Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PSYCHOLOGICAL HEALTH ASSESSMENT MODEL USING WEIGHTED COUPLED MULTILEVEL VARIABLE SET WITH T5 TRANSFORMER

[1]M PRAVARSHA REDDY, [2]LAKSHMI RAMANI BURRA

Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram Andhra Pradesh, INDIA
mekalapravarshareddy@gmail.com
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram Andhra Pradesh, INDIA
ramanimythili@gmail.com

ABSTRACT

The system of psychological health assessment and problem identification primarily consists of multivariate, multilevel variables with strong coupling, and is characterized by complex nonlinear interactions among different factors. The field of psychological quality and the methodologies used to assess it has seen substantial growth with its widespread application. Unfortunately, the conclusions of the present evaluation methods, which rely heavily on basic statistical analysis, are inaccurate. College and university student affairs departments must so actively seek out and address adults' mental health issues as soon as they manifest. With the advancements in intelligent technology, it is now feasible to assess and forecast adults' mental states in real time using multimodal data and deep learning models. The use of transformer-based language models in healthcare contexts is still in its infancy, nonetheless, despite their fast advancement. Text-to-Text Transfer Transformer (T5) model is used in this research, which makes use of an encoder-decoder architecture. The decoder produces the output text after processing the input text by the encoder. The encoder's job is to take in the syntactic and semantic information from the input text and store it in a contextualized representation. This research proposes a Psychological Health Assessment Model using Weighted Coupled Multilevel Variable Set with T5 Transformer (PHA-WCMVS-T5) for accurate health assessment. The proposed model when compared with the traditional models performs better in psychological health assessment.

Keywords: *Psychological Health Assessment, Multilevel Variables, Psychological Quality, Adult's Mental Health, Multimodal Data, Deep Learning, Text-to-Text Transfer Transformer*.

1. INTRODUCTION

The goal of Psychological disease prediction using text is to detect mental problems from textual data using deep learning algorithms and Natural Language Processing (NLP) techniques [1]. Unfortunately, the present systems' use of generic language models and standard embedding techniques to produce text embeddings is the main cause of their unsatisfactory performance. There has to be domain-specific pretrained language models that fully grasp the context of posts made by people with mental disabilities in order to overcome this constraint [2]. Posts made by people with mental illness frequently include metaphorical terms, which makes it difficult for current models to comprehend this type of discourse. Anxiety and sadness are the most prevalent abnormalities in psychological health, which is defined as a harmonious state in which individual psychological processes are balanced [3].

College students are vulnerable to a range of mental health issues, with depression being among the most common, as they navigate the transition from campus life to adulthood and face numerous forms of external pressure [4]. Evaluation of college students' psychological health, promotion of education on college students' psychological health, and, finally, the realization of healthy development of college students' psychology are all greatly aided by a scientific and flawless assessment and analysis model of college students' psychological health [5]. To begin with, the input text is enhanced by incorporating part-of-speech information into the word vector; this serves to disambiguate words and improves the input quality for the neural network [6]. In contrast, the model uses a segmentation pooling technique to extract the most aspects of sentences while taking the text's structural information into account [7]. Unfortunately, there aren't enough scientific and consistent standards in the present model for assessing and analyzing college students' mental health, and there aren't enough evaluation signs. Psychological health evaluation is a challenging field because of the variety, complexity, multiplicity, urgency, and contingency of issues affecting college students' mental health [8]. Therefore, it is important to develop a professional analysis model for evaluating college students' mental health and conduct thorough research into the psychological traits of today's college students based on an examination of the variables that impact their mental health [9].

The use of Convolutional Neural Networks (CNNs) to evaluate college students' psychological health levels offers several benefits. With the CNN-based training algorithm, the issue of gradient disappearance in mental health analysis is effectively resolved, resulting in a longer effective training period [10]. In a huge data setting, the CNN algorithm's great complexity and capability provide good generalization, which is necessary for the assessment and study of college students' mental health. CNNs have several useful applications in mental health analysis, such as improving the efficiency of data extraction from large datasets, improving the accuracy of psychological assessments, and breaking down complex interrelated factors into simpler, more manageable components [11]. Indicators can be aggregated into layers or element groups with basically the same attributes; there is no mutual influence or dominance between elements within the layers; and the number of mapping relationships in the model can predict the pros and cons of single-layer and multi-layer model accuracy [12]. In order to enhance the training performance of the CNN for the evaluation and analysis of college students' mental health, it incorporates a feature learning component onto the foundation of the conventional multi-layer neural network and makes use of the spatial relative relationship to decrease the number of parameters [13].

The system of psychological health assessment and problem identification primarily consists of multivariate, multilevel variables with strong coupling, and is characterized by complex nonlinear interactions among different factors [14]. CNNs have been a game-changer in artificial intelligence by streamlining algorithms, reducing parameter counts, and making otherwise difficult issues easier to understand and solve. As a result, CNN offers significant benefits in developing the model for evaluating the mental health of college students. However, complex, multimodal psychological data, which may incorporate text, physiological signs, and behavioral patterns, is not well-suited to CNNs because of their primary intended for processing structured spatial data, like images [15]. In order to attain high accuracy, they often need a lot of labeled training data, which isn't always easy to come by in the mental health area since datasets are usually small, diverse, or sensitive because of privacy concerns. CNNs are notoriously difficult to interpret because they are essentially black boxes. This is especially problematic in the healthcare industry, where clinicians rely on model predictions but need to know why in order to act on them. Mood tracking and patient-reported outcomes are examples of time-series data that CNNs may find difficult to handle because of their potential inability to capture the temporal dynamics and long-range dependencies in psychological states [16]. Neural networks struggle to generalize their predictions to other populations due to their sensitivity to demographic imbalances and their propensity for overfitting on limited datasets. These constraints imply that CNNs can extract useful information, but that they frequently necessitate additional models or modifications tailored to the domain in order to be fully effective in evaluating mental health [17].

T5 Transformer is designed to handle sequence modeling challenges, particularly in the NLP sector. Transformer uses a scaled dot-product attention method that can be parallelized. This one-of-a-kind attention mechanism enables pretraining on a massive scale. It was also possible to train transformers without expensive annotations by using self-supervised pretraining on big unlabeled datasets using techniques like input masking [18]. While Transformers were initially developed for NLP, they have now been adapted for use in computer vision, remote sensing, time series, speech processing, and multimodal learning, among other fields. Unfortunately, no one has yet done a thorough literature search for any and all healthcare-related works that have used the Transformer architecture. Understanding common patterns and features in the rapid incorporation of the transformer architecture into nearly every healthcare domain is of the utmost importance, and this is happening at the perfect time [19].

A T5 Transformer integrates text-to-text tasks within a text-to-text framework. T5 transforms all jobs into a text-to-text output format, unlike older models that are tailored for specialized tasks like categorization, translation, or summarization. As a result, it can tackle a broad variety of issues related to natural language with ease and power. The encoder in T5 takes in text with its context and makes sense of it, while the decoder uses that information to make sense of the output. There are several benefits to using T5 for psychological health assessments. The T5 model is able to grasp complex linguistic patterns, which allows for more precise identification of signs associated with mental health issues like anxiety, depression, and stress. Because mental health symptoms are frequently nonlinear and interrelated, it is crucial that this system be able to identify correlations between multilevel and multivariate components. Its contextual embedding method makes this possible. The transformer model workflow is shown in Figure 1.

