2024) [11] monitored behavior patterns and predictive Artificial Intelligence (AI) models to identify and enhance early mental health evaluation based on a text analysis. In this research, they employed Support Vector Machine (SVM), Random Forest (RF) and Neural Network algorithms of Machine Learning (ML) and Deep Learning (DL), which trained on text features. Nevertheless, the research did not dwell much on the issue of diagnosis of depression using multimodal emotional risk analysis but in relation to the diagnosis of the ADHD. The results indicated that the diagnosis was more accurate than the current methods of diagnosis.

Zohuri et al. (2020) [12] investigated the possibility of utilizing AI to enhance the mental health monitoring systems through mood analysis, depression and suicide risk management. These authours research revealed the threat of developing mental illnesses on the basis of behavioral information based on predictive analytics via AI, emotional pattern recognition, and decision-support models. However, the research lacked real-time intervention functionalities and person-specifically prescribed systems and results demonstrated that AI Perform a major role in early depressions and suicide risk identification.

Thekkekara et al. (2025) [13] introduced an Attention-Based Bidirectional Encoder Representations from Transformers- Convolutional Neural Network- Bidirectional Long Short-Term Memory (BERT- CNN - BiLSTM) system to learn how to recognize depression using emojis in social media text, where BERT is the contextual representations, CNN is the feature extractors, and BiLSTM is the sequence models, to respond to emoji-induced cues of depressions. It was shown that the research has not conducted the analysis of non-textual signals, was limited in the case of multimodal integration, and achieved higher than baseline performance in depression detection conditions with the use of the DL model.

Hemtanon et al. (2022) [14] developed a proactive mechanism of identifying depression that relied on Facebook text content and user behavior patterns to identify the initial signs of depression. Natural Language Processing (NLP), sentiment and MLclassifiers are used as the methods of research by these authours to analyse the expression of emotions and behavioral cues. However, the research focused on Facebook data only and was unable to extrapolate across the platforms and make adaptive interventions and the results showed that it was possible to identify depressive behavior through the analysis of social media.

Ullah et al. (2020) [15] proposed a combination of textual and emoticon-powered emotion expression algorithm to improve on the classification of emotional states. These were NLP, emoticon feature extraction and ML models to derive sentiment polarity. However, the methodology failed to encompass contextual emotional risk testing and adaptive training of features using the deep learning; the results revealed that the sentiment system performed more favorably as compared to the current text-only based systems.

Rongali et al. (2025) [16] studied how AI can be applied to healthcare to detect medical threats, namely, the identification of psychological and behavioral health threats. The identification of threats was done using AI-based predictive models, anomaly detection, and smart health monitoring systems. However, no direct research was conducted on the problem of predicting the risk of depression based on the analysis of emotional contexts, and the results showed that AI increased the score in terms of the precision of healthcare threats identification and the risk early identification.

Noor et al. (2025) [17] offered a stack based ensemble-based DL technique in the detection of depression on textual data through sentiment analysis. This research had previously embraced the DL model using the ensemble learning models as LSTM, CNN and stacking classifiers to obtain improved outcomes. The approach, however, lacked adaptive customized intervention responses after the identification of the depression and the findings displayed that the approach possessed a higher detection power as compared to the single DL models.

The research by Guo et al. (2021) [18] analyzed the behavioral and environment-related symptoms of psychological distress and depression during the COVID19 pandemic. In this study, the statistical analysis, psychological assessment tools and the data analysis of behaviors have been implemented to determine the risk factors of depression. However, the research did not present AI-based predictive research of depression and found that behavioral stresses had strong relations with the presence of more symptoms of depression.

Selim et al. (2025) [19] developed an emotional text-synthesis model of speech generation, which was in hybrid DL and artificial hummingbird optimization. The hybrid DL models, modeling and optimization algorithms of emotional features were used to enhance the quality of speech emotion. Yet, this research was focused on the synthesis of emotional speech and not on depression or identification of the emotional risk, and the results confirmed that there was a greater accuracy of the synthesis of emotional speech with the speech synthesis.

The research by Haque et al. (2022) [20] is the comparative research of the suicidal ideation detection using the NLP, ML, and DL. This research elicited NLP features, supervised learning models and Deep Neural Networks (DNN) to identify suicidal intent on text. However, adaptive recommendations of mental health intervention were not included in the research and the findings showed that the DL models were more precise than the current ML models after the detection.

**Table 1:** Background Research Background Study on Depression Detection with ML and DL and Multimodal Approaches

| **Authors & Year** | **Concept** | **Methods Used** | **Research Gap** | **Limitations** | **Results** |
|---|---|---|---|---|---|
| Sang et al., 2025[21] | Depression decoding from different brain regions | Hybrid ML models on brain-region features | Lacks emotional and behavioral context integration | Brain-region dependency, limited generalization | Improved regional-level depression prediction |
| Kshirsagar et al., 2025[22] | Early depression detection using sarcastic text | Multiple deep learning models (HEDL framework) | Sarcasm handled, but broader emotional dynamics ignored | High computational cost, domain dependency | Enhanced sarcasm-based depression detection |
| Ksibi et al., 2023[23] | EEG-based depression detection | Multiple ML classifiers on EEG signals | Does not model temporal emotional transitions | Sensitive to noise and EEG artifacts | Competitive accuracy on EEG datasets |
| Aleem et al., 2022 | Survey on ML for depression diagnosis | Traditional ML and DL algorithms | No unified adaptive or context-aware framework | Review-based, lacks experimental validation | Identified trends and challenges in depression ML |
| Haq et al., 2024 | Adaptive ML for early depression detection | Adaptive machine learning models | Emotional semantics and hierarchy not considered | Limited scalability across datasets | Improved early-stage depression recognition |
| Cruz et al., 2022 | Social media usage and depression | Statistical and behavioral analysis | No automated intelligent detection model | Observational study, not predictive | Established correlation between social media use and depression |
| Mahdi & Ahmed, | MDD diagnosis using | AI-based EEG PSD image | Focused only on frequency-domain | Lacks multimodal and | High accuracy in EEG-based |

| | | | | | |
|---|---|---|---|---|---|
| 2022 | EEG PSD imaging | classification | EEG features | contextual modeling | MDD diagnosis |
| Yousufi et al., 2024 | Multimodal EEG–audio depression recognition | Modified DenseNet121 with multimodal fusion | Limited emotional hierarchy modeling | High computational complexity | Improved MDD recognition accuracy |
| Lim et al., 2025 | Lightweight cross-modality depression detection | Cross-modal lightweight DL models | Reduced depth affects deep emotional understanding | Trade-off between speed and accuracy | Efficient and fast depression detection |
| Li et al., 2026 | Dynamic multimodal depression detection | Multi-scale time-series modeling, multimodal fusion | Requires large-scale labeled multimodal data | High training complexity | Superior performance in multimodal depression detection |

The above Table 1 provides a concise description of the recent studies on depression detection with various AI tools of EEG, text in social media, and multimodal data. It comprises the authors, year, reference and even the concept, methods to be employed, gaps in the research, limitations and the crucial results. The table will use in establishing trends, challenges and opportunities of the existing research and this will provide a clear background when it comes to developing advanced frameworks of depression detection. Having summed up a number of approaches within a single field, it offers the comparative perspective of accuracy, methodologies and the efficacy of different models.

## 3. Proposed Methodology

In this study, the proposed algorithm based on CAERI, its overall workflow, the general system structure as well as its design gets described. The methodology is also systematic in the way that every processing step such as processing of emotional texts has been narrated in such a way that it can be understood and interpreted with ease. The transformations and decision making procedures in the context of depression identification is presented in a manner that is mathematically provided. In the proposed algorithm, also, there will be structured pseudocode that describes how the algorithm was step by step implement.

