IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 4, August 2026, pp. 3202~3217
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i4.pp3202-3217

Deep transfer learning architecture for depression detection using PHQ-9 analysis in Bengaluru urban population

U. Ananthanagu, Pooja Agarwal
Department of Computer Science and Engineering, PES University, Bengaluru, India

Article Info

Article history:

Received Jun 30, 2025
Revised May 21, 2026
Accepted Jul 9, 2026

Keywords:

Depression
Patient health questionnaire-9
Population-based analysis
Quantitative research
ResNet-50
Statistical analysis

ABSTRACT

In the metropolitan cities in India, there is a scarcity of population-based statistics on depression prevalence. This study utilized the patient health questionnaire-9 (PHQ-9) instrument to look at the prevalence and risk factors for depression in major metropolitan cities. This study's population-based analysis identifies major demographic and socioeconomic characteristics associated with depression in an Indian urban context. This quantitative inquiry conducted a population-centric analysis of 987 people from a metropolitan city. Depression was measured and validated using the PHQ-9 questionnaire to calculate the prevalence rates and sociodemographic characteristics. A deep transfer learning framework integrating visual and textual features was proposed for automated classification of depression, to assess mental health patterns in urban populations. Depression symptoms were common, with many respondents reporting persistent symptoms like a persistent lack of interest, pervasive sorrow, sleep difficulties, and suicidal ideation. Bivariate studies found relationships between these symptoms and other demographic and socioeconomic characteristics, whereas multivariate analysis revealed significant predictors. In addition, the classification of people's stress in metropolitan regions is analyzed through various pre-trained models to show better outcomes. The substantial link between sociodemographic characteristics and depression levels shows the importance of mental health interventions for this urban Indian population.

Corresponding Author:

U. Ananthanagu
Department of Computer Science and Engineering, PES University
Electronic City Campus, Bengaluru, Karnataka, India
Email: pes1pd20cs006@pesu.pes.edu

1. INTRODUCTION

Depression is a major mental health disorder that affects emotional well-being, cognitive function, and daily life activities of millions worldwide [1]. Common symptoms include sadness, sleep disturbances, fatigue, and loss of interest, which negatively impact personal, social, and occupational functioning [2]. Socioeconomic status, including income, education, employment, and housing conditions, plays an important role in mental health outcomes, particularly depression [3]. Factors such as education and occupation influence access to healthcare, coping abilities, and social support systems, thereby affecting psychological well-being [4].

Previous studies have reported that lower socioeconomic status is strongly associated with higher levels of depression [5]. However, most investigations were conducted in Western populations and often overlooked cultural and regional influences on mental health outcomes [6]. Cultural pressures, occupational stress, and traditional social roles may significantly influence depression patterns in different populations [7]. In addition, many earlier studies relied on conventional statistical approaches, which are limited in capturing the complex and non-linear relationships among socioeconomic and psychological factors [8], [9].

Traditional depression assessment methods mainly rely on standardized questionnaires such as the hamilton depression rating scale and beck depression inventory [10]. Although widely used in clinical and research settings, these approaches are often affected by subjective responses, literacy limitations, and respondent bias [11]. In addition, conventional statistical techniques and basic machine learning models, including support vector machines and logistic regression, have limitations in handling complex, high-dimensional, and multimodal depression data [12], [13].

Recent advances in artificial intelligence and deep learning have improved the analysis of large and complex healthcare datasets by identifying hidden patterns and non-linear relationships that traditional statistical approaches may overlook [14]. Deep learning models can effectively process high-dimensional and multimodal data, including text and images, leading to improved mental health prediction performance. Therefore, this study proposes a deep transfer learning framework for accurate depression severity classification in urban populations using patient health questionnaire-9 (PHQ-9) data [15]: i) to analyze depression prevalence in metropolitan populations using PHQ-9 data, ii) to identify demographic factors associated with depression severity, iii) to support data-driven mental health interventions and public health planning, and iv) to develop a deep transfer learning-based framework for automated depression classification using multimodal feature analysis.