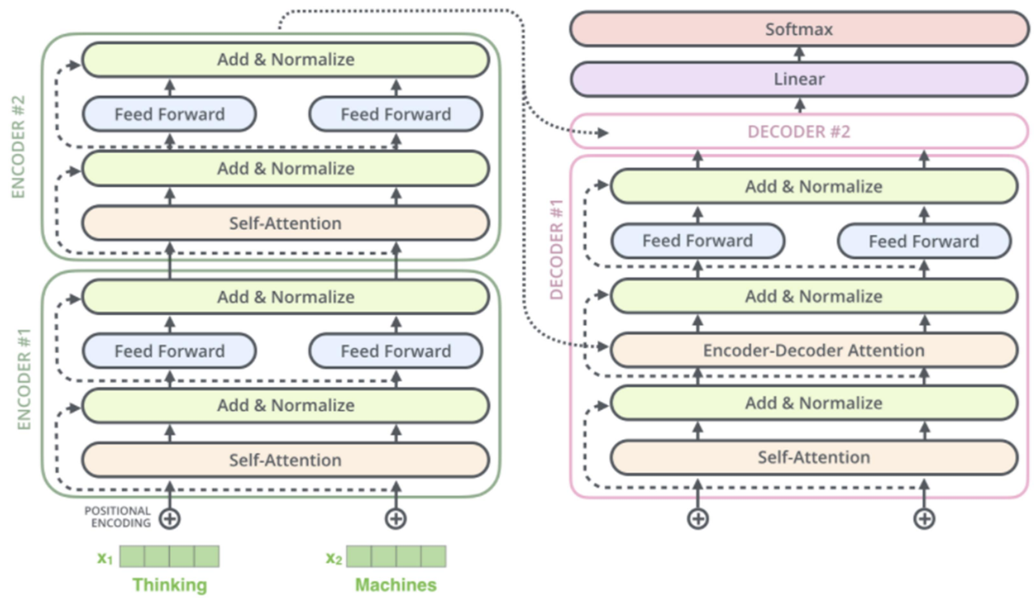

Fig 1: T5 Transformer Model Architecture

The T5 model is crucial for interpreting language data and producing accurate health forecasts when it is an integral part of the PHA-WCMVS-T5. T5 uses its extensive knowledge of context to generate reliable evaluations, and the weighted linked multilevel variable set architecture aids in simplifying complicated relationships among psychological components. This process reduces the complexity of the model's parameters and boosts its computing efficiency, making it suitable for scalable and real-time applications. By adjusting to different parameters, T5-based model can also help with tailored psychological evaluation. This makes sure that tests are correct from a technical standpoint, but they are also considerate of the fact that people's mental health might manifest in different ways. Adults dealing with mental health issues may soon have access to real-time, dependable, and adaptive help with the PHA-WCMVS-T5, an innovative system that combines deep learning with transformer-based designs.

The major contribution of this work is to address the limitations of existing psychological health assessment models by integrating weighted multilevel feature representation with an enhanced T5 Transformer architecture. Unlike previous approaches that mainly focus on single-task prediction or conventional deep learning models, the proposed framework explicitly models coupled feature interactions through dedicated weighted linking and normalization layers, thereby improving contextual understanding, prediction accuracy, and computational efficiency for psychological health assessment.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. LITERATURE SURVEY

In recent years, Automatic Text Summarizing (ATS) has emerged as an essential tool for extracting more effective and valuable information from large texts by condensing them into brief and informative summaries. To aid in the study and comprehension of psychological contents, ATS condenses statistical material into a more comprehensible and concise version, with a particular emphasis on psychological text summary in order to extract insights and emotional states. With this in mind, Khan et al. [1] suggested a novel hybrid model, T5-LSTM FusionNet, to improve psychological text summarization. The increasing availability and breadth of online psychological literature is driving the need for novel approaches to reliably and rapidly extracting important findings. Th top pick is the T5-LSTM FusionNet model, which brings together the best features of T5 and LSTM. The dataset of 5480 records, collected from several websites related to psychology, has been utilized to assess its efficacy. T5-LSTM FusionNet is tested against many state-of-the-art models, including T5, LSTM, BERT, and DistilBERT, to determine its overall performance.

Identifying people who are having suicidal thoughts is an important area of study with enormous promise for better mental health support networks. Accessing large-scale, annotated datasets that are essential for training successful machine learning models is challenging due to the sensitivity surrounding data related to suicide. Ghanadian et al. [2] presented a novel approach to overcome this shortcoming by making use of generative AI models like Llama, ChatGPT, and Flan-T5 to generate artificial data for the purpose of detecting suicidal thoughts. To guarantee that crucial information on suicidal thoughts is covered, the data generation approach is based on social elements taken from the psychological literature. The author compared the results to those of cutting-edge NLP classification algorithms, with a focus on BERT-based models. These traditional models typically produce F1-scores between 0.75 and 0.87 when trained on the UMD real-world dataset. Based on social aspects, the synthetic data-driven approach consistently achieves F1-scores of 0.82 for both models, indicating that the variety of synthetic data subjects can help close the performance gap when dealing with models of varying complexity.

The state of mental health, particularly among college students, is a growing concern for millions of people around the world. Community Question Answering for Psychological Wellness (CQA-PW) uses query-response pairing to improve mental health services, and Ma et al. [3] provided a new model called Psychological Snapshot Guided Pairing (PSG-Pair) to meet this demand. In contrast to conventional approaches, PSG-Pair incorporates role-based psychological snapshots culled from the past posts of those seeking assistance and those offering it. In the first stage, known as screening, the model employs a BERT-based retrieval model to weed out irrelevant supportive posts. In the second stage, called pairing, psychological snapshots are used in conjunction with a stacked attention mechanism to fine-tune conceptual pairings according to users' psychological traits. Extensive studies on the CLPsych 2022 Shared Task dataset show that PSG-Pair improves pairing accuracy and memory more than conventional single-phase models. Adding psychological snapshots improves the model's ability to handle complicated psychological health scenarios, which is particularly helpful for university students. This, in turn, makes automated psychological support systems more successful. The study does, however, have a number of caveats. One issue is that the dataset used for testing has a lot of irrelevant negative examples, which makes it noisy and imbalanced. This could affect how well the model performs. Secondly, although the method shows potential when applied to mental health, it is unclear whether the model can be applied to other fields or uses.

As a result of inadequate resources for early detection of depression, many people are struggling with mental health concerns. Conditions like anxiety disorders, bipolar disorders, sleep problems, depression, and, in extreme circumstances, self-harm and suicide, are largely caused by this delay in seeking therapy. This makes the already formidable challenge of locating people with mental health issues and providing timely assistance even more formidable. So, to help with early depression prediction, Almutairi et al. [4] presented a new hybrid deep-learning approach. An LSTM and a CNN model are the major components of the hybrid deep learning model that the author presented for depression prediction in this research. The feature attention mechanism is used to apply Two-State LSTM (TS-LSTM), an improved version of the LSTM technique. The suggested architecture improves the TS-LSTM method's relationship identification and keyword extraction capabilities for depression detection by integrating a feature attention mechanism. Using this methodology, the author analyzed a massive dataset collected from a youth-oriented, publicly-available online platform. Questions posed by younger platform users in text form make up this dataset. From robust markers of possible depressive symptoms that were predefined by medical and psychiatric professionals, the author extracted features using a one-hot encoding method. When tested against more traditional methods, our technology consistently outperforms them.

Depression and negative emotional cycles have a lasting impact on both the quality of life and productivity at work. But it's not easy to spot depression; most of the tools used today depend on scales, which don't allow for direct and immediate measurement of degree of depression. Finding the precise locations and points of feedback electrodes for brainwave stimulation that may be associated with depression is the overarching goal of this empirical investigation. According to brainwave data gathered from mood-induction methods, the front and occipital lobes specifically, the Fp1 and Fp2 positions and the O1 and O2 positions play a crucial role in the regulation of depressive moods. The minimum value of approximated signals is significantly influenced by the wavelet brainwave bands, whereas the α and θ bands primarily influence the Fourier brainwave bands. Chiang et al. [5] presented eight possible models for evaluating depression. The models built using deep neural networks outperformed the others in terms of stability and accuracy. The signal processing methods used in the study include multilayer perceptron, deep neural network, deep belief network, and long short-term memory. Thus, this model can be used as a supplementary approach to quickly and objectively evaluate underlying depression, which can help with self-management of emotions and early diagnosis and treatment of depression.

The effects of mental health illnesses on a person's and society's capacity for rational thought, emotional regulation, and social cohesion are far-reaching and long-lasting. An exciting new avenue for early detection and intervention in mental health is sentiment analysis, which is gaining attention as a growing global concern. The complexity, domain-specific subtleties, and specialist vocabulary of literature pertaining to mental health make traditional approaches inadequate. In response to these difficulties, Hossain et al. [6] presented BERT-Fuse, a hybrid deep learning model that integrates state-of-the-art neural architectures such as Bidirectional Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Transformer blocks, and CNN with more conventional feature extraction methods using TF-IDF and Bag-of-Words. Within BERT-Fuse, the author tested three word embedding methods: BERT, GloVe, and FastText. The author discovered that BERT embeddings outperformed the other two in every measure and looked at: accuracy, precision, recall, and F1-score. A well-rounded dataset with text samples covering seven mental health problems was used to train the algorithm.