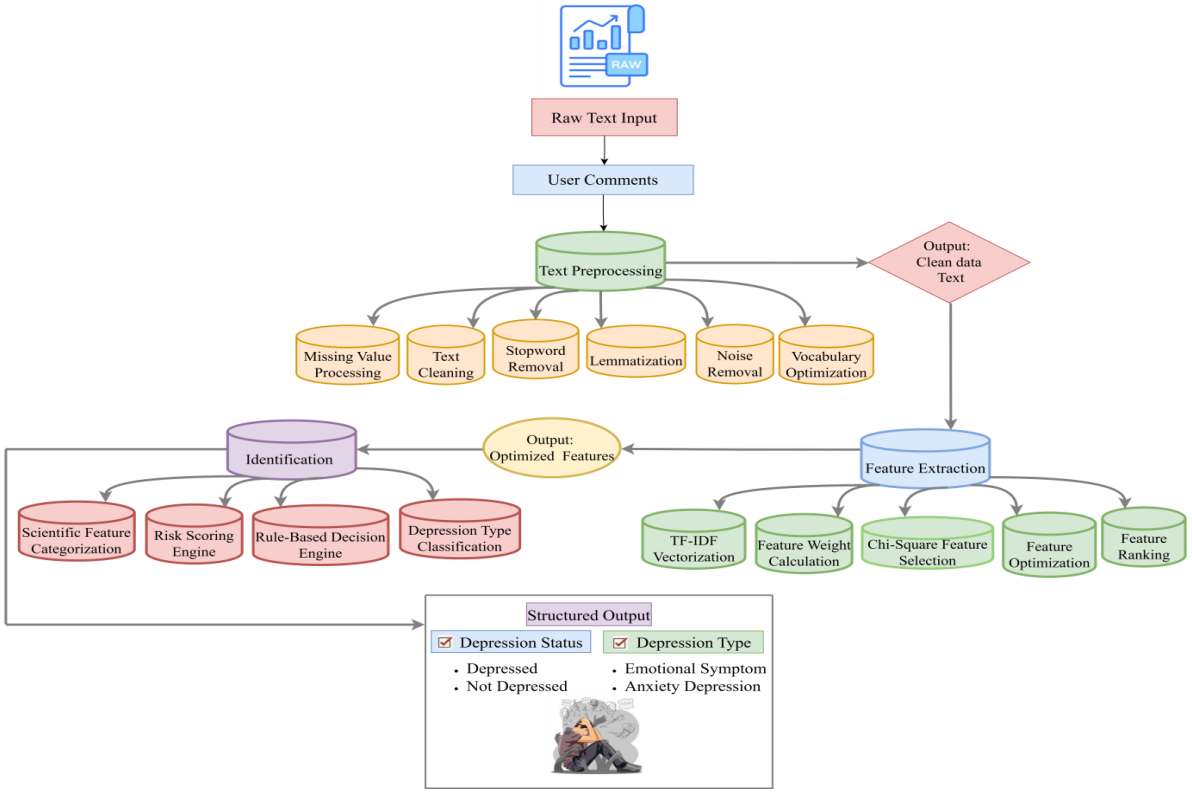


**Figure 1:** The CAERI Depression Identification Framework Architecture

The above Figure 1 indicates the workflow of the proposed model of CAERI of detecting depression and classifying the type thereof on an end-to-end basis. Then an initial layer of preprocessing is fed into the raw emotional text data to eliminate noise and even the linguistic disparity to create clean emotional text. In the second stage, the discriminative emotional characteristics are then filtered and refined by use of statistical and weighting techniques so as to render into more relevant. Finally, the scientific identification layer is based on the risk scoring and rule based reasoning to identify the existence of depression and the exact brand of depression to give accurate and understandable mental health results.

### 3.1 Dataset Description

Sentiment Analysis of Mental Health dataset https://www.kaggle.com/datasets/suchintikasarkar/sentiment-analysis-for-mental-health presented on Kaggle, includes textual data, which is gathered within the social media networks, and each text record depicts emotional and mental state of the respondents, i.e., depression, stress, anxiety, or nothing. It consists of the types of sentiments, which are tagged and utilize the researchers to train and test both the ML and DL models to identify mental illnesses based on language patterns. The development of automated mental health monitoring systems, early depression detection system, and sentiment systems in AI-based healthcare system are all trending in the acquisition of the information.

### 3.2 Pre-processing

The preprocessing of the text is one of the most important activities of the mental-health analysis since the texts left by users are generally noisy, informal and colored by emotions include abbreviations and spelling alternatives, emojis, and wrong constructions. Otherwise, in cases where this noise is not reduced, it is mask smaller emotional indications and have an adverse effect on deep-model learning. The text preprocessing CAERI is directed towards the cleaning of the text and the preservation of emotional semantics, necessary in the processing of context of depression. Normalization gives the linguistic consistency of diverse styles of writing, and tokenization gives the manifestation of feelings to be interpreted with fine-grained precision. The stop-words are processed with much care, to eliminate the non-informative words, but not the psychologically important words. Broadly, the step transforms raw mental-health text into a normal emotion-sensitive representation, which can be deep-context modelled.

$$T_n = f_{norm}(T_r) \tag{1}$$

In equation 1 means that $T_r$is a raw input text obtained by user comments, $T_n$is normalization text of lowercasing, noise elimination, and spelling standardization, and $f_{norm}$ is a normalization function. This formula converts raw text that contains a lot of noise into a standard format to minimize the variability that occurs during informal writing.

$$\mathrm{W} = \{\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_\mathrm{n}\} = \mathrm{f}_{\mathrm{tok}}(\mathrm{T}_\mathrm{n}) \tag{2}$$

In equation 2 represents that $T_n$ is normalized text, $W$ is collection of extracted tokens (words/subwords), $w_i$Individual token,$f_{tok}$ and is tokenization function. This stage breaks down normalized text into extractable units, which make it possible to extract word-based emotional features.

$$W_l = \{l_1, l_2, \ldots, l_3\} = f_{lem}(W) \tag{3}$$

In equation 3, $W$ represents tokenized words, $W_l$ is lemmatized word set, $l_i$ is the root form of a token $w_i$and $f_{lem}$ represents a lemmatization function. Lemmatization is used to map words onto the base forms in order to enable CAERI to identify semantically similar emotional expressions at all times.

### 3.3 Feature Extraction

After the preprocessing, CAERI will be applied to isolate emotion-centric features to remove fine-grained and holistic emotion links in mental-health text. To create the system that identifies subtle features like stress, fatigue, or hopelessness, word-level embeddings are initially created to model the semantic and emotional meaning of individual tokens. Embeddings at the sentence-level are then calculated to reflect the emotional tone of the whole comment and the contextual intent of the entire comment as opposed to individual words. The contextual composition processes combine word interactions to reflect semantic dependencies which are complex to reflect actual emotional expression. Hierarchical emotional organization is maintained in this dual level representation where the lower level word emotion can make a significant contribution to higher level sentence meaning. CAERI provides a powerful and emotionally aware feature space to be used in downstream by jointly learning the word-level accuracy and sentence-level context.

$$e_i = f_{emb}(w_i) \tag{4}$$

The equation 4 is that $w_i$lemmatized word $i^{th}$ token, $e_i$dense embedding vector of word $w_i$, and $f_{emb}$ is a word embedding function. The words are mapped to a dense and representational semantic and emotive meaning.

$$E = [e_1, e_2, \dots, e_m] \quad (5)$$

The equation 5 shows that $E$ is word embedding matrix, $m$ is a word count in the sentenceand $e_i$ is embedding of the $i^{th}$word. Such a matrix will contain the full text emotional data of the text on word level.

$$h_i = f_{ctx}(E, i) \quad (6)$$

In equation 6 denotes that $h_i$ is a context-sensitive encoding of the $i^{th}$word,$f_{ctx}$ context a modeling of context, $E$ is word embedding matrix. The embedding of words is optimized with the use of surrounding words to obtain emotional dependencies in the context of the word.

$$s = f_{sent}(h_1, h_2, \dots h_m) \quad (7)$$

It is shown in equation 7 that $s$ is a sentence-level embedding vector,$h_i$ is contextual word representations, $f_{sent}$ terms of sentence aggregation. Emotional cues on a word-level are combined to create a general sentence emotion representation.

$$F_c = [E + s] \quad (8)$$

Equation 8 means that $F_c$ combines emotional feature representation, $+$ is the concatenation operator in the features, $E$ is an word-level feature matrix, $s$is is the sentence-level embedding. This combination maintains the finer and inclusive emotional data.

$$H = f_{hier}(F_c) \quad (9)$$

In equation 9, $H$ represents hierarchically encoded emotional features, and $f_{hier}$shows a hierarchically encoding function. Emotional elements are arranged in hierarchies to indicate nearly the connection of words and sentence-level emotions.

$$Z = f_{final}(H) \quad (10)$$

In equation 10 represents that$Z$ is final extracted emotional feature vector, $f_{final}$ is a feature refinement function. The feature vector is a fined result to be used as a solid input in respective and risk forecasting.