Some of the recent studies related to our work are reviewed as follows. The authors [16]–[18] highlighted the increasing prevalence of depression and common mental disorders in India, emphasizing the need for early screening, mental health awareness, and effective clinical assessment. Their findings further demonstrated significant associations between depression and factors such as medical comorbidities, academic stress, anxiety, age, and socio-demographic conditions. Furthermore, Shenoy *et al.* [19] identified an elevated suicide risk among individuals with common mental disorders and chronic illnesses, underscoring the importance of targeted mental health interventions and preventive care strategies. Srinivasan *et al.* [20] reported reduced depression levels following collaborative care interventions, while the authors [21]–[23] validated the reliability and effectiveness of PHQ-9 across different demographic and cultural populations. In addition, the researchers [24]–[26] highlighted the impact of medical and sociodemographic factors on depression assessment and screening. The authors [27], [28] further demonstrated the applicability of PHQ-9 in diverse populations, emphasizing the importance of culturally adaptive mental health evaluation approaches. Uyulan *et al.* [29] developed a deep convolutional neural network (CNN)-based electroencephalography (EEG) framework using ResNet-50, MobileNet, and Inception-v3 models for major depressive disorder detection, demonstrating the effectiveness of deep learning in depression diagnosis. Similarly, Sofia *et al*. [30] applied machine learning techniques such as decision tree, k-nearest neighbors, and naive Bayes to survey-based depression assessment, highlighting the importance of early detection and automated mental health evaluation. Siddiqua *et al.* [31] developed machine learning-based approaches for depression prediction among university students using questionnaire data, while Li *et al.* [32] examined the relationship between subjective well-being and depression. In addition, Kim [33] analyzed the association between socioeconomic factors and depression using PHQ-9 data from a large population study, highlighting the influence of income, education, marital status, and work-related factors on mental health outcomes.

Previous studies primarily relied on traditional statistical approaches or basic machine learning techniques for depression analysis. Limited research has integrated validated clinical measures such as PHQ-9 with multimodal deep transfer learning frameworks that combine textual and visual features for automated depression detection. In addition, the scalability and real-world applicability of such models in urban mental health settings remain underexplored. Existing studies also provide limited attention to convolutional block attention module (CBAM)-based pretrained architectures for improving classification performance. Therefore, this study proposes multimodal deep transfer learning-based framework integrating ResNet-50, bidirectional encoder representations from transformers (BERT)-CNN, and CBAM-enhanced InceptionV3 architectures to improve depression classification accuracy and support scalable urban mental health assessment. The integration of multimodal deep transfer learning with bio-inspired optimization establishes a clinically interpretable and scalable artificial intelligence (AI)-driven framework for depression assessment and healthcare decision support.

2. METHOD

This study employs a quantitative research design that utilizes population-based methodologies. The objective was to examine the prevalence and associated factors of depression among individuals aged 17 to 40 in the city of Bengaluru, Karnataka. The sample comprised 987 respondents, selected through stratified random sampling to ensure representation across various demographic groups.

Data were collected using the PHQ-9, which evaluates depressive symptoms experienced over the preceding fortnight. The PHQ-9 comprises a series of nine inquiries rated according to a continuum of responses: "Not at all," "Several days," "More than half the days," and "Nearly every day" [34]. The responses are categorized into four distinct levels of depressive severity: normal (0-9), mild (10-13), moderate (14-20), and severe (21-27) [35].

The collected data were analyzed using SPSS major techniques. In a previously published SPSS study, a bivariate analysis was conducted to investigate the varied associations between demographic factors and depression rating. The demographic factors were gender, level of education, age, income and marital status. Crosstabs, symmetric measures, and to determine whether a correlation exists at statistically significant levels, chi-square tests were employed. Covariate effects on depression, on the other hand, were examined using multiple regression and multivariate analysis to assess all predictors while correcting for confounding variables.