The medical industry now requires Internet-Delivered Psychological Treatment (IDPT). Training models to function at the level of clinicians requires big, diverse patient populations for deep neural networks (DNNs). When applied to new data sets, however, DNN models trained on small datasets perform poorly in clinical settings. Therefore, it is critical to increase the availability of varied and unique training data. A structural hypergraph and an emotive lexicon for word representation are proposed by Ahmed et al. [7]. A federated learning-based embedding model was created for the purpose of detecting symptoms related to mental health. The model views text data in a sequential word form. Keeping the contextual relationship in mind, the model proceeds to train a continuous vector with low dimensions. Experiments are subsequently conducted to evaluate the produced models that incorporate attention-based mechanisms and federated learning. This method works well for dynamic lexicon analysis, grammatical word representation, and vocabulary diversification. Making use of an attention network model, the author aimed to generate semantic word representations. Afterwards, the text is marked by embedding it using clinical techniques.

In order to automatically detect whether a Twitter user lives with one of nine mental health problems, Villa-Pérez et al. [8] analyzed their messages using classical and deep learning algorithms. The participants' native languages are English and Spanish. The author made two sets of data, one in English and one in Spanish. The author collected the timelines of 1,500 individuals who have stated

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

in at least one post that they have been diagnosed with ADHD, Anxiety, Autism, Bipolar, Depression, Eating disorders, Obsessive-Compulsive Disorder, Post-Traumatic Stress Disorder, or Schizophrenia. The 1,700 users who were chosen at random and who did not report a diagnosis make up the control set. The author trained both traditional machine-learning and deep learning classifiers for two tasks: binary classification to distinguish between diagnosed and non-diagnosed users, and multiclass classification to identify the specific diagnosis. From the collected data, the author extracted a variety of text features, including n-grams, q-grams, Part-of-speech (POS) tags, topic modeling, Linguistic Inquiry and Word Count (LIWC), and word embeddings. When comparing the two classification tasks, XGBoost and CNN came out on top.

2.1 Research Gap and Motivation

Although recent studies have demonstrated significant progress in psychological health assessment using deep learning and transformer-based architectures, several important research gaps still exist. Existing approaches such as T5-LSTM primarily focus on psychological text summarization rather than comprehensive psychological health assessment. Transformer-based methods including BERT, Flan-T5, and LLM-based frameworks mainly concentrate on specific tasks such as suicidal ideation detection or depression classification without considering the complex relationships among multilevel psychological variables. Similarly, CNN- and LSTM-based models exhibit limitations in capturing nonlinear interactions among heterogeneous psychological factors and often fail to model feature dependencies effectively. Most existing studies also assign equal importance to all input features, ignoring the varying contribution of psychological, behavioral, and contextual variables. Furthermore, current approaches lack an efficient mechanism to jointly model feature weighting, coupled variable interactions, and contextual language understanding within a unified framework.

To overcome these limitations, this research proposes the Psychological Health Assessment Model using Weighted Coupled Multilevel Variable Set with T5 Transformer (PHA-WCMVS-T5). The proposed framework introduces a Weighted Feature Similarity Linked Layer (WFSL2), Dual Linked Interface Layer (DLIL), and Normalization Vector Layer (NVL) before the T5 Transformer encoder. These components enable adaptive feature weighting, explicit modeling of coupled multilevel variables, dimensionality reduction, and efficient contextual representation learning. Consequently, the proposed model improves feature extraction, similarity linking, transformer processing efficiency, and overall psychological health assessment accuracy compared with existing state-of-the-art methods.

3. PROPOSED MODEL

The importance of Mental Health awareness and support has grown on a global scale, and there are now significant avenues for communication and expression through online platforms [20]. In order to construct inclusive support systems, it is crucial to understand and manage Mental Health variances across varied emotional circumstances. In order to help those in need in a timely manner, precise case predictions of Mental Health are paramount [21]. While self-reporting and clinical diagnosis are common components of traditional Mental Health evaluation approaches, they may miss more enhanced shifts in a person's mental health over time. Individuals sometimes conceal or underreport their Mental Health signs due to the complexities of mental health issues, which further impedes correct diagnosis [22]. Knowing how mental health develops over time is crucial for a number of reasons, including early intervention, public health planning, improved coping mechanisms, and overall life satisfaction [23]. This is because Mental Health influences an individual's cognitive, emotional, and social functioning, which in turn affects their overall well-being [24].

Depressive symptoms include an overwhelming sense of despair, a lack of motivation to do anything, impatience, trouble focusing, negative thought patterns, extreme weariness, disturbed sleep and eating habits, inexplicable bodily discomfort, and a slowing of movement. On the other hand, unique patterns in text data are often indicators of depression. People who suffer from depression often have a flattened affect and less emotional expressiveness in their posts. Several linguistic characteristics allow for the identification of depressive symptoms in text data. People who are depressed tend to express themselves more negatively, using words like sad, hopeless, and worthless more frequently. Extremist words like always, never, and completely may also be more common in depressed people's texts. Shorter sentences and a simpler vocabulary could also be signs of cognitive exhaustion and trouble focusing

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

in their writing. The individual's written expressions reveal depressive thought patterns through the increased repetition of themes associated to loss, failure, and poor self-evaluation.

Sentiment analysis has emerged as a potent technique in the dynamic digital context for assessing and analyzing the feelings, opinions, and attitudes expressed in written content. The importance of sentiment analysis in mental health is being more recognized, despite its widespread use in marketing, consumer feedback, and social media monitoring. By delving into the tone of mental health-related information, new insights into people's emotional health can be monitored and find creative ways to help them. Evaluating the efficacy of mental health therapies has made use of sentiment analysis, which tracks changes in emotional expression through time. Based on the feelings conveyed in patient-generated content, this method offers a quantitative assessment of therapy efficacy. The proposed model architecture is shown in Figure 2.

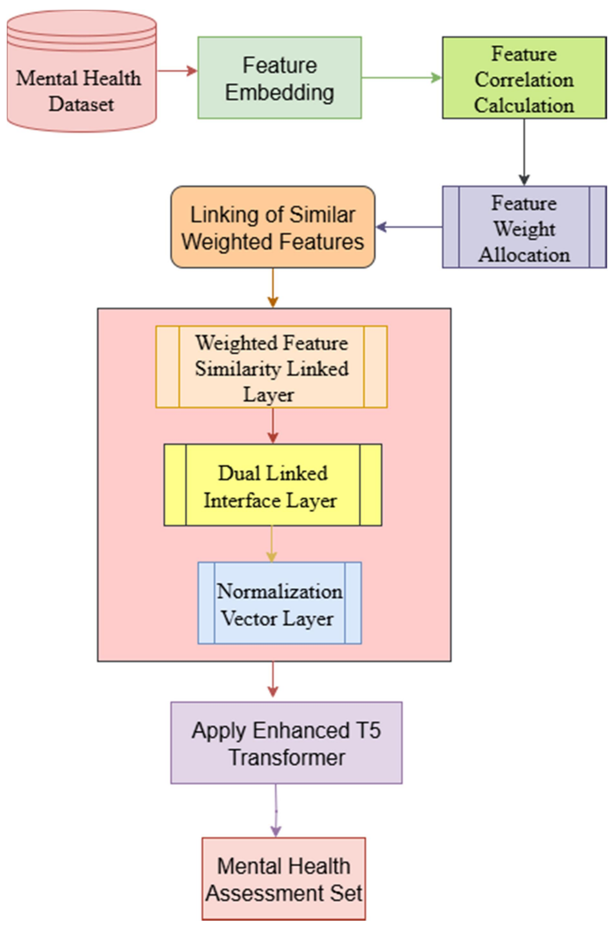

Fig 2: Proposed Model Architecture

The processing capabilities of T5 are defined by its layer configurations, which adhere to the standard encoder-decoder transformer design. Each of the two sets of blocks in the T5 model the encoders and the decoders that include a number of sublayers that process and produce text. A position-wise feed-forward network follows each encoder block, which also has a bidirectional self-attention mechanism. Layer normalization and residual connections are performed throughout. By enabling all points in the input sequence to pay attention to each other at the same time, self-attention captures intricate contextual links within the text. Masked self-attention, encoder-decoder cross-attention, and position-wise feed-forward networks are the three primary parts of T5's more intricate decoder blocks. Positions cannot attend to future positions during text production due to the masked self-attention, which preserves the autoregressive feature necessary for sequential output generation. Because of encoder-decoder cross-attention, every decoder position can pay attention to any position in the encoder's output, giving the decoder complete input context for every token it generates. The Pseudo code for the PHA-WCMVS-T5 model clearly explains the process of Mental Health Assessment.