**3.4 Identification**

In this step, CAERI converts the extracted emotional features into explicit depression-type categories modeling hierarchical and temporal emotional features. Capsule Networks are utilized to obtain part-whole emotional relations so that the system can know how the low-level emotions like stress or frustration can lead to the higher levels of depression. In comparison to the existing pooling, capsules preserve spatial and semantic dependencies and increase the interpretability and robustness. TCNs are then used to encode sequential emotional interactions within sentences or time-varying comments, and emotional dynamics and sustainability. This time argument is critical in differentiating between the passing sadness and long-term emotional pain. Using hierarchy modeling and time modeling, CAERI effectively categorizes the types of depression as work related, situational, and emotional distress in real-time context.

$$u_i = f_{cap}(Z) \quad (11)$$

In equation 11 is where $Z$a resultant emotional feature vector is, $u_i$ is an output vector of the $i^{th}$ primary capsule, $f_{cap}$is capsule feature. Emotional features are converted into capsule vectors of emotion low-level patterns.

$$v_j = squash\left(\sum_i c_{ij} u_i\right) \quad (12)$$

In equation 12 means that, $v_j$ is a product of the $j^{th}$ higher level capsule, $c_{ij}$ is routing coefficient between capsules $i$ and $j$, $u_i$ is a primary capsule outputs, $squash$ is a non-linear normalization function. Dynamic routing combines emotional capsules associated with an emotion in order to create a representation of a higher level of depression.

$$C = [v_1, v_2, \dots, v_n] \quad (13)$$

In equation 13 is an indication that $C$ is an emotional representation in the form of a capsule, $v_j$ is an elevated level of capsules and $n$ is a count of capsules that represent depression. Every capsule will be associated with a depression-type feature representation.

$$T_t = f_{TCN}(C_{t-k:t}) \quad (14)$$

In equation 14, $T_t$ is temporal feature representation at time $t$,$C_{t-k:t}$ is a capsule outputs within a time window, $k$ is size of the temporal receptive field and $f_{TCN}$ is a temporal convolution filter. TCN identifies long-range emotional relations between consecutive text inputs.

$$\hat{y} = argmax\left(f_{cls}(T_t)\right) \quad (15)$$

In equation 15 denotes that $\hat{y}$ is a predicted depression type label, $f_{cls}$ is the function, $T_t$ is a temporal emotional features. The likelihood of depression category is chosen depending on emotional and time patterns learned.

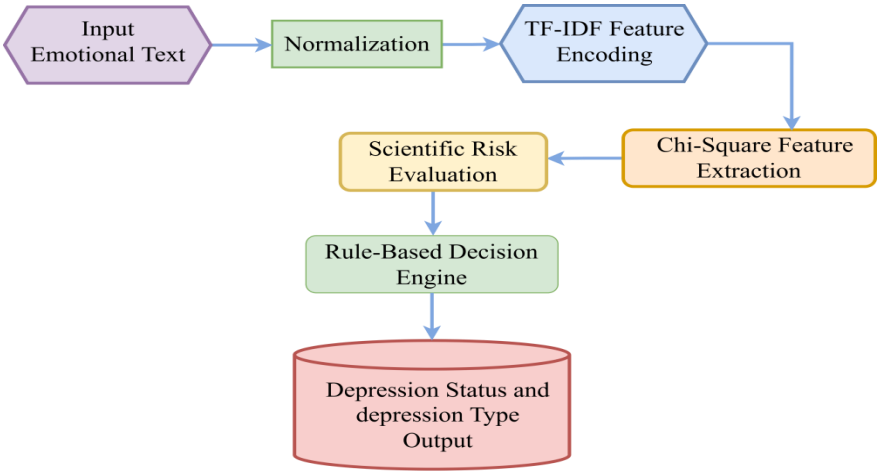


**Figure 2:** Proposed Algorithms Workflow to Depression Status and Type Identify

In above figure 2 demonstrates the overall work process of the proposed algorithm of analyzing depression with emotional text data. The process consists of input emotional text that is normalized and encoded with the use of TF-IDF to obtain term significance and Chi-Square feature extraction to obtain statistically significant features. The resulting features are then assessed by a scientific risk assessment module and decisions are finally arrived at by a rule based decision engine. The workflow will provide the general state of depression and the particular type of depression, which will make the mental health assessment accurate and interpretable.

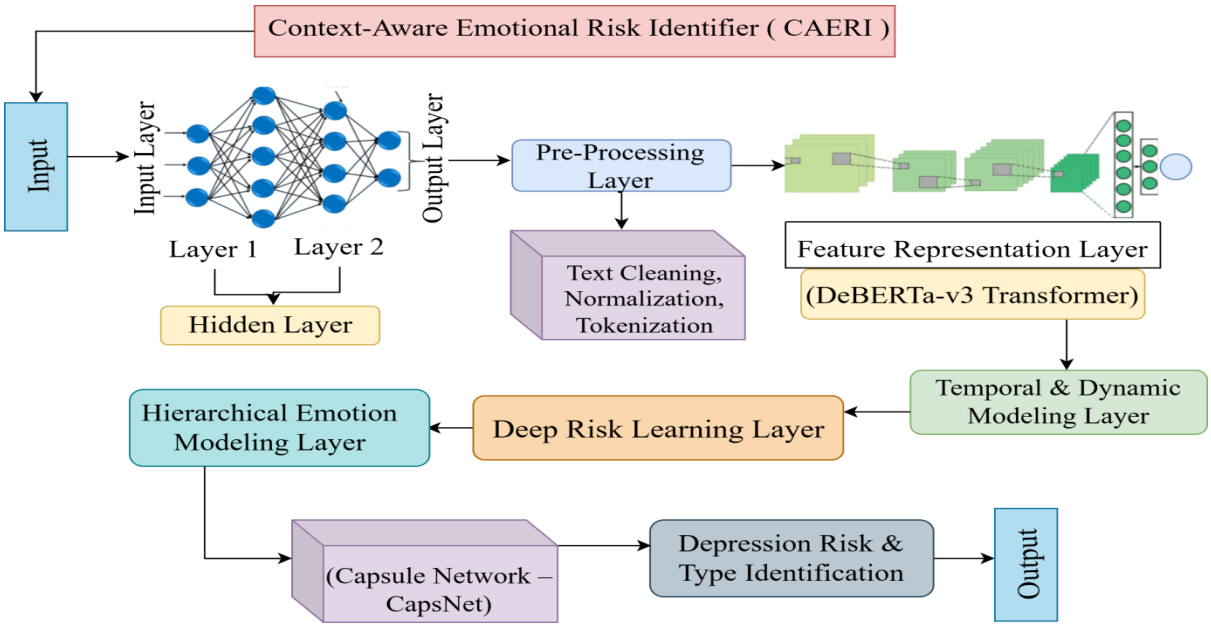


**Figure 3:** Layered Architecture of the Proposed CAERI Algorithm

In the above figure 3, the proposed CAERI architecture is a sequence of layers that takes in input text and first does preprocessing to clean and normalize the data and then does DeBERTa-v3 to extract depression-related contextual features. Neural ODE models the dynamics of depression symptoms over time, Res-DNN learns deep depression risk patterns and CapsNet learns hierarchical depression features, making it accurate in detecting the presence and type of depression.

**Algorithm:** Context-Aware Emotional Risk Identification (CAERI)

Input:

D → Emotional text dataset

Each record contains:

statement → emotional text

status → emotion label

Begin

1. Load Dataset

Read dataset D from file

Extract text column → statement

Extract label column → status

2. Handle Missing Values

for each record r in D do

IF r.statement is Null Then

Replace with "unknown emotional state"

ENDIF

ENDFOR

Remove remaining NULL records

3. Text Cleaning And Preprocessing

Initialize stopword list SW

Initialize lemmatizer L

for each text t in D do

Convert t to lowercase

Remove URLs

Remove special characters and numbers

Tokenize text into words

Remove stopwords using SW

Apply lemmatization using L

Store cleaned text in Dc

endfor

4. Filter Text Length

```
    for each cleaned text c in Dc DO
        Calculate word length
        if length > maximum threshold then
            Remove c from Dc
        endif
    endfor
5. Label encoding
    Convert status labels into numeric values
    Store encoded labels in Y
6. Feature Extraction Using Tf-Idf
    Initialize TF-IDF vectorizer with parameters:
        max_features = N
        ngram_range = (1,3)
    Transform cleaned dataset Dc into feature matrix F
7. Feature Selection Using Chi-Square
    Apply Chi-square feature selection on F and Y
    Select top K important features
    Store selected features in Fs
8. Depression Risk Identification
    Define strong emotional indicators
    Define moderate emotional indicators
    Define weak emotional indicators
    Initialize counters:
        strong_count ← 0
        moderate_count ← 0
    for each selected feature f in Fs DO
        if f belongs to strong indicators then
            Mark as Depressed
            strong_count ← strong_count + 1
        else if f belongs to moderate indicators then
            Mark as Moderate Risk
            moderate_count ← moderate_count + 1
        else
            Mark as Normal
```

```
        endif
    endfor
9. Final Decision
    if strong_count ≥ 1 THEN
        Overall status ← Clinical Depression
    else if moderate_count ≥ 2 THEN
        Overall status ← Moderate Depression Risk
    else
        Overall status ← Normal Emotional State
    Endif
10. Return Output
    Return cleaned dataset Dc
    Return selected features Fs
    Return depression identification result R
    Return processing and feature metrics PM
End
Output:
    Dc → Cleaned dataset
    Fs → Selected emotional feature set
    R → Depression risk identification result
    PM → Processing efficiency and feature analysis metrics
```

The above CAERI Pseudocode algorithm runs through emotional text data by completing missing data treatment, text cleaning, and feature extraction using TF-IDF and feature selection using Chi-square in determining the most appropriate emotional indicators. Depending on the existence and intensity of primary emotional features, the algorithm identifies the level of depression risk and generates structured emotional risk detection outcomes and the processing efficiency indices.