2.1. Mental health analysis by pre-trained models

This proposed work presents a new framework based on deep transfer learning, which includes multiple pretrained models organized using a parallel architecture to successfully recognize depression. The model features a dual-branch ResNet-50 backbone. In the first branch, a BERT-CNN hybrid module combines contextual semantic (textual) and low-level visual information features. Advanced contextual embedding techniques were incorporated to improve semantic feature extraction from PHQ-9 responses. Future extensions may integrate transformer-based models such as RoBERTa and DistilBERT for enhanced linguistic understanding. In addition, focal loss and cost-sensitive learning strategies can improve minority class detection, particularly for severe depression categories. In the second branch of the dual-branch architecture, InceptionV3 is combined with the CBAM to identify important spatial and channel-wise features. The features from both branches are concatenated and passed through a fully connected layer for final label classification. This multimodal combination supports depression symptom analysis and promotes improved early-stage depression identification and mental health monitoring. Furthermore, the proposed method utilizes the adaptive osprey optimization algorithm (A-OOA) to optimize important hyperparameters such as epochs, batch size, and learning rate, thereby improving classification performance and convergence speed.

2.1.1. Preprocessing

Preprocessing is the first step in our proposed work, confirming that both textual and visual data are cleaned, consistent, and ready for feature extraction. It also enhances the quality of the data and allows the model to better classify depression in a more consistent way. In this work, data cleaning, normalization, and missing value imputation are performed.

i) Data cleaning: data cleaning eliminates irrelevant entries, segmented PHQ-9 responses, duplicates, and noisy visual samples. Data preparation for text data entails removing irrelevant characters, while poor samples or corrupt image files are simply excluded during data cleaning. This restricts the training set to valid, meaningful inputs for the model.
ii) Normalization: normalization guarantees the numeric features such as PHQ-9 scores or image-based values, are treated identically in training. Normalization rescales the features to a particular range that are now all in the range of [0,1]. This reduces the effect of high-range values and allows the model to learn in a more stable environment. Normalization, using the definition of narrowing range min-max normalization, is defined as (1).

$$X_n = \frac{X - min(X)}{\max(X) - min(X)} \tag{1}$$

iii) Missing value imputation: incomplete records or missing PHQ-9 scores are common in survey data. In these instances, counts of missing values are not simply recorded, and the units are not discarded. To fill in missing values that exist, the scores are filled with a statistical estimate to maintain integrity in data (use of mean or median). This method ensures there is consistent input for deep learning architecture.

2.1.2. Parallel ResNet-50 for feature extraction

After preprocessing, the output proceeds in parallel to two distinct but complementary deep transfer learning pathways ResNet with BERT along with convolutional neural network (ResBCNN), and attention-based InceptionResNet (A-IRes). In path one, a ResNet-50 model is used first to build hierarchical feature representations. The extracted features are then separated into two parallel branch submodules. The first branch implements BERT to model contextual and sequential dependencies, and the second branch implements a CNN pipeline to build latent local patterns for the input data. The outputs of BERT and CNN are fused to create an enhanced semantic representation of the original input. Branch two also begins with

ResNet-50 layers, but the next modules are an inception module to collect multi-scale convolutional features from the extracted features. To further enhance the model's attention to informative spatial patterns and channel-based patterns, a CBAM follows a skip-connection max-pooling operation to enable the model to dynamically learn and guide the most useful features for classification. The last outputs from each path, first the Resnet with BERTCNN path, and second the attention-based InceptionResNet path, can be concatenated into a weighted feature vector, yielding a final array feature vector. The combined vector then goes through a dense fully connected layer combined with the activation function sigmoid, before providing the final classification output based on severity of depression. The parallel ResNet architecture is depicted in Figure 1.