Pseudo code: PHA-WCMVS-T5

Input:

Multilevel Variables: $T = \{V_1, V_2, \ldots, T_M\}$

Text Sequence: $T = \{T_1, T_2, \ldots, T_N\}$

Output:

- Predicted psychological health assessment $\hat{y}$

for i = 1 to M do

$r_i \leftarrow f\_emb(x_i)$

$\tilde{w}_i \leftarrow a^T \sigma(W r_i + b)$

end for

Normalize weights:

$w_i = \exp(\tilde{w}_i) / \Sigma_j \exp(\tilde{w}_j), \quad \forall i \in \{1,...,M\}$

Compute similarity:

for i = 1 to M do

for j = 1 to M do

$S_{ij} = (r_i^T r_j) / (||r_i|| \, ||r_j||)$

end for

end for

$\hat{A} = D^{(-1/2)}(S + I)D^{(-1/2)}$

Weighted embeddings:

$Z = diag(w) R, \quad Z\_link = \hat{A} Z$

for i = 1 to M do

for j = 1 to M do

$B_{ij} = r_i^T U r_j$

$\beta_{ij} = \exp(v^T [r_i ; r_j]) / \Sigma_p\Sigma\_q \exp(v^T [r_p ; r\_q])$

end for

end for

$u = \Sigma_i\Sigma_j \beta_{ij} W\_b [r_i ; r_j]$

for i = 1 to M do

$\bar{r}_i = \Sigma_j \hat{A}_{ij} r_j$

$g_i = \sigma(W\_g [r_i ; \bar{r}_i] + b\_g)$

$h_i = g_i \odot r_i + (1 - g_i) \odot \bar{r}_i$

end for

$H = \{h_1, h_2, \ldots, h\_M\}$

$Q = [u ; pool(H)]$

$y = \varphi(QP)$

$\hat{y} = LayerNorm(y)$

$\tilde{y} = Dropout(\hat{y}, p = 0.1)$

$\gamma = W\_\gamma \tilde{y}, \quad \beta = W\_\beta \tilde{y}$

for k = 1 to N do

$E'_{tk} = \gamma \odot E_{tk} + \beta$

end for

for layer ℓ = 1 to L_enc do

for head h = 1 to H do

$Q\_h = E'_t W\_Q^{(\ell,h)}$

$K\_h = E'_t W\_K^{(\ell,h)}$

$V\_h = E'_t W\_V^{(\ell,h)}$

$Attn\_h = softmax(Q\_h K\_h^T / \sqrt{d\_k}) V\_h$

end for

$MHA = Concat(Attn_1,\ldots,Attn\_H) W\_O$

$Z = LayerNorm(E'_t + Dropout(MHA))$

$Z = LayerNorm(Z + Dropout(FFN(Z)))$

end for

for layer ℓ = 1 to L_dec do

$\text{Masked\_Attn} = \text{softmax}((Q\_h\ K\_h^T + M) / \sqrt{d\_k})\ V\_h$

$\text{Cross\_Attn} = \text{softmax}(Q\times\_h\ (K\times\_h)^T / \sqrt{d\_k})\ V\times\_h$

$Y = \text{LayerNorm}(Y + \text{Dropout}(\text{Masked\_Attn}))$

$Y = \text{LayerNorm}(Y + \text{Dropout}(\text{Cross\_Attn}))$

$Y = \text{LayerNorm}(Y + \text{Dropout}(\text{FFN}(Y)))$

end for

$O = Y\ W\_lm + b\_lm$

$p(y\_k \mid y\_{<k}, t, x) = \text{softmax}(O\_k)$

Loss:

$L\_T5 = -\Sigma \log p(y^*\_k \mid y^*\_{<k}, t, x)$

$L = L\_T5 + \lambda_1 \|w\|_1 + \lambda_2 w^T L w$

$\hat{y} = \text{argmax}\_y\ p(y \mid t, x)$

end if

Return $\hat{y}$

Adding three more architectural layers to the ordinary T5 transformer increases its capabilities to process a Weighted Coupled Multilevel Variable Set. To ensure that the multivariate and strongly linked features of mental health assessment are correctly represented before being fed into the T5 encoder, these layers bridge the gap between raw psychological data and the transformer. The first component is a Weighted Feature Similarity Linked Layer (WFSL2) that factors in emotional aspects, among others, on several levels. In this layer, the relative importance of each feature variable in affecting mental health is accounted for by their weights. For example, signs of anxiety or stress may be more heavily weighted than factors related to daily habits. This makes sure that important traits have a bigger impact while encoding. This layer improves the T5 model's initial representation by producing a weighted embedding matrix.

A Dual Linked Interface Layer (DLIL) is another enhancement and it is integrated before the encoder stack. Its purpose is to capture interactions and nonlinear relationships between feature variables that conventional self-attention can miss. For instance, there may be a strong correlation between stress and emotional stability and academic achievement, illustrating the complicated ways in which psychological health issues can interact. This layer creates joint embeddings that represent the interdependencies and individual relevance of variables by explicitly modeling such couplings. Dual Linking is done between DLIL and WFSL2 and the Normalization Vector Layer (NVL).

The connected embeddings are kept computationally efficient and redundancy-free with an NVL. This layer reduces the dimensionality of the enriched embeddings while keeping the most useful features, as coupling variables increase dimensionality. By standardizing the representations and further stabilizing the training process, normalization prepares the data for the T5 encoder. With these updates, T5's encoder-decoder system may now generate more trustworthy psychological health evaluations than before by simplifying nonlinear interactions, drawing attention to important variables, and decreasing noise.
The attention mechanisms in each T5 block use 8 attention heads with 4 more newly added heads, which enables the model to concentrate on multiple relationship types at once. A feed-forward network can internally expand to 3072 dimensions before projecting back down. Because of this extension, the model can learn efficient non-linear transformations despite their complexity. The model uses a conventional dropout rate of 0.1 to minimize overfitting and applies dropout at several points throughout the design, including feed-forward networks, skip connections, and attention weights.

This research proposes a Psychological Health Assessment Model using Weighted Coupled Multilevel Variable Set with T5 Transformer (PHA-WCMVS-T5) for accurate health assessment.

Algorithm PHA-WCMVS-T5

Input: Multilevel Variables: $V = \{V_1, V_2, \ldots, V_M\}$ and Text Sequence: $T = \{T_1, T_2, \ldots, T_N\}$

Output: Predicted Psychological Health Assessment $\{PHA_{set}\}$

Consider the dataset with Variables set V and Text Sequence represented as T.

The features are extracted from the dataset and the feature embedding is performed to learn the priorities and feature values range.

$$FeatEmbed[M] = \sum_{f=1}^{M} getattr(T_f) + f_{embd}(V_f) \in \{V, T\}$$

Here T is the text sequence, V is the variable se, f is considered as the current feature.

For each features, attribute ranges are calculated and the correlation calculation is performed to identify the range of dependency of features that is performed as

$$f_{corr}[M] = \sum_{i=1}^{M} \frac{\sum(T_i - \overline{T})\,(V_i - \overline{V})}{\sqrt{\sum(T_i - \overline{T})^2 \sum(V_i - \overline{V})^2}} + \frac{\sum \frac{\max(V_i)}{T}}{\sum \frac{\min(T_i)}{V}} + \sqrt{\frac{maxVal(FeatEmbed(T, V)}{len(V)}}$$

The correlation calculation model finds the correlation factor of one feature with the other. The feature weight allocation is performed based on the correlation factor and feature embedding ratio. The process is performed as

$$\omega[M] = \frac{\min(f_{corr}(f, f+1))}{M} + \frac{\exp(f_{embd}(T, T+1))}{\sum_{f=1}^{M} \exp(\overline{T})} + Wei_f(FeatEmbed(f)) + b_f$$

The WFSL2 performs linking of similar features with common weights that represents a layer for feature processing. The linking of similarly weighted features is performed as

$$Sim_{max,min}[M] = \sum_{f=1}^{M} \frac{\max(\omega(V))}{R} + \cos(\frac{\omega_{max}(f)}{\omega_{min}(f)})$$

$$S_{Link}[M] = \gamma((Sim_{max}(f) + \omega_{max}(f)), (Sim_{max}(f+1) + \omega_{max}(f+1))$$