### 3.5 Ablation and Methodological

Ablation research was done to measure the contribution of each element of the proposed CAERI framework such as feature extraction, hierarchical emotion modelling, and temporal sequence analysis. Methodological justification provides a reason why Capsule Networks were utilized to maintain part-whole emotional relationships and why Temporal Convolutional Networks were utilized to learn sequential dependencies. The performance improvements made by each module as compared to the baseline models indicate the need to have the integrated architecture. This strategy will guarantee that each component will have a meaningful contribution to the better accuracy and strength of depression detection.

## 4. Results and Discussions

The results section provides the description of the experiment outcomes according to the application of the proposed CAERI framework in Python, which is reproducible and scalable due to the fact that the entire preprocessing feature extraction and Identification are realized with the use of Python libraries. The outcomes are a critical analysis of the CAERI performance on working with the emotional text data and the identification of a status and type of a depression. The proposed approach is more effective and stronger and can be utilized more successfully as compared to the existing ML and DL approaches.

### 4.1 Pre-Processing

The preprocessing stage manages to cleanse, standardize and structure unrefined emotional text to receive quality input with which to proceed with appropriate analysis of depression.

| Before Cleaning | After Cleaning |
|---|---|
| oh my gosh | oh gosh |
| trouble sleeping, confused mind, restless heart. All out of tune | trouble sleeping confused mind restless heart tune |
| All wrong, back off dear, forward doubt. Stay in a restless and restless | wrong back dear forward doubt stay restless restless place |
| I've shifted my focus to something else but I'm still worried | shifted focus something else still worried |
| I'm restless and restless, it's been a month now, boy. What do you mean | restless restless month boy mean |

**Figure 4:** CAERI Result of text pre-processing and cleaning

The transformation of the raw emotional text into the normalized representation as shown in the above figure 4 is done at the CAERI stage of pre-processing. The noisy components that may include the stop words, punctuations, repetitions and colloquial expressions are removed and do not affect negatively in semantics of emotion. The resulting clean text results are used as an input to the further extraction of features and the analysis of depression.

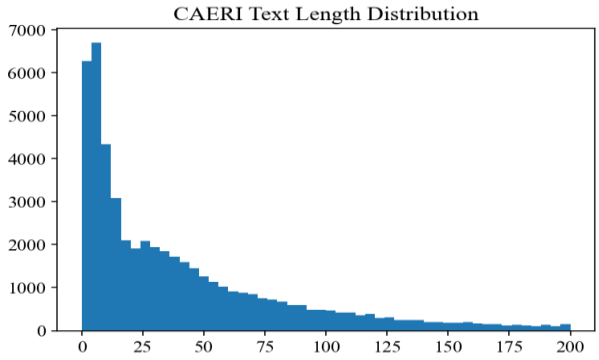


**Figure 5:** Character Length Distribution CAERI Text

In the above Figure 5 above indicates that the length of the text in the CAERI dataset was distributed after pre-processing. It is proposed by the skewed form that long-form expressions and short emotional texts are fewer in number and inform the selection of the preferred sequence length and padding strategy. This type of analysis ensures an efficient training of models with no information on significant emotional factors being overlooked.

**Table 2:** Text Pre-processing Effects on CAERI Dataset

| Step | Before | After |
|---|---|---|
| Total rows | 53043 | 51141 |
| Special characters | Present | Removed |
| Stop words | Present | Removed |

Table 2 above shows the outline of effect of CAERI pre-processing stage to the raw emotional text dataset. Data quality and consistency is greatly improved since most items such as noise such as special characters and stop words are highly reduced. The existence of duplicates is ensured to be non-existent consequently making the data integrity assured. The obtained pre-processed data can be processed to obtain features in a stable manner and label the depression.

**Table 3:** Text Optimization and Dataset Efficiency after CAERI Preprocessing

| Metric | Before | After | Efficiency / Reduction |
|---|---|---|---|
| Vocabulary | 131,811 | 45,640 | 65.37% Reduction |
| Average Text Length | 90.79 | 40.67 | 55.21% Reduction |
| Dataset Rows | 53,043 | 51,141 | 96.41% Retention Rate |

The efficiency of pre-processing step of CAERI in the dataset maximization process are displayed in the table 3 above. A massive reduction in the size of vocabulary and the mean text length are the signs of a successful noise elimination and linguistic normalisation. Generally, the preprocessing ensures that the high quality inputs of small size are made in order to extract the emotional features correctly.

**Table 4:** Comparison of Pre-processing Efficiency between the current methods and the one proposed by CAERI

| Methods / Metric | Vocabulary Reduction Ratio (%) | Text Length Reduction Ratio (%) | Noise Removal Efficiency (%) | Missing Data Processing Efficiency (%) | Dataset Retention Rate (%) |
|---|---|---|---|---|---|
| Logistic Regression (LR) [1] | 28.45 | 24.31 | 30.12 | 82.64 | 92.10 |
| Temporal Convolutional Networks (TCN) [3] | 46.78 | 41.65 | 48.90 | 89.23 | 94.85 |
| Convolutional Neural Networks (CNN) [4] | 50.36 | 45.28 | 52.14 | 90.75 | 95.12 |
| Deep Neural Networks (DNN) [5] | 53.92 | 48.11 | 56.87 | 92.40 | 95.68 |
| Naïve Bayes (NB) [7] | 32.64 | 29.08 | 35.76 | 85.91 | 93.04 |
| Proposed Algorithm (CAERI) | 65.37 | 55.21 | 65.37 | 96.77 | 96.41 |

The table 4 above has made a comparison between the effectiveness of different ML and DL methods in the pre-processing of emotive text data. Vocabulary reduction measures, text reduction length and noise removal efficiency measure are the measures that indicate how each method can use to remove unnecessary and redundant information. The absence of data processing efficiency and data retention rate is an indicator of strength based on the incomplete data processing and retention of useful samples. The CAERI algorithm proposed is superior to the existing ones because it has a high preprocessing efficiency, and does not lose information significantly.

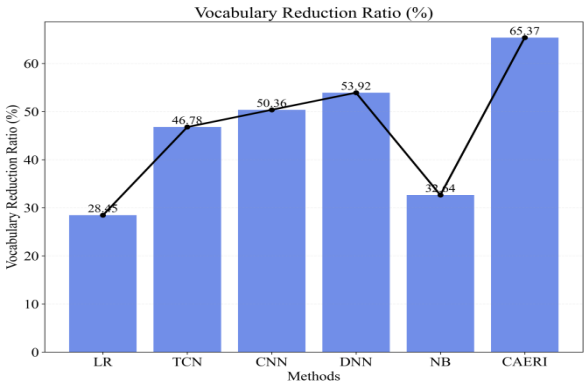


**Figure 6:** Comparison of Algorithms ratio between vocabularies Reduction

The above figures 6 are actually constructed through plotting of ratios of vocabulary reduction of the pre-processing process with assistance of the different algorithms. The x-axis represents the methods and algorithms (LR, TCN, CNN, DNN, NB and CAERI) and the y-axis represents the percentage of vocabulary reduction. The mark is used to denote how much of a decrease was achieved in each technique. The visualization shows that CAERI is the most useful when it comes to decreasing the vocabulary during the preprocessing stage.

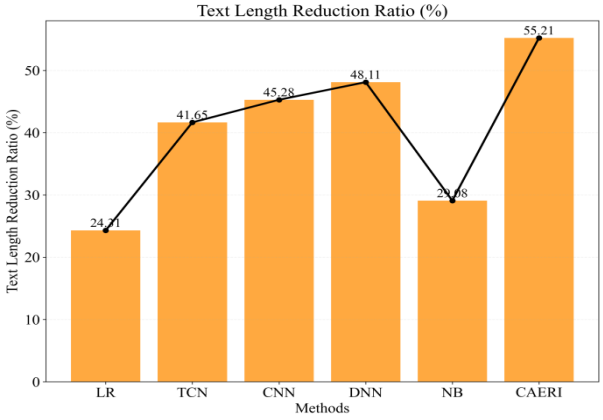


**Figure 7:** Comparison of the Ratio of text length reduction to the length of the algorithm

In above figure 7 indicates how the average length of text in each algorithm reduces as it is pre-processed. The processing before methods are shown as the x-axis and the percentage ratio of the text length reduced as the y. The direction of the curve of the text compression efficiency is shown through the plotting line, whereas the measure of the reduction was emphasized through the filled area. CAERI has the optimum text length reduction implying that it possesses greater text normalization.