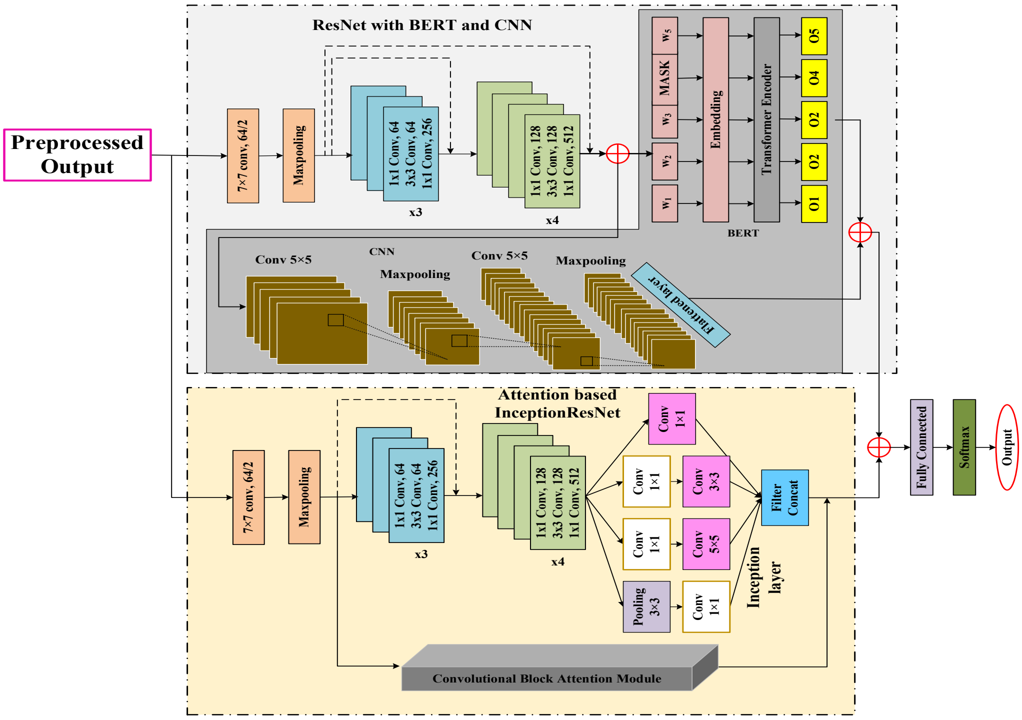


Figure 1. Architecture of parallel ResNet-50

ResNet with BERTCNN: the ResBCNN module starts with the preprocessed input applied to a convolutional layer 7×7 containing 64 filters followed by a max-pooled operation. The max-pooled operation reduces the overall spatial dimension of input data while maintaining essential local features of the overall input. The resulting output from this operation is passed through two groups of major residual blocks. Each group is made of sets of convolutions with 1×1 kernels, a 3×3 kernel, and a 1×1 kernel. The first group were performed three times (×3) and the second group was performed four times (×4). The included residual connections will efficiently share features, counter the vanishing gradient problem, and allow low-level and mid-level spatial features to easily propagate throughout the network by maximizing feature-sharing/casting capabilities. The resultant structure provides a solid backbone for early visual feature extracting as a ResNet-50 module. Once the residual layers have completed, we split the extracted features into two parallel pathways, going to a pretrained BERT model and the CNN pathway.

For the BERT pathway, the input features continue to a transformer encoder, which includes the encoder portion only of the original transformer architecture. This model learns semantic and contextual relationships in both directions, by predicting masked tokens using information from both past and future positions in a sequence. The transformer encoder is made up of several important submodules. The transformer encoder in the BERT model is made up of a collection of complex sublayers that together represent contextual relationships in the input features, as shown in Figure 2. The first component is input embedding, a layer transferring discrete tokens or features into continuous dense vectors, representing the fact that the model understood semantic similarity. The addition of positional encoding to the embeddings preserves the positional context of each token in the encodings, given that the transformer itself is not inherently sequentially aware.

The central part of the model is the multi-head self-attention mechanism, which consists of the query (Q), key (K), and value (V) matrices attending to many representation subspaces in parallel and were able to capture dependencies across the sequence. After the output of the multi-head attention, both residual connections and layer normalization (add and norm) were performed to increase stability. A feed-forward network (FFN) performs nonlinear transformations independently to each position first, followed by another add and norm layer. These components together enable BERT to effectively encode both local and global context in the same input. Simultaneously, the CNN pathway refines the representations of spatial features through the application of numerous 5×5 convolutional layers and max-pooling operations. The convolution operation takes the input matrix and performs an element-wise multiplication with the learnable filters (kernels), adds the bias and a non-linear activation function typically rectified linear unit (ReLU) to introduce non-linearity. The mathematical representations of the operations are as in (2).

$$Y = f(\sum_{i=1}^{n} X_i * W + b_i) \tag{2}$$

Where X_i is the input, W is the convolutional filter, b_i is the bias, f is the ReLU activation function, and * denotes the convolution operation.