The DLIL process establishes the dual link among the similar weighted features and gated fusion in the feature set. The process is initiated as

$$\beta_{P,Q}[M] = \frac{\exp(V^T(S_{Link(P)}, S_{Link(Q)}))}{\sum_{f=1}^{M} \log(\max(\omega(V)) + ||V_P * V_Q||}$$

The gated fusion in the feature set is calculated as

$$g_{fus}[\beta] = \sum_{f=1}^{M} (P_f \odot Q_f) + 1 - \frac{\sum_{f=1} \log(\beta(f, f+1))}{\overline{T} \odot \overline{V}}$$

The dual linking of the feature set with the fused features is performed to consider the final feature vector as

$$D_{Link}[M] = \prod_{f=1}^{M} \max(g_{fus}(f) + \frac{\max(\beta(P, Q))}{\sum_{f=1}^{M} \sum_{i=1}^{f} \exp(\max(g_{fus})) + \log(\max(\omega(V))}$$

The Normalization Vector Layer is used to scale down the feature values for easy process and to reduce the computational time. The NVL processing is performed as

$$Fdim_{red}[M] = \sum_{f=1}^{M} \phi(D_{Link}(f, f+1) + \max(g_{fus}(f))$$

$$f_{Norm}[Fdim_{red}[M]] = \mu(D_{Link}(P, Q)) + \tau(f)$$

These NVL feature vectors are embedded intoT5 token embeddings

$$\delta = \sum_{f=1}^{M} f_{Norm}(f) + \max(w(f, f+1))$$

$$\eta = \sum_{f=1}^{M} f_{Norm}(f) + \max(g_{fus}))$$

The T5 encode self attenuation, adding of extra residual and head attenuations is performed as

$$T_h[M] = H_t + W^k(\ell, h) + \delta$$

$$V_h[M] = H_t + W^R(\ell, h) + \eta$$

$$D_h[M] = H_t + \mathrm{W}^{\mathrm{C}}(\ell, \mathrm{h})$$

$$atten_h = softmax\left(\frac{T_h V_h}{\sqrt{d_k}}\right) * D_h + (\delta + \eta) + B$$

The T5 decoder attenuation mechanism is performed as

$$atten_h[M] = softmax\left(\frac{\bar{T}_h \bar{D}_h}{\sqrt{d_k}}\right) * V_h + B$$

The final mental health assessment is performed as

$$PHA_{set}[M] = \sum_{f=1}^{M} atten_h(f) + \max\big(f_{Norm}(f, f+1)\big) + \max\big(S_{Link}(f, f+1)\big) + \frac{\exp(\max(g_{fus}))}{log\,(\max(\omega(V)))}$$

4. RESULTS

The proposed model is implemented in Python and executed in Google Colab. The dataset is considered from the link https://www.kaggle.com/datasets/sonia22222/students-mental-health-assessments and https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media. The obtained dataset is used to assess the suggested hybrid transformer design. There was no overfitting problem because the model training was smooth. As the number of epochs increases, the training and validation accuracy curves show a progressive improvement. The training was halted after 250 epochs since the curves had stabilized. In comparison, the loss curve shows inconsistent behavior, beginning with high values and then settling after 250 epochs. As the number of epochs increases, these curves show that the model is learning efficiently.

Using a dataset that was randomly split 80:20 between a training set and a test set, the performance and training dynamics of the Transformer model is thoroughly tested. The evaluation's robustness and reliability were guaranteed by regularly employing this division over numerous training iterations. Every part of the test proves that the Transformer is better than the traditional models at representing complicated data. When it comes to difficult predicting tasks, particularly those involving mixed data types, the Transformer's potential as an excellent tool is highlighted by its exceptional performance across multiple evaluation parameters. With its sophisticated attention mechanisms, the model may dynamically select data segments according to their task-specific importance. By directing processing power to the most illuminating portions of the data, this feature increases the model's efficiency while simultaneously increasing accuracy. In addition, the T5 Transformer model can adjust to both the size and complexity of the data it processes with its robust learning dynamics. The model can process numerous relationships concurrently with a multi-head attention architecture, which allows for a more accurate depiction of the input features.

One other significant improvement is how easily the T5 Transformer model can be understood. The Transformer's attention mechanism makes it easy to see which features are impacting predictions right now, in contrast to many deep learning models where decision-making processes can be opaque. This research proposes a Psychological Health Assessment Model using Weighted Coupled Multilevel Variable Set with T5 Transformer (PHA-WCMVS-T5) for accurate health assessment. The proposed model is compared with the traditional Next-Generation Text Summarization: A T5-LSTM FusionNet Hybrid Approach for Psychological Data (T5-LSTM), Socially Aware Synthetic Data Generation for Suicidal Ideation Detection Using Large Language Models (SID-LLM) and Hybrid Deep Learning Model for Predicting Depression Symptoms From Large-Scale Textual Dataset (HDL-LSTD).

The feature extraction accuracy for record sizes varying from 10,000 to 60,000 using four different models: PHA-WCMVS-T5, T5-LSTM, HDL-LSTD, and SID-LLM is shown in the Table and Figure. The most accurate and stable method is PHA-WCMVS-T5, which maintains an accuracy level over 98.5% throughout and peaks at about 99% with 50,000 recordings. Starting at around 95%, T5-LSTM drops to around 93% after 50,000 records, and then makes a small comeback to 94.2% after 60,000. Across all record sizes, HDL-LSTD keeps an accuracy that is quite consistent, ranging from 92.6% to 93.5% with small fluctuations. On the other hand, SID-LLM is the worst performer; it starts at 91.8%, drops to 90.2% at 20,000 records, peaks at 93% at 40,000 recordings, and then drops back down to about 91% as the record counts go up. In every case, PHA-WCMVS-T5 manages to significantly surpass the other models.

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Feature Extraction Accuracy (%)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	99.14	94.94	91.75	92.59
20000	98.64	94.05	90.24	92.6
30000	98.57	93.44	91.59	92.9
40000	98.96	93.18	92.95	93.55
50000	99.16	93.03	90.73	93.17
60000	98.81	94.19	90.7	92.83

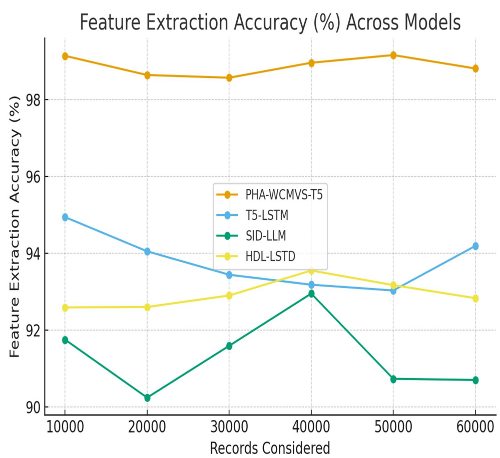

Fig 3: Feature Extraction Accuracy (%)

Using record sizes ranging from 10,000 to 60,000, the graph compares the Similar Feature Linking Accuracy (%) for four models: PHA-WCMVS-T5, T5-LSTM, SID-LLM, and HDL-LSTD. With little fluctuations between 98.2% and 98.8%, PHA-WCMVS-T5 consistently achieves the greatest and most steady accuracy, which is above 98%. Starting at 91.3%, T5-LSTM reaches a peak of 92.5% for 40,000 records, and then slightly drops to 91.9% for 60,000 records, demonstrating reasonable performance. At 10,000, SID-LLM's performance is 90.7%; at 30,000, it's 92.8%; and at 60,000, it drops dramatically to 89.5%. From 91% to 93%, HDL-LSTD maintains a consistent range, with a maximum result of 93% for 60,000 recordings. Over all record sizes, PHA-WCMVS-T5 routinely beats the competition.