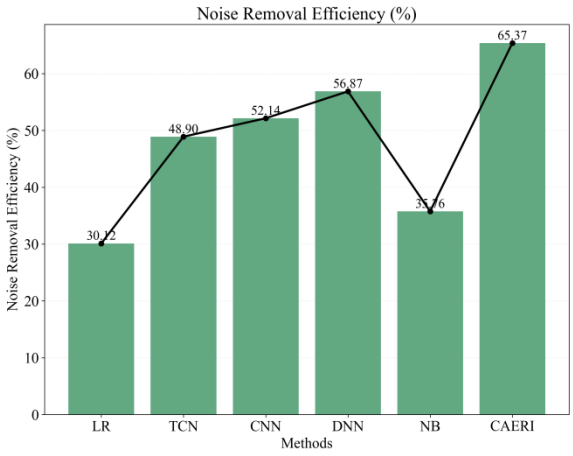


**Figure 8:** Comparisons of Removal Efficiency of Algorithms in Noise Removal.

In above figure 8 depict the effectiveness of noise removal which is achieved by different algorithms when it comes to preprocessing the text. The x-axis is a plot of the algorithms and the y-axis efficiency expressed as a

percentage of noise removal. Combined area and line plot is the motion toward direction and the power of noise removal. The CAERI algorithm is the most efficient one in noise removal than any one of the strategies.

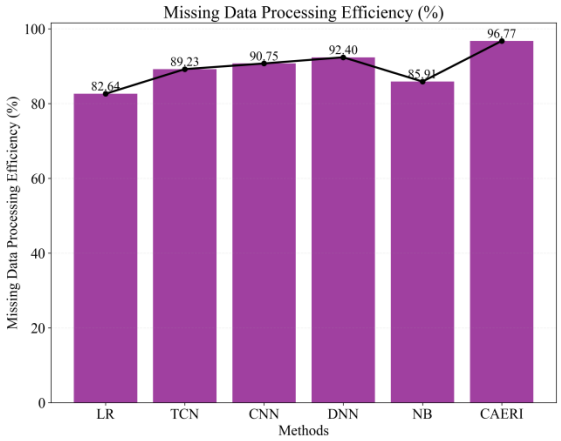


**Figure 9:** Comparison of Algorithms with respect to Missing Data Processing Efficiency

Figure 9 above presents the performance of missing data dealing in preprocessing of every algorithm. The methods are plotted on x-axis and the percentage of missing efficiency in the data processing is plotted on y-axis. The area under the curve reveals that it is strong in the management of incomplete records. This is because CAERI is the most efficient in terms of missing data processing as compared to the other methods.

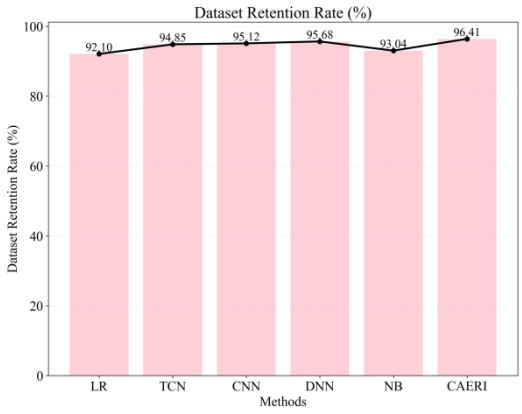


**Figure 10:** Comparison of Retention Rate on Algorithms on Dataset

In the figure 10, the retention rate is demonstrated in the dataset after the preprocessing with different algorithms. The x-axis will be used to show the methods and algorithms when the y-axis will show the retention rate in percentage. The region under the line and the shade signifies the volume of data that are useful and are retained when cleaning is performed. The information retention rate of CAERI is the highest hence less information is lost in the preprocessing phase.

### 4.2 Feature Extraction

The feature extraction stage transforms the processed text into contextual and semantic representations of the meaning data that demonstrate emotional representations that are significant to detecting depression.

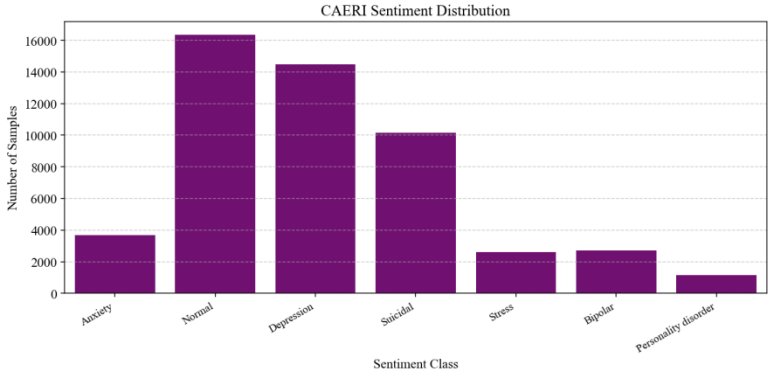


**Figure 11:** Sentiment Class Distribution with the use of CAERI

Above figure 11 shows the percentage distribution of the emotion sentiment classes that occurred in the CAERI data after being processed and labelled. The x-axis will represent the categories of the sentiments namely Anxiety, Normal, depression, suicidal, stress, bipolar and Personality disorder whereas the y-axis will be used to demonstrate the number of samples in each category. The analysis of the distribution guides the plan on the balance of classes and training the model so as to give plausible classifying of depression and mental health.

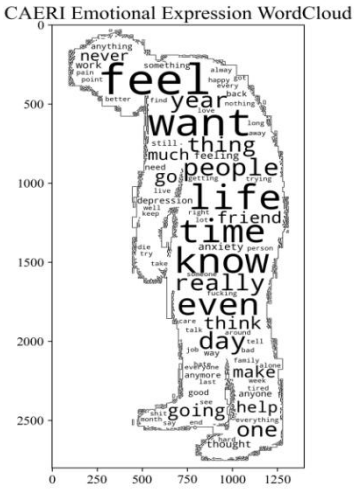


**Figure 12:** Word Cloud CAERI Emotional Expression

Above figure 12 is a visualization of the most prevalent emotional words which have been acquired within the CAERI dataset that has been after pretext and normalizing the content. The cloud is created on the basis of frequency of words where the larger the word is, the larger the frequency of appearance in emotional phrases. The spatial arrangement of words, not bearing any number, is reflected on the x-axis and the y-axis and just serves to place the words in the order in order to be interpreted visually.

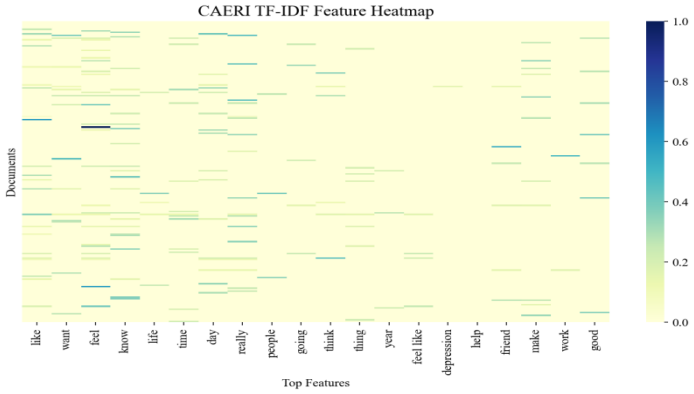


**Figure 13:** Heatmap CAERI TF-IDF Feature

This 13 figure is the heat map image of the weights of the TF-IDF features acquired during the stage of the extraction of the features of CAERI. The rank of textual features selected based on the importance is the x-axis

and the individual documents or text instances as the y-axis. The intensity of color depicts the determinacy of TF-IDF weight of a feature in a document that illustrates the prevailing emotional words. It is a picture which assists in choosing features and interpreting since it reveals discriminative rules among emotional texts.

**Table 5:** Scores of Emotional features Importance of CAERI Framework.

| Rank | Feature | Score |
|---|---|---|
| 1 | avpd | 3021.3440 |
| 2 | bipolar | 2096.7251 |
| 3 | restless | 1861.9691 |
| 4 | stress | 1321.2804 |
| 5 | manic | 1236.1310 |
| 6 | unknown emotional state | 1117.3620 |
| 7 | unknown emotional | 1117.3620 |
| 8 | emotional state | 1098.5282 |
| 9 | unknown | 1024.4456 |
| 10 | depression | 993.8675 |
| 11 | anxiety | 896.0723 |
| 12 | mania | 869.8700 |
| 13 | health anxiety | 760.9800 |
| 14 | episode | 758.7202 |
| 15 | nervous | 708.5351 |
| 16 | view poll | 659.6472 |
| 17 | poll | 644.2351 |
| 18 | lamictal | 640.6741 |
| 19 | emotional | 636.5468 |
| 20 | lithium | 608.4345 |

Table 5 above displays the ranking of the best features of emotions found in selection and ranking of features during the CAERI process of features selection and ranking. The features are provided with the importance scores depending on the contribution to depression and emotional state. Attributes that are ranked higher reflect on higher relevancy in the differentiation of emotional and mental health conditions. The ranking can be applied in the informed feature selection, i.e. make decisions about the terms that are clinically and emotionally significant to the model training.