Pooling layers, and max-pooling in particular, down sample the feature maps by selecting the maximum value per sliding window. The pooling operation effectively retains the most important features, yet reduces computational burden. At the end of the model architecture, this fuses the output from the BERT and CNN modules through a concatenation, indicating that we have combined semantic features from BERT with spatial features from the CNN. This complete representation is then passed onto the classification pipeline. Adding residual learning, BERT's semantic understanding, and a hierarchical visual feature map together strengthens the overall feature extraction process towards recognizing depression levels.

Attention-based Inception ResNet: the Attention-InceptionResNet starts extracting features from input features through a ResNet-50 backbone, which consists of a 7×7 convolution layer, followed by a max-pooling layer and a series of residual convolutional blocks with repeated sequences of three convolutions, having a 1×1, 3×3, then another 1×1 convolution layer in rows, stacked in several stages. After going through the 4th residual block, the high-level feature representation output is passed into a custom modification of the Inception module, thus, enhancing feature extraction across multiple scales. The Inception module allows for multi-path feature extraction by passing the input through multiple convolutional filters (1×1, 3×3, 5×5), and pooling would be undertaken in parallel. This allows the model to acquire both local fine-grained characteristics of the input data and larger spatial context at the same time through each of the various paths.

The outputs of the numerous branches are concatenated in the depth dimension to produce a dense and varied set of features. Additionally, the inclusion of 1×1 convolution operations assists with reducing dimensions and computational load while maintaining salient information. As a modular architecture, Inception modules enable scales of depth and width of the residual parts of ResNet, while also decreasing issues of overfitting, vanishing gradients, and memory issues that arise due to a deep architecture. This ultimately leads to greater discriminative power and greater generalization towards complex visual patterns, valuable in the context of subtle depression detection with fine-grained classification methods.

To further refine the extracted features and improve the model's focus, the CBAM is incorporated right after the Inception module. CBAM contains two attention mechanisms that act sequentially in a way that the model can focus on important features not only due to where to `attend` to, but also what to `attend` to from the output of the last three layers of the inception module. The structure of CBAM is shown in Figure 3.

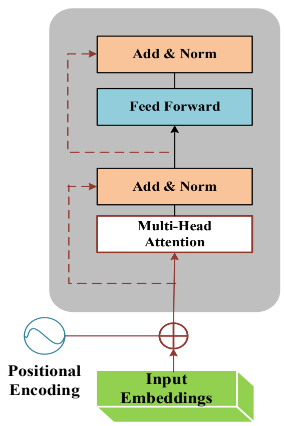


Figure 2. Structure of transformer encoder

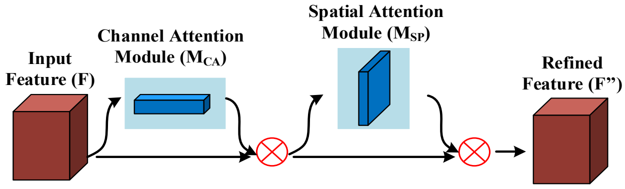


Figure 3. Convolutional block attention module

Let $F \in R^{c \times h \times w}$ be the input feature map. First, the channel attention module generates a 1D attention map $M_{CH} \in R^{c \times 1 \times 1}$ by applying both global average pooling and max pooling across spatial dimensions. These pooled descriptors F^c_{avg} and F^c_{max}, are passed through a shared multi-layer perceptron (MLP) with reduction ratio R, and their outputs are combined and activated using a sigmoid function, as defined in (3).

$$M_{CH}(F) = \sigma\left(MLP\big(AvgPool(F)\big) + MLP\big(MaxPool(F)\big)\right) \tag{3}$$

This mechanism learns which channels are the most relevant and therefore enhances their importance in later processing, as defined in (4).

$$F' = M_{CH}(F) \otimes F \tag{4}$$

Next, the spatial attention module determines where in the feature map to attend as defined in (5). It pools F' along the channel axis using average and max pooling, resulting in two 2-D maps (F^s_{avg}, F^s_{max}), which are concatenated and processed through a convolution layer with a 7×7 kernel. As a result of CBAM, the model learns context-aware and locality-sensitive features that only attend to the most salient parts of the feature map in each of the channel and spatial dimensions. This representation enhanced by attention is important for downstream classification or detection tasks such as finding depression-related patterns in high dimensional feature spaces.