Table 2: Similar Feature Linking Accuracy (%)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	98.87	91.28	90.71	90.95
20000	98.39	92.3	92.57	91.24
30000	98.43	92.29	92.78	92.6
40000	98.27	92.48	91.01	90.75
50000	98.35	92.28	91.5	91.45
60000	98.76	91.94	89.46	92.96

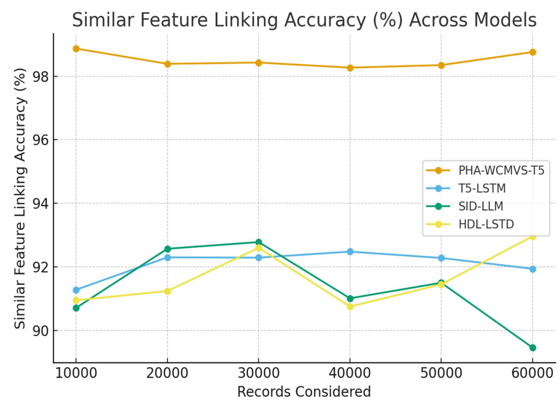

Fig 4: Similar Feature Linking Accuracy (%)

The Table and Figure shows the Dual Linking Time (in seconds) for four models, ranging from 10,000 to 60,000 records, using record sizes such as PHA-WCMVS-T5, T5-LSTM, SID-LLM, and HDL-LSTD. With times ranging from 1.95s to 2.75s, PHA-WCMVS-T5 is the most efficient model since it delivers the lowest and most steady times. The performance of T5-LSTM is moderate, being within the range of 4.95s to 5.8s. Starting at 10,000, SID-LLM reports the highest times (6.35s), which peak at 50,000 (6.65s), and then drop to 60,000 (5.25s). The instability is indicated by HDL-LSTD, which starts at 6.1s, drops to 5.1s at 30,000, and then jumps to 6.25s at 60,000. In terms of continuously low dual linking time, PHA-WCMVS-T5 performs better than the others, although SID-LLM and HDL-LSTD are the slowest.

Table 3: Dual Linking Time (s)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	2.42	5.01	6.34	6.1
20000	2.54	4.96	6.17	5.21
30000	1.94	5.13	6.25	5.13
40000	2.79	5.52	6.35	5.77
50000	2.52	5.81	6.68	4.96
60000	2.45	5.27	5.17	6.31

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

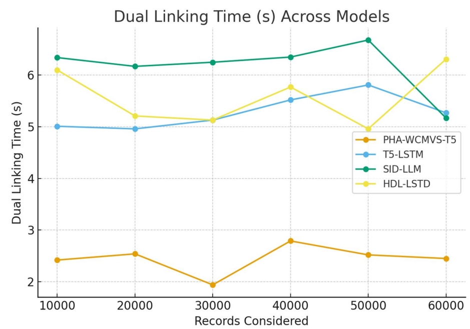

Fig 5: Dual Linking Time (s)

The graph shows Normalization Accuracy (%) for four models—PHA-WCMVS-T5, T5-LSTM, SID-LLM, and HDL-LSTD—from 10,000 to 60,000 records. PHA-WCMVS-T5 performs best across all data sizes with accuracy between 98.6% and 99.2%. T5-LSTM starts at 92.5%, peaks at 94.1% at 20,000, and stabilizes near 93.2% at 60,000. Between 92% and 93.4%, HDL-LSTD is slightly lower. SID-LLM has the lowest accuracy, starting at 93.7%, dropping to 90.2% at 40,000, and finishing at 91.6%. PHA-WCMVS-T5 performs best, whereas SID-LLM is least reliable.

Table 4: Normalization Accuracy (%)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	99.19	92.46	93.7	93.11
20000	99.3	94.09	93.37	93.99
30000	98.52	92.96	91.43	92.07
40000	99.15	94.08	90.17	93.29
50000	99.11	93.66	91.22	92.78
60000	98.67	93.17	91.59	93.08

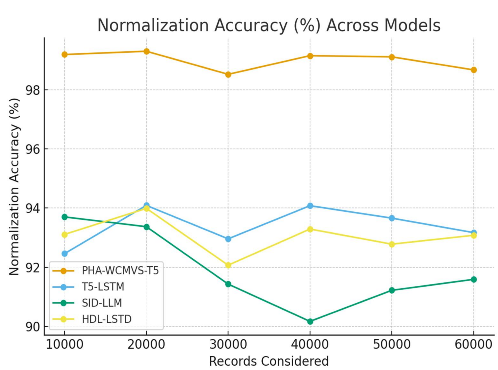

Fig 6: Normalization Accuracy (%)

The graph shows T5 Transformer Feature Processing Accuracy (%) for four models—PHA-WCMVS-T5, T5-LSTM, SID-LLM, and HDL-LSTD—from 10,000 to 60,000 records. PHA-WCMVS-T5 has the best accuracy, 98.7% to 99.3%. T5-LSTM starts at 92.8%, rises to 94.6% at 50,000, and falls to 93.7%. HDL-LSTD starts at 92%, drops to 91% at 50,000, and rises to 94% by 60,000. SID-LLM performs worst, starting at 93.2%, falling to 89.9% at 50,000, then returning to 93% at 60,000. PHA-WCMVS-T5 outperforms all models in processing accuracy.

Table 5: T5 Transformer Feature Processing Accuracy (%)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	98.85	92.32	93.27	92.02
20000	99.1	92.39	90.92	92.66
30000	98.94	92.97	90.58	92.74
40000	99.04	93.98	90.77	92.56
50000	99.21	94.54	89.86	91.01
60000	99.37	93.66	93.15	93.97

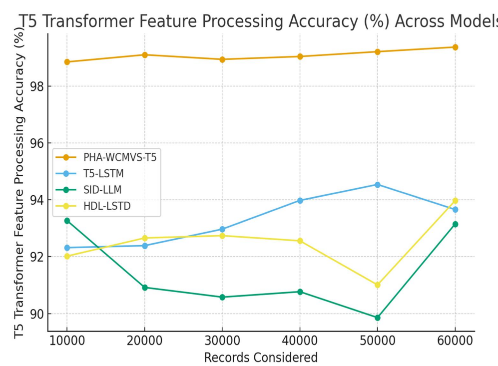

Fig 7: T5 Transformer Feature Processing Accuracy (%)

Mental Health Assessment Accuracy (%) for four models—PHA-WCMVS-T5, T5-LSTM, SID-LLM, and HDL-LSTD—from 10,000 to 60,000 records is shown in the graph. PHA-WCMVS-T5 had the highest accuracy, starting at 99.8%, dropping to 99% around 20,000–30,000 records, then returning to 99.8% at 60,000. Between 93.7% and 95.4%, T5-LSTM peaks at 30,000 records. SID-LLM starts at 92.4%, peaks at 93.9% at 30,000, then decreases to 90.8% at 40,000 before concluding at 90.8%. Starting at 92.3%, HDL-LSTD drops to 91.2%, then rises to 93.8% at 50,000. PHA-WCMVS-T5 surpasses all models in accuracy.

Table 6: Mental Health Assessment Accuracy (%)

Records	PHA-	T5-	SID-	HDL-

Considered	WCMVS-T5 Model	LSTM Model	LLM Model	LSTD Model
10000	99.71	93.77	92.4	92.35
20000	99.01	94.69	93.46	92.64
30000	99.09	95.42	93.93	91.28
40000	99.32	94.18	90.32	91.89
50000	99.7	95.19	91.71	93.78
60000	99.78	93.48	90.82	92.71

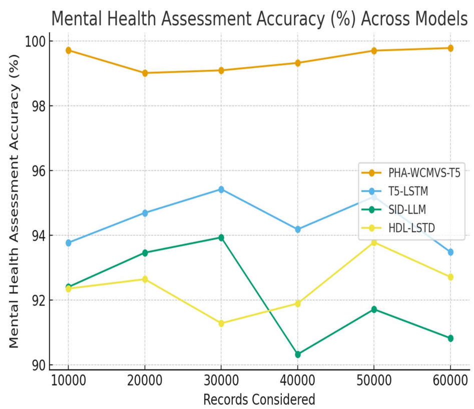

Fig 8: Mental Health Assessment Accuracy (%)

Over datasets ranging from 10,000 to 60,000 records, the graph displays the Precision (%) of four models: PHA-WCMVS-T5, T5-LSTM, SID-LLM, and HDL-LSTD. The most accurate and consistent results are obtained by PHA-WCMVS-T5, which begins at 98.7%, reaches 99.1% at 20,000 records, dips slightly to 98.1% at 40,000, and finally reaches 99.2%. T5-LSTM reaches its maximum at 20,000 records and fluctuates between 92.4% and 94.3%. HDL-LSTD exhibits significant variation, beginning at 92.7%, falling to 91.3%, and then reaching 93.8%, in contrast to SID-LLM, which varies from 91.8% to 93.6%. When it comes to accuracy, PHA-WCMVS-T5 is the clear winner, regardless of the record size.