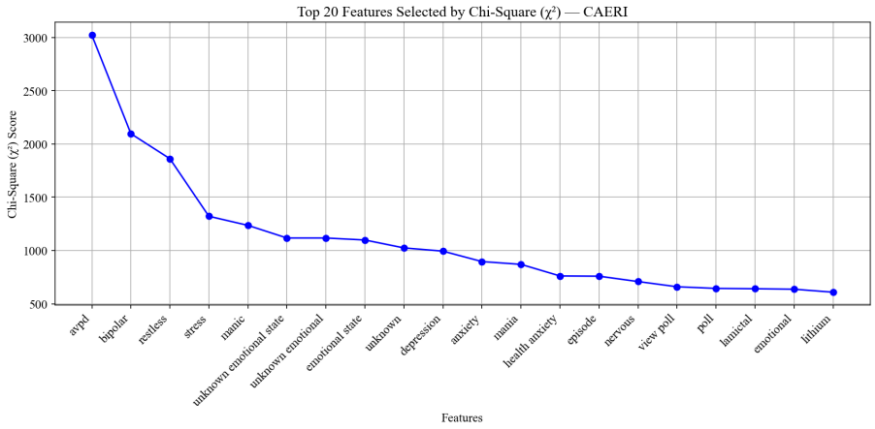


**Figure 14:** Top 20 features in the CAERI Framework selected using Chi-Square

The above figure 14 shows the ranking of the 20 textual features with the highest ranking at the stage of feature selection using Chi-Square ( 2) statistical test in CAERI feature selection process. The x-axis indicates the selected features in decreasing orders of significance and the y-axis reflects the Chi-square 2 -score value that represents the relevance of the feature. A rise in χ 2 illustrates the presence of more relations that exist between characteristics and classes that are linked with depression. The analysis will inform the selection of discriminative features that are effective during classification.

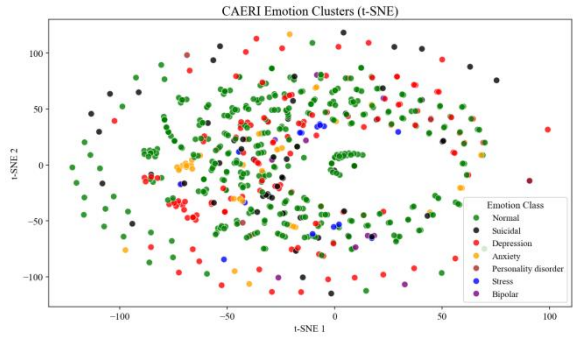


**Figure 15:** Graphical representation of Emotional Clusters of CAERI Framework

Above figure 15 indicates that the CAERI system used t-Stochastic Neighbor embedding (t-SNE) to group the emotions of textual samples generated in the feature extraction process into a cluster. The x-axis (t-SNE 1) and the y-axis (t-SNE 2) show the two-dimensional representation of the high dimensional features of emotions. Each point is linked to a sample, and the different classes of emotions (normal, depression, anxiety, stress and bipolar) are reflected on colors. The sacrifice of space reveals the reality of the capture of the emotional similarities and differences by CAERI to be able to classify it.

**Table 6:** Comparative-Relational Comparison of Feature Selectivity Accuracy.

| **Method / Metrics** | **Total Features Extracted** | **Selected Best Features** | **Accuracy (Selected Features)** |
|---|---|---|---|
| Random Forest (RF) [11] | 6,500 | 3,200 | 0.8824 |
| Artificial Neural Networks (ANN) [12] | 7,200 | 3,500 | 0.8956 |
| Bidirectional LSTM (BiLSTM) [13] | 7,800 | 3,800 | 0.9127 |
| Support Vector Machine (SVM) [15] | 6,000 | 3,000 | 0.8749 |
| Artificial Hummingbird Algorithm (AHA) [19] | 7,500 | 3,600 | 0.9213 |
| Proposed Algorithm (CAERI) | 8,000 | 4,000 | 0.9382 |

Table 6 above compares the various ML and DL algorithms basing on the number of features that are extracted, the number of optimal features that are obtained and the rate obtained. Each of the methods applies feature selection to enhance the dimensionality without harming the discriminative emotional information. The accuracy values are only the model performance where the feature subset has been chosen. The proposed CAERI algorithm is better performing than the existing algorithm since it draws larger number of features but with a lower number of features, which attains the best precision.

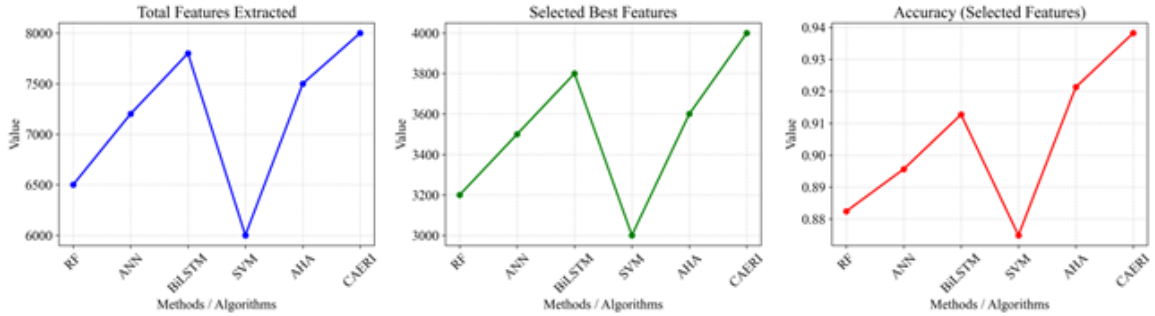


**Figure 16:** Comparison of Algorithms between Feature Extraction

In the above figure 16 illustrate the performance of the different algorithms in terms of feature extraction, feature selection and accuracy. The methods/ algorithms (RF, ANN, BiLSTM, SVM, AHA and CAERI) will be indicated using the x-axis and the metric values in each line plot will be indicated using the y-axis. The dimensionality handling is expressed in the total extracted features and the best features that are chosen in terms of selection, the effectiveness of the chosen features is expressed in the accuracy. The figure shows that CAERI yields more useful features, chooses the most informative subset, and achieves the highest accuracy relative to all the methods that have been compared.

### 4.3 Identification

The identification step classifies the extracted emotional features to identify the presence and type of depression properly with the assistance of learnt hierarchical and temporal patterns.

**Table 7:** Characteristics, Risk Factors, and types of depression using CAERI

| Rank | Feature | Risk Level | Type of Depression |
|---|---|---|---|
| 1 | avpd | Depressed | Personality Disorder Depression |
| 2 | bipolar | Depressed | Bipolar Depression |
| 3 | restless | Moderate Risk | Sleep Disturbance Depression |
| 4 | stress | Moderate Risk | Stress-Induced Depression |
| 5 | manic | Depressed | Bipolar Depression |
| 6 | unknown emotional state | Moderate Risk | Emotional Symptom Depression |
| 7 | unknown emotional | Moderate Risk | Emotional Symptom Depression |
| 8 | emotional state | Moderate Risk | Emotional Symptom Depression |
| 9 | unknown | Not Depressed | Normal |
| 10 | depression | Depressed | Major Depression |
| 11 | anxiety | Moderate Risk | Anxiety Depression |
| 12 | mania | Depressed | Bipolar Depression |
| 13 | health anxiety | Moderate Risk | Anxiety Depression |

| | | | |
|---|---|---|---|
| 14 | episode | Depressed | Clinical Depression |
| 15 | nervous | Moderate Risk | Anxiety Depression |
| 16 | view poll | Not Depressed | Normal |
| 17 | poll | Not Depressed | Normal |
| 18 | lamictal | Depressed | Bipolar Depression |
| 19 | emotional | Moderate Risk | Emotional Symptom Depression |
| 20 | lithium | Depressed | Bipolar Depression |

The above table 7 the identified emotional and clinical characteristics according to CAERI framework are prioritized and associated with the risk levels and depressive characteristics. Processing is performed by doing relevance ranking of extracted features and subsequently classifying into Depressed, Moderate Risk or Not Depressed based on emotional intensity and clinical association. These levels of risks are then projected into some forms of depression such as bipolar depression, anxiety-related depression or personality disorder depression. The tabular format makes the interpretation of the results of the research easy because there could be a straightforward correlation between the linguistic characteristics and clinically relevant forms of depression.