$$M_{SP}(F) = \sigma\big(f^{7 \times 7}([AvgPool(F); maxPool(F)])\big) \tag{5}$$

The final refined feature map is computed as (6).

$$F'' = M_{SP}(F') \otimes F' \tag{6}$$

2.1.3. Classification

After deep feature extraction steps, the outputs of both the ResBCNN and A-IRes modules were concatenated to yield a combined feature representation that allows both molecular information (from ResBCNN) as well as multi-scale spatial features from A-IRes. The combined feature vector provides a variety of combinations of learned features containing richer and more comprehensive features that encode contextual dependencies at a higher level of semantics, while also encoding fine-grained visual minutiae. The fused feature vector is then run through a fully connected layer that alters dimensionality and is non-linear in nature, effectively maps the learned high-dimensional feature representations to be appropriate for final prediction. Dropout regularization can be applied in this layer to avoid overfitting and maintain stable learning through propagation. The final output layer applies a softmax activation function to compute the predicted probability distribution across the defined classes of depression levels. The softmax function transforms the model output values into probabilities class-wise, where the maximum predicted probability identifies the model's class label. The classification pipeline produces predictions that are both accurate and meaningful based on an integrated set of deep features, enabling recognition and classification of the values of depression levels of the specified input data.

2.1.4. Adaptive-osprey optimization algorithm

The osprey optimization algorithm is a bio-inspired metaheuristic method that is based on the predatory and adaptive foraging behavior of ospreys. There are two principal phases of the algorithm: global exploration (positioning and fishing) and local exploitation (moving the fish to a safe place). In the application of the algorithm, each solution is regarded as an individual osprey, and the algorithm iteratively updates the position of all the ospreys in the search space to minimize some predefined objective function, and, in the case of classification problems, that objective function might be the cross-entropy loss.

The cross-entropy loss function is used as the objective function to improve the classification performance of the model. Cross-entropy loss quantifies how much the predicted probability differs from the label given, and thus, strives to minimize the error in prediction as (7).

$$J_{bce} = -\frac{1}{M}\sum_{m=1}^{M}\left[y_m \times log\left(h\theta(x_m)\right) + (1 - y_m) \times log\left(1 - h\theta(x_m)\right)\right] \quad (7)$$

Where M stands for the total amount of training samples in the dataset. Each training sample is characterized by two components: y_m which denotes the target label corresponding to the training example m, and x_m which represents input features for that example. h is model prediction and where θ is model parameter.

Initialization: the algorithm begins by randomly initializing the positions of all ospreys within a defined search space. Each osprey's position $X_{i,j}$ in the $j - th$ dimension is determined as in (8).

$$X_{i,j} = lb_j + r_{i,j} \cdot \left(ub_j - lb_j\right), i = 1,2, \dots, N; j = 1,2, \dots, D \quad (8)$$

Initial position of the ith osprey ($X_{i,j}$) in the jth dimension, lower (lb_j) and upper (ub_j) bounds of the jth problem variable, respectively. $r_{i,j}$ as a random number between [0, 1], N is the population size, D is the problem's solution dimension, and j indicates the specific dimension. Each osprey serves as a candidate solution, and its quality is evaluated by the fitness value of the objective function F, as defined in (9).

$$F_i = F(X_i), i = 1,2, \dots, N \quad (9)$$

Where fitness value (F_i) of ith osprey, and the position of ith osprey (X_i).

i) Phase 1: positioning and fishing. At this stage, the osprey finds targets (fish) and tries to catch them which involves a large shift in position. The osprey chooses the best performing individual, or the overall best individual in the population, as its target as defined in (10).

$$FP_i = \{X_k | k \in \{1,2, \dots, N\} \cap F_k < F_i\} \cup \{X_{best}\}, i = 1,2, \dots, N \quad (10)$$

Location of ith osprey FP_i, population size N, and fitness values of kth and ith ospreys (F_k). Best osprey's position (X_{best}). In the search space, osprey finds a fish at random and attacks it. Position update when osprey approaches fish is simulated by (11).