Table 7: Precision (%)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	98.69	92.5	92.66	92.7
20000	99.13	94.34	93.55	91.25
30000	99.11	92.86	92.79	92.75
40000	98.07	92.92	91.76	93.44
50000	99.06	94.0	91.75	92.01
60000	99.26	92.33	93.06	93.78

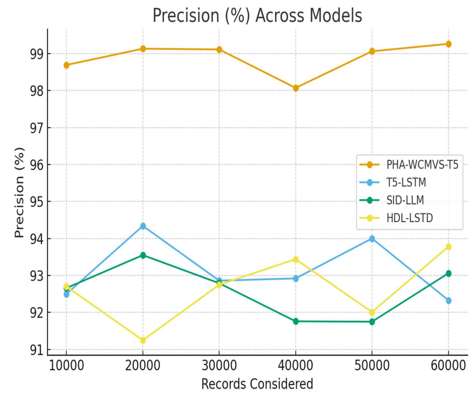

Fig 9: Precision (%)

Using datasets ranging from 10,000 to 60,000 records, the graph displays the Recall (%) of four models: PHA-WCMVS-T5, T5-LSTM, SID-LLM, and HDL-LSTD. From a low of 99.1% to a steady range of 98.9% to 99.2%, PHA-WCMVS-T5 always manages to get the best recall. T5-LSTM demonstrates substantial fluctuation, starting at 94.7% and down to 92.3% before ending at 93.2%. HDL-LSTD begins at 91.0%, reaches a peak of 93.8%, and ends at 93.2%, whereas SID-LLM ranges from 90.5% to 93.3%. Regardless of the record size, PHA-WCMVS-T5 consistently achieves superior recall compared to its competitors.

Table 8: Recall (%)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	99.13	94.73	92.25	91.02
20000	98.86	92.39	90.49	92.65
30000	99.13	92.25	90.81	93.8
40000	98.12	92.42	93.25	92.75
50000	99.29	93.2	91.87	91.62
60000	99.23	93.27	93.23	93.15

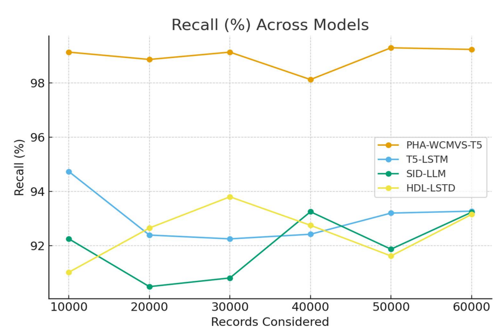

Fig 10: Recall (%)

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The error rate of the proposed model is compared with the traditional models and the Table and Figure represents that the proposed model error rate is minimum when compared to the traditional models.

Table 9: Error Rate (%)

Records Considered	PHA-WCMVS-T5 Model	T5-LSTM Model	SID-LLM Model	HDL-LSTD Model
10000	0.88	6.42	8.31	7.88
20000	1.17	5.37	7.72	7.06
30000	0.53	6.63	6.31	6.78
40000	1.14	5.2	8.1	7.86
50000	0.53	6.96	7.98	6.73
60000	1.24	7.99	6.15	6.94

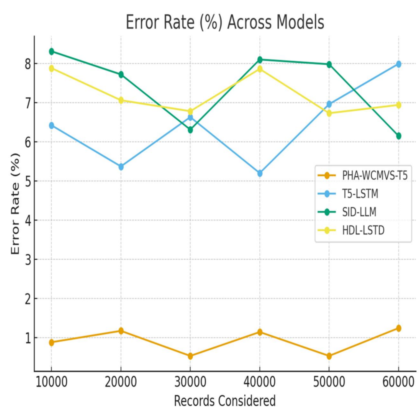

Fig 11: Error Rate (%).

4.1 Critical Analysis and Limitations

Although the proposed PHA-WCMVS-T5 model achieved superior performance across all evaluation metrics, several limitations should be acknowledged. First, the model was evaluated using publicly available Kaggle datasets, which may not fully represent the diversity of real-world psychological health conditions across different demographic, cultural, and clinical populations. Consequently, the generalizability of the proposed framework requires further validation using larger, multi-institutional datasets. Second, the incorporation of the WFSL2, DLIL, and NVL layers improves prediction accuracy but also increases architectural complexity, resulting in higher computational requirements during model training. Third, the proposed framework primarily relies on textual and multilevel variable information, whereas other important modalities such as speech, facial expressions, physiological signals, and wearable sensor data were not considered. Moreover, although the T5 Transformer provides strong contextual learning capabilities, the overall model remains relatively difficult to interpret from a clinical perspective, which may limit its direct adoption in healthcare environments where explainability is essential. Future research will therefore focus on improving model interpretability through explainable AI techniques, reducing computational complexity, and validating the framework using large-scale multimodal clinical datasets collected from real healthcare settings.

5. CONCLUSION

This research presented a novel Psychological Health Assessment Model using a Weighted Coupled Multilevel Variable Set with T5 Transformer (PHA-WCMVS-T5) for accurate and efficient psychological health assessment. The novelty of the proposed work lies in integrating three dedicated architectural components—Weighted Feature Similarity Linked Layer (WFSL2), Dual Linked Interface Layer (DLIL), and Normalization Vector Layer (NVL)—with the T5 Transformer to effectively model multilevel psychological variables and their nonlinear interdependencies. Unlike conventional CNN-, LSTM-, and transformer-based methods that primarily focus on text representation or single-task prediction, the proposed framework jointly performs adaptive feature weighting, coupled feature interaction modeling, and contextual language understanding within a unified architecture. Experimental evaluation demonstrated superior performance over T5-LSTM, SID-LLM, and HDL-LSTD, achieving approximately 99% mental health assessment accuracy, high precision and recall, and a significantly lower error rate. The proposed model has considerable practical impact by providing a scalable, accurate, and computationally efficient framework for early psychological health assessment, enabling timely intervention and supporting intelligent decision-making in healthcare and educational environments. These contributions establish the proposed framework as an effective solution for next-generation AI-assisted psychological health assessment. By utilizing a Weighted Coupled Multilevel Variable Set with T5 Transformer, the PHA-WCMVS-T5 outperforms more conventional models such as T5-LSTM, SID-LLM, and HDL-LSTD when it comes to properly assessing and forecasting mental health disorders. The model reliably obtained feature extraction accuracies over 98.5%, reaching a maximum of 99.3%, across datasets with 10,000 to 60,000 entries. Compared to T5-LSTM (95.4%), HDL-LSTD (93.8%), and SID-LLM (93.9%), the PHA-WCMVS-T5 model consistently outperformed the others in terms of mental health assessment accuracy, which was maintained near 99.8%. Recall rates ranged from 98.9% to 99.2% and accuracy values from 98.1% to 99.2%, both of which demonstrate the model's reliability and robustness. The model's stability and accuracy were further demonstrated by its lowest error rates, which ranged from 0.53% to 1.24%. Its scalability for real-time applications was demonstrated by its smooth convergence in training dynamics, stable learning after 250 epochs without overfitting, and efficient dual linking times of 1.94-2.79 seconds. The model accurately captures nonlinear interactions within multimodal data by integrating improved attention processes and multilayer feature coupling. This enables high-precision psychological assessments. To better suit the needs of real-world healthcare systems, future studies will aim to make the PHA-WCMVS-T5 model more private and adaptable. Data sharing and collaboration between institutions can be made secure without sacrificing patient confidentiality by integrating homomorphic encryption and other privacy-preserving approaches. To make psychological health assessments more accurate and less biased by demographics or culture, it is necessary to train the model on bigger, more varied, and multilingual datasets. It is possible to gain a more complete picture of mental health issues by combining data from several sources, such as physiological signals, patterns of behavior, and audio or video inputs.

This study makes a significant contribution to the existing body of knowledge by extending transformer-based psychological health assessment beyond conventional text classification and prediction approaches. Unlike previously published studies that primarily employ CNN-, LSTM-, or standard transformer-based architectures for isolated mental health tasks, the proposed PHA-WCMVS-T5 framework introduces a unified methodology that combines weighted multilevel feature representation, coupled variable interaction modeling, and contextual language understanding within a single architecture. The integration of the Weighted Feature Similarity Linked Layer (WFSL2), Dual Linked Interface Layer (DLIL), and Normalization Vector Layer (NVL) demonstrates an effective strategy for modeling complex nonlinear psychological relationships while improving prediction accuracy and computational efficiency. These findings provide new insights into designing transformer-based frameworks capable of handling heterogeneous psychological data and establish a foundation for future research on intelligent, multimodal, and explainable psychological health assessment systems.