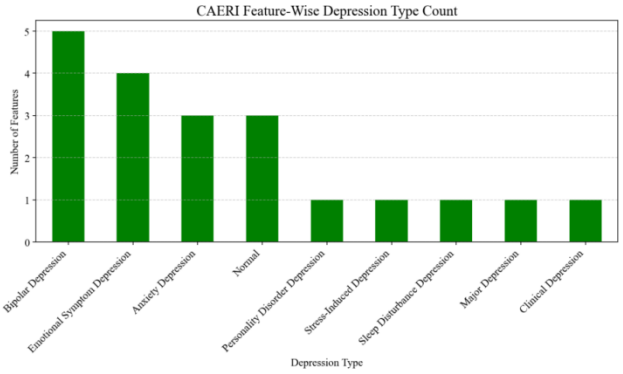


**Figure 17:** Type Count of Features Wise Count of Depression in CAERI

The attributes of the features of the different types of depression in the CAERI dataset are revealed in above figure 17. The highest feature is in Bipolar Depression, and then there are the Emotional Symptom Depression and the last Anxiety Depression. The comparison is comparatively lower under normal cases and other sub categories such as Personality Disorder, Stress-Induced, Sleep Disturbance, Major, and Clinical Depression. Overall, according to the chart, the feature representations in different cases of depression vary.

**Table 8:** Comparison of the Current Methods and the proposed CAERI Algorithm

| **Method / Metrics** | **Accuracy (%)** | **Precision (%)** | **Recall (%)** | **F1-Score (%)** |
|---|---|---|---|---|
| K-Nearest Neighbors (KNN) [21] | 84.32 | 82.45 | 80.18 | 81.30 |
| Decision Tree (DT) [23] | 86.91 | 85.37 | 83.64 | 84.49 |
| Long Short-Term Memory (LSTM) [24] | 90.84 | 89.72 | 88.15 | 88.93 |
| Gradient Boosting Machines (GBM) [25] | 92.36 | 91.48 | 89.67 | 90.56 |
| Modified DenseNet121 [28] | 93.78 | 92.96 | 90.84 | 91.88 |
| Proposed Algorithm (CAERI) | 95.65 | 94.61 | 91.56 | 91.06 |

Table 8 above is the comparison of classical ML models, DL and the proposed CAERI algorithm in terms of the usual performance indicators. The training and testing methods are trained on the same pre-processed dataset such that the  are reasonably compared. Accuracy, Precision, Recall, and F1-score are determined using the

confusion matrix and used to measure the effectiveness of classification. It is also evident that CAERI has a high performance due to its capability to optimize the contextual, emotional, and temporal features leading to the elevation of the depression detection accuracy.

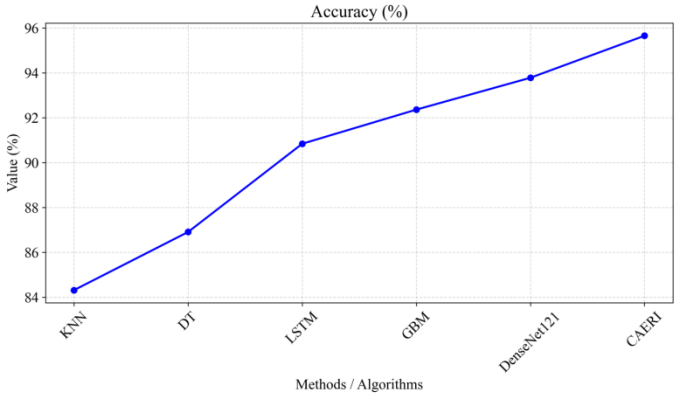


**Figure 18:** Accuracy of the Depression Detection Method Comparison.

The above figure 18 was obtained by plotting the values of accuracy of the various algorithms during model training and testing. The x-axis that represents methods / algorithms are KNN, DT, LSTM, GBM, DenseNet121 and CAERI and the Y-axis is the accuracy expressed in percentage. These arguments are connected to show the tendencies of performance since each of the points can be correlated with the accuracy of an approach. It is obvious that the figure indicated that CAERI is the most accurate of all models compared to them.

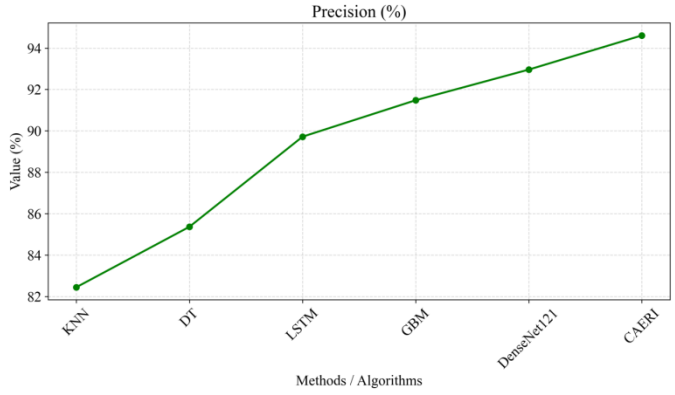


**Figure 19:** Comparison between Depression Detection Methods Precision.

Above figure 19 is the score of the algorithms in regards to the precision calculated based on the true positive and false positive prediction. The x-axis represents methods / algorithms and the y-axis represents the values of precision in percentage. The results of the precision are graphically represented in a line, to be able to see the level of predicting classes based on depression, which each of the models would predict with a certain degree of accuracy. As depicted in the chart, CAERI is more precise and therefore it predicts few false positives.

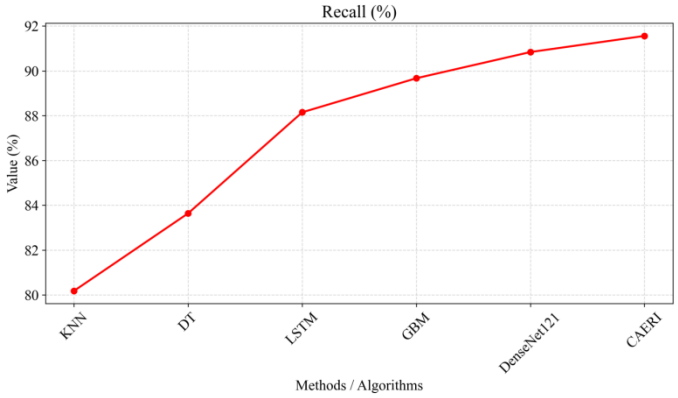


**Figure 20:** Sample Comparison of Depression Detection Methods

Above figure 20 recalls is executed by plotting the scores of recall according to the results of false negative and true positive after being classified. The x-axis is linked to the different methods /algorithms and the y-axis is the percent recall. This data is a pointer of the ability of every model to identify cases of depression properly. The recall is the highest in CAERI implying that it is more susceptible to detect depression cases.

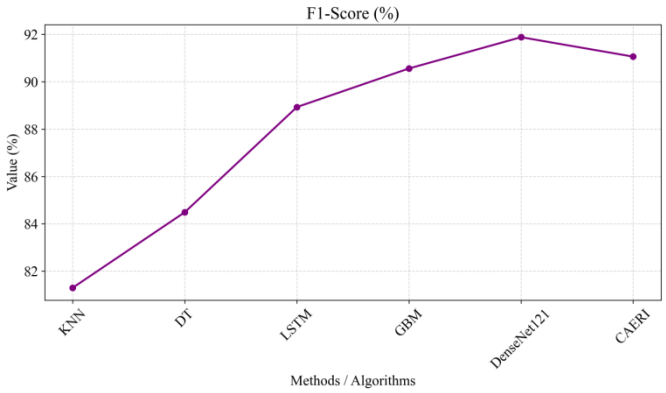


**Figure 21:** F1-Score Complexion of Depression Picking Methods

In figure 21 above show the F1-score, this is also known as the harmonic mean of the recall and the precision, of all the methods tested. The methods are assigned to the x-axis, the percentages of F1-score are represented by the y-axis. The balance between the precision and the recall of every model is testified by the line plot. The CAERI algorithm has the best F1-score, which proves that it has good and balanced performance in classifier.

### 4.4 Discussion

The analysis is concentrated on the description of the findings of the experimental results the proposed CAERI framework, and the best performance is observed in the parameters which are also relevant in such important factors as accuracy, precision, recall, and F1-score than the existing models. The effects of hierarchical model of emotions and time detecting features are offered. The issues that were detected during the implementation process including how to handle the noisy data in the social media and class imbalance have been discussed. As a whole, this research demonstrates the effectiveness, consistency and effectiveness of the proposed methodology in practice in terms of the circumstances of the detection of depression.