$$X_{i,j}^{P1} = X_{i,j} + r_{i,j} \cdot \left(SF_{i,j} - I_{i,j} \cdot X_{i,j}\right), i = 1,2, \dots, N; j = 1,2, \dots, D \quad (11)$$

Where ith osprey's new location in stage 1 ($X_{i,j}^{P1}$), where $X_{i,j}^{P1}$ is its jth dimension. Osprey's initial location (X_i), and its jth dimension is denoted by ($X_{i,j}$). Fish selected by 1st osprey ($SF_{i,j}$), and ($r_{i,j}$) is a random number in $[0, 1]$. $I_{i,j}$ is chosen at random from the $\{1, 2\}$ collection. According to (12), the updated position is modified if it is outside the boundary; if it is below the lower bound, it is set to the lower bound; if it is above the upper bound, it is set to the upper bound.

$$X_{i,j}^{P1} = \begin{cases} X_{i,j}^{P1}, lb_j \le X_{i,j}^{P1} \le ub_j \\ lb_j, X_{i,j}^{P1} < lb_j \\ ub_j, X_{i,j}^{P1} > ub_j \end{cases} \quad (12)$$

Prior position is replaced if revised position's fitness value, as resolute by (11) and (12), is superior. The updated position of osprey is derived from process, which is in (13).

$$X_i^1 = \begin{cases} X_i^{P1}, F_i^{P1} < F_i \\ X_i, F_i^{P1} \ge F_i \end{cases} \quad (13)$$

Where fitness value (F_i^{P1}) of i osprey's new location following phase 1. Osprey's position following phase 1 is denoted by X_i^1.

ii) Phase 2: move fish to proper spot. This state represents the osprey's behavior as it carries the prey to a safe place which is equivalent to local search and refining already suggested candidate solutions. This has an essential part to play in optimism because it will ensure that each osprey will be exploring the local area of the candidate search space, looking for improvement. The positions of each candidate solution will be adjusted based on a small perturbation that depends on the current iteration t, which decreases its size with an increase in iterations, encouraging tuning, as defined in (14).

$$X_{i,j}^{P2} = X_{i,j}^{1} + \frac{lb_j + r_{i,j} \cdot (ub_j - lb_j)}{t}, t = 1,2,\dots,T \tag{14}$$

In phase 2, $i-th$ osprey's new location is $X_{i,j}^{P2}$, and its jth dimension is $X_{i,j}^{P2}$. T is the maximum no. of iterations, t is current iteration, and $r_{i,j}$ is a random number in $[0, 1]$. To ensure that solutions remain within valid bounds, as defined in (15).

$$X_{i,j}^{P2} = \begin{cases} X_{i,j}^{P2}, lb_j \leq X_{i,j}^{P1} \leq ub_j \\ lb_j, X_{i,j}^{P2} < lb_j \\ ub_j, X_{i,j}^{P2} > ub_j \end{cases} \tag{15}$$

The final position of osprey after local development is selected based on fitness evaluation as in (16).

$$X_i^2 = \begin{cases} X_i^{P2}, F_i^{P2} < F_i^{P1} \\ X_i^1, F_i^{P2} \geq F_i^{P1} \end{cases} \tag{16}$$

Where F_i^{P2} is the fitness value of the position X_i^{P2}, and X_i^2 is the position of the osprey after phase 2. To impose more strength on the ability of the optimizer to strike the exploration/exploitation balance so far, it implements a chain foraging mechanism. The chain foraging mechanism enables ospreys to potentially either follow the global best or mimic the individual that is directly before them in the chain. The A-OOA by adding a chain foraging mechanism as (17).

$$x_i^d(t+1) = \begin{cases} x_i^d(t) + r \cdot \left(x_{best}^d(t) - x_i^d(t)\right) + \alpha \cdot \left(x_{best}^d(t) - x_i^d(t)\right) i = 1 \\ x_i^d(t) + r \cdot \left(x_{i-1}^d(t) - x_i^d(t)\right) + \alpha \cdot \left(x_{best}^d(t) - x_i^d(t)\right) i = 2,\dots,N \end{cases} \tag{17}$$

Where $x_i^d(t)$ is the position of the ith osprey in dimension d, $x_{best}^d(t)$ is global best solution, $r \in [0,1]$ is the random value. The adaptive coefficient is calculated as in (18).

$$\alpha = 2 \cdot r \cdot \sqrt{|\log(r)|} \tag{18}$$

By adding the chain foraging mechanism to the osprey optimization algorithm for adaptive tuning, the model is able to adaptively tune the parameters classification parameters such as epoch, batch size, and learning rate. This approach improves convergence because the first osprey will always follow the global best solution. The rest of the ospreys will balance exploration with the global best solution while drifting from the previous osprey. This will improve classification accuracy while avoiding premature convergence and keeping the population diverse while optimizing. This would enhance the accuracy in classifying depression and robustness when analyzing an urban population.

3. RESULTS AND DISCUSSION

3.1. Descriptive statistics

The study sample's descriptive statistics, which include demographic factors and PHQ-9 questionnaire results. The mean age of participants is 2.00, with a standard deviation (SD) of 1.007, showing a distinct age group. Sex and married status have mean values of 1.50 and 1.52, respectively, with SDs of 0.500, indicating a balanced representation in both categories. The educational level has a mean of 3.73 and an SD of 1.289, indicating varying educational levels. Income is calculated using a mean of 2.05 and an SD of 0.801. The mean scores for the PHQ-9 items range from 2.23 (PHQ1) to 2.41 (PHQ9), with SDs of 0.759 to 0.823. These scores represent a moderate level of depression symptoms reported by participants, with responses varying across questions [6], [21], [22]. The constant mean scores across PHQ questions indicate that the sample's experience with depression symptoms is broadly homogeneous.

3.2. Multivariate tests

The multivariate tests that evaluate the influences of independent variables (IDV) on combined dependent variables (DV). An intercept has a highly significant effect across all multivariate tests

(Wilks' lambda, Hotelling's trace, Roy's largest root, and Pillai's trace) with a p-value of 0.000, suggesting that the model explains variance. However, variable "sex" had no significant effect on the DVs, with p-values of 0.926 across all tests, indicating that gender does not significantly contribute to the variance in the outcomes. Thus, while the model is generally effective, "sex" as a variable has little effect.

3.3. Estimated marginal means

The grand mean shows the average scores for each item on the PHQ. The mean scores range from 2.232 (PHQ1) to 2.411 (PHQ9), with standard errors of around 0.025, demonstrating similar measurement precision across items. The 95% confidence intervals are narrow, indicating that the population means may be reliably estimated. The very equal means across PHQ items suggest that patients reported similar levels of symptoms across all questionnaire items, indicating a consistent experience with depression symptoms. PHQ9 had the highest mean (2.411), which may reflect a somewhat higher prevalence of the symptoms indicated by that item. The narrow confidence intervals further confirm the reliability and consistency of the estimated PHQ-9 item scores. The mean for gender compares the average scores of males and females on the PHQ's nine components. The mean score differences between males and females are modest, indicating that both genders experience equal levels of depression symptoms. Most PHQ items have modest variances that fall within the standard error range, indicating that there is no significant gender variation. Males had significantly higher mean scores on PHQs 1, 4, 5, 6, and 8, whereas females had slightly higher means on PHQs 2, 3, 7, and 9. However, these differences are modest, and the overlapping 95% confidence ranges indicate that they are statistically insignificant.

3.4. Bivariate analysis

3.4.1. Crosstab

The crosstabulation of age distribution of the answers to the PHQ items in four age groups: 17–25, 26–40, 40–50, and 50+. PHQ items are categorized into three response categories: "Nearly every day," "Several days," and "More than half the days," which represent the frequency of symptoms, respectively. PHQ1 results suggest that most participants in all age groups report feeling symptoms "Nearly every day," with the 17–25 age group reporting the highest frequency (18.7%). Other PHQ items also show a similar pattern, with younger age groups (17–25 and 26–40) typically reporting a higher frequency of depression symptoms. In PHQ2, for instance, the largest percentage of respondents 21.1% are in the 17–25 age range and report feeling symptoms "Nearly every day". The frequency