This study significantly advances existing knowledge by introducing the proposed PHA-WCMVS-T5 framework, which integrates weighted multilevel feature representation with the T5 Transformer for psychological health assessment. Unlike existing CNN-, LSTM-, and standard transformer-based methods, the proposed approach effectively models coupled psychological variables, resulting in improved accuracy, robustness, and computational efficiency. These findings provide a strong foundation for future intelligent and explainable mental health assessment systems.

REFERENCES

[1] B. Khan, M. Usman, I. Khan, J. Khan, D. Hussain and Y. H. Gu, "Next-Generation Text Summarization: A T5-LSTM FusionNet Hybrid Approach for Psychological Data," in IEEE Access, vol. 13, pp. 37557-37571, 2025, doi: 10.1109/ACCESS.2025.3540590.

[2] H. Ghanadian, I. Nejadgholi and H. A. Osman, "Socially Aware Synthetic Data Generation for Suicidal Ideation Detection Using Large Language Models," in IEEE Access, vol. 12, pp. 14350-14363, 2024, doi: 10.1109/ACCESS.2024.3358206.

[3] Y. Ma, X. Sun and A. Ma, "PSG-Pair: A Psychological Snapshot Guided Model for Improved Query-Response Pairing in Psychological Wellness Communities for University Students," in IEEE Access, vol. 13, pp. 41505-41523, 2025, doi: 10.1109/ACCESS.2025.3545902.

[4] S. Almutairi, M. Abohashrh, H. H. Razzaq, M. Zulqarnain, A. Namoun and F. Khan, "A Hybrid Deep Learning Model for Predicting Depression Symptoms From Large-Scale Textual Dataset," in IEEE Access, vol. 12, pp. 168477-168499, 2024, doi: 10.1109/ACCESS.2024.3496741.

[5] H. -S. Chiang, M. -Y. Chen and L. -S. Liao, "Cognitive Depression Detection Cyber-Medical System Based on EEG Analysis and Deep Learning Approaches," in IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 2, pp. 608-616, Feb. 2023, doi: 10.1109/JBHI.2022.3200522.

[6] M. M. Hossain, Sanjara, M. S. Hossain, S. Chaki, M. S. Rahman and A. B. M. S. Ali, "Revolutionizing Mental Health Sentiment Analysis With BERT-Fuse: A Hybrid Deep Learning Model," in IEEE Access, vol. 13, pp.

85428-85446, 2025, doi: 10.1109/ACCESS.2025.3568340.

[7] U. Ahmed, J. C. -W. Lin and G. Srivastava, "Hyper-Graph Attention Based Federated Learning Methods for Use in Mental Health Detection," in IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 2, pp. 768-777, Feb. 2023, doi: 10.1109/JBHI.2022.3172269.

[8] M. E. Villa-Pérez, L. A. Trejo, M. B. Moin and E. Stroulia, "Extracting Mental Health Indicators From English and Spanish Social Media: A Machine Learning Approach," in IEEE Access, vol. 11, pp. 128135-128152, 2023, doi: 10.1109/ACCESS.2023.3332289.

[9] F. Benrouba and R. Boudour, Emotional sentiment analysis of social media content for mental health safety, Soc. Netw. Anal. Min. 13(1) (2023) 17. Crossref, Google Scholar

[10] A. Zunic, P. Corcoran and I. Spasic, Sentiment analysis in health and well-being: Systematic review, JMIR Med. Inf. 8(1) (2020) e16023. Crossref, Google Scholar

[11] J. Park, M. H. Tsou, A. Nara, S. Cassels and S. Dodge, Developing a social sensing index for monitoring place-oriented mental health issues using social media (Twitter) data, Urban Inf. 3(1) (2024) 2. Crossref, Google Scholar

[12] S. Khan and S. Alqahtani, Hybrid machine learning models to detect signs of depression, Multimed. Tools Appl. 83 (2023) 1–19. Crossref, Google Scholar

[13] C. T. Oyewale, J. D. Akinyemi, A. O. Ibitoye and O. F. Onifade, Predicting Suicide Ideation from Social Media Text Using CNN-BiLSTM, in K. K. Patel, K. Santosh, A. Patel, A. Ghosh (eds.), Soft Computing and Its Engineering Applications. icSoftComp 2023, Communications in Computer and Information Science, Vol. 2030 (Springer, Cham, 2024), pp 274–286. https://doi.org/10.1007/978-3-031-53731-8_22 Crossref, Google Scholar

[14] A. De and S. Mishra, Augmented intelligence in mental health care: Sentiment analysis and emotion detection with health care perspective, in Augmented Intelligence in Healthcare: A Pragmatic and Integrated Analysis (Springer, Singapore, 2022), pp. 205–235. Google Scholar

[15] M. Colizzi, A. Lasalvia and M. Ruggeri, Prevention and early intervention in youth mental health: Is it time for a multidisciplinary and trans-diagnostic model for care? Int. J. Mental Health Syst. 14 (2020) 1–14. Crossref, Google Scholar

[16] M. Salvi, H. W. Loh, S. Seoni, P. D. Barua, S. García, F. Molinari and U. R. Acharya, Multi-modality approaches for medical support systems: A systematic review of the last decade, Inf. Fusion (2023) 102134. Google Scholar

[17] R. Abd Rahman, K. Omar, S. A. M. Noah, M. S. N. M. Danuri and M. A. Al-Garadi, Application of machine learning methods in mental health detection: A systematic review, IEEE Access, 8 (2020) 183952–183964. Crossref, Google Scholar

[18] Ilias L, Askounis D (2022) Explainable identification of dementia from transcripts using transformer networks. IEEE J Biomed Health Inform 26(8):4153–4164. https://doi.org/10.1109/JBHI.2022.3172479

[19] Ilias L, Askounis D (2023) Multitask learning for recognizing stress and depression in social media. Online Social Netw Media 37–38:100270. https://doi.org/10.1016/j.osnem.2023.100270

[20] Ilias L, Mouzakitis S, Askounis D (2023) Calibration of transformer-based models for identifying stress and depression in social media. IEEE Trans Comput Social Syst. https://doi.org/10.1109/TCSS.2023.3283009

[21] Ji Y, Pan S, Liu Q, Zhou H (2022) Mentalbert: Publicly available pretrained language models for mental healthcare. arXiv preprint arXiv:2208.03156

[22] Kabir M, Ahmed T, Hasan MB, Laskar MTR, Joarder TK, Mahmud H, Hasan K (2023) DEPTWEET: A typology for social media texts to detect depression severities. Comput Human Behav 139:107503. https://doi.org/10.1016/j.chb.2022.107503

[23] Kamath R, Ghoshal A, Eswaran S, Honnavalli P (2022) An enhanced context-based emotion detection model using RoBERTa. In: 2022 IEEE international conference on electronics, computing and communication technologies (CONECCT), pp. 1–6. https://doi.org/10.1109/CONECCT55679.2022.9865796

[24] Kamite SR, Kamble VB (2020) Detection of depression in social media via twitter using machine learning approach. In: 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC), pp. 122–125. https://doi.org/10.1109/ICSIDEMPC49020.2020.9299641

[25] Malviya K, Roy B, Saritha SK (2021) A transformers approach to detect depression in social media. In: 2021 international conference on artificial intelligence and smart systems (ICAIS), pp. 718–723. https://doi.org/10.1109/ICAIS50930.2021.9395943
[26] Naseem U, Dunn AG, Kim J, Khushi M (2022) Early identification of depression severity levels on reddit using ordinal classification. In: Proceedings of the ACM Web Conference 2022. WWW ’22, pp. 2563–2572. Association for Computing Machinery, New York https://doi.org/10.1145/3485447.3512128
[27] Parapar J, Martín-Rodilla P, Losada DE, Crestani F (2021) Overview of erisk 2021: Early risk prediction on the internet. In: Candan KS, Ionescu B, Goeuriot L, Larsen B, Müller H, Joly A, Maistro M, Piroi F, Faggioli G, Ferro N (eds) Exp IR Meet Multiling Multimodality Interact. Springer, Cham, pp 324–344
[28] Rao G, Zhang Y, Zhang L, Cong Q, Feng Z (2020) MGL-CNN: A hierarchical posts representations model for identifying depressed individuals in online forums. IEEE Access 8:32395–32403. https://doi.org/10.1109/ACCESS.2020.2973737