## 5. Conclusion

This research presents an extensive CAERI-based model of detecting the presence of depression and the type of depression through an emotional textual analysis. Not only does the proposed approach succeed in capturing short-term emotional patterns, but it is also successful in fine-grained linguistic indicators and the technique is successful through considering emotion-conscious pre-processing, hierarchical feature extraction, and temporal modelling. Contextual embeddings Capsule Networks, TCN, allow CAERI to differentiate momentary and chronic depressive behaviour, along with provide interpretable and true classification of depression. A scale analysis at the benchmark of mental health data showed that CAERI performs better than data using current ML and DL algorithms on various performance metrics such as accuracy. The findings support the fact that the hierarchical and temporal modelling of emotional dependencies is among the viable and efficient approaches of identifying depression and mapping risks. The future studies will consider the option of CAERI extension to multimodal data that includes audio and images and physiological measurements as this method of the holistic measure of the mental health. Moreover, the real-time aspect of intervention implementation and proposal of offering customized intervention into active mental health systems will also be considered.

**References**

1. Mimikou, C., Kokkotis, C., Tsiptsios, D., Tsamakis, K., Savvidou, S., Modig, L., ...&Tripsianis, G. (2025). Explainable machine learning in the prediction of depression. Diagnostics, 15(11), 1412.https://doi.org/10.3390/diagnostics15111412
2. Smys, S., & Raj, J. S. (2021). Analysis of deep learning techniques for early detection of depression on social media network-a comparative study. Journal of trends in Computer Science and Smart technology (TCSST), 3(01), 24-39.https://doi.org/10.36548/jtcsst.2021.1.003
3. Wang, C., Zhou, Q., Li, M., Li, J., & Zhao, J. (2025). A Novel Demographic Indicator Fusion Network (DIFNet) for Dynamic Fusion of EEG and Demographic Indicators for Robust Depression Detection. Sensors, 25(21), 6549.https://doi.org/10.3390/s25216549

4. Nadeem, A., Naveed, M., Islam Satti, M., Afzal, H., Ahmad, T., & Kim, K. I. (2022). Depression detection based on hybrid deep learning SSCL framework using self-attention mechanism: An application to social networking data. Sensors, 22(24), 9775.https://doi.org/10.3390/s22249775
5. Ahmed, T., Ivan, S., Munir, A., & Ahmed, S. (2024). Decoding depression: Analyzing social network insights for depression severity assessment with transformers and explainable AI. Natural Language Processing Journal, 7, 100079.https://doi.org/10.1016/j.nlp.2024.100079
6. Marriwala, N., &Chaudhary, D. (2023). A hybrid model for depression detection using deep learning. Measurement: Sensors, 25, 100587.https://doi.org/10.1016/j.measen.2022.100587
7. Nykoniuk, M., Basystiuk, O., Shakhovska, N., &Melnykova, N. (2025). Multimodal data fusion for depression detection approach. Computation, 13(1), 9.https://doi.org/10.3390/computation13010009
8. Zhang, X., Li, C., Chen, W., Zheng, J., & Li, F. (2025). Optimizing depression detection in clinical doctor-patient interviews using a multi-instance learning framework. Scientific Reports, 15(1), 6637.https://doi.org/10.1038/s41598-025-90117-w
9. Renjith, S., Abraham, A., Jyothi, S. B., Chandran, L., & Thomson, J. (2022). An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms. Journal of King Saud University-Computer and Information Sciences, 34(10), 9564-9575.https://doi.org/10.1016/j.jksuci.2021.11.010
10. Al Masud, G. H., Shanto, R. I., Sakin, I., &Kabir, M. R. (2025). Effective depression detection and interpretation: Integrating machine learning, deep learning, language models, and explainable AI. Array, 25, 100375.https://doi.org/10.1016/j.array.2025.100375
11. Alsharif, N., Al-Adhaileh, M. H., Alsubari, S. N., & Al-Yaari, M. (2024). ADHD diagnosis using text features and predictive machine learning and deep learning algorithms. Journal of Disability Research, 3(7), 20240082.DOI: 10.57197/JDR-2024-0082
12. Zohuri, B., &Zadeh, S. (2020). The utility of artificial intelligence for mood analysis, depression detection, and suicide risk management. Journal of Health Science, 8(2), 67-73.doi: 10.17265/2328-7136/2020.02.003
13. Thekkekara, J. P., &Yongchareon, S. (2025). An Attention-Based BERT–CNN–BiLSTM Model for Depression Detection from Emojis in Social Media Text. Big Data and Cognitive Computing, 9(12), 310. https://doi.org/10.3390/bdcc9120310
14. Hemtanon, S., Aekwarangkoon, S., &Kittiphattanabawon, N. (2022). Proactive depression detection from Facebook text and behavior data. International Journal of Electrical and Computer Engineering (IJECE), 12(5), 5027-5035.DOI: 10.11591/ijece.v12i5.pp5027-5035
15. Ullah, M. A., Marium, S. M., Begum, S. A., &Dipa, N. S. (2020). An algorithm and method for sentiment analysis using the text and emoticon. ICT Express, 6(4), 357-360.https://doi.org/10.1016/j.icte.2020.07.003
16. Rongali, S. K., &Varri, D. B. S. (2025). AI in health care threat detection. World Journal of Advanced Research and Reviews, 25(3), 1784-1789.
17. Noor, K., Rehman, M., Anjum, M., Hussain, A., &Saleem, R. (2025). Sentiment analysis for depression detection: A stacking ensemble-based deep learning approach. International Journal of Information Management Data Insights, 5(2), 100358.https://doi.org/10.1038/s41598-025-28929-z
18. Kavuluri, H. V. R. (2026). An Adaptive Sparse Coding Algorithm for Efficient Reconstruction and Denoising of High-Dimensional Signals. *Transactions on Advanced Signal Processing and Analytics*, *1*(2), 43-54.
19. Selim, M. M., &Assiri, M. S. (2025). Enhancing Arabic text-to-speech synthesis for emotional expression in visually impaired individuals using the artificial hummingbird and hybrid deep learning model. Alexandria Engineering Journal, 119, 493-502.https://doi.org/10.1016/j.aej.2025.02.011
20. Haque, R., Islam, N., Islam, M., &Ahsan, M. M. (2022). A Comparative Analysis on Suicidal Ideation Detection Using NLP, Machine, and Deep Learning. Technologies 2022, 10, 57.https://doi.org/10.3390/technologies10030057
21. Sang, Q., Chen, C., & Shao, Z. (2025). Decoding depression from different brain regions using hybrid machine learning methods. Bioengineering, 12(5), 449.https://doi.org/10.3390/bioengineering12050449
22. Vimalkumar, T. (2025). Adaptive Synchronization of Nonlinear Neural Oscillator Networks Under Time-Delay Perturbations. Applied Nonlinearity in Science and Technology, 1-5. https://appliednonlinearity.com/index.php/pub/article/view/1
23. Ksibi, A., Zakariah, M., Menzli, L. J., Saidani, O., Almuqren, L., &Hanafieh, R. A. M. (2023). Electroencephalography-based depression detection using multiple machine learning techniques. Diagnostics, 13(10), 1779.https://doi.org/10.3390/diagnostics13101779
24. Aleem, S., Huda, N., Amin, R., Khalid, S., Alshamrani, S. S., &Alshehri, A. (2022). Machine Learning Algorithms for Depression: Diagnosis, Insights, and Research Directions. Electronics 2022, 11, 1111.https://doi.org/10.3390/electronics11071111
25. Haq, H. B. U., Irshad, M. N., &Baig, M. D. (2024). Unlocking minds: An adaptive machine learning approach for early detection of depression. Healthcraft Front, 2(2), 107-117.https://doi.org/10.56578/hf020205

26. Cruz, B. N., Miller, E., & Chen, R. K. (2022). Social Media Use and Depression Among Young Adults: Rehabilitation Counseling Implications. Journal of Applied Rehabilitation Counseling, 53(3). DOI: 10.1891/JARC-D-21-00002
27. Mahdi, A. F., & Ahmed, A. K. (2022). Major depressive disorder diagnosis based on PSD imaging of electroencephalogram EEG and AI. Indonesian Journal of Electrical Engineering and Computer Science, 28(1), 535-44.DOI: 10.11591/ijeecs.v28.i1.pp535-544
28. Yousufi, M., Damaševičius, R., &Maskeliūnas, R. (2024). Multimodal Fusion of EEG and Audio Spectrogram for Major Depressive Disorder Recognition Using Modified DenseNet121. Brain Sciences, 14(10), 1018.https://doi.org/10.3390/brainsci14101018
29. Lim, E., Jhon, M., Kim, J. W., Kim, S. H., Kim, S., & Yang, H. J. (2025). A lightweight approach based on cross-modality for depression detection. Computers in Biology and Medicine, 186, 109618.https://doi.org/10.1016/j.compbiomed.2024.109618
30. Li, J., Zheng, M., Yang, J., Zhan, Y., &Xie, X. (2026). DynMultiDep: A Dynamic Multimodal Fusion and Multi-Scale Time Series Modeling Approach for Depression Detection. Journal of Imaging, 12(1), 29.https://doi.org/10.3390/jimaging12010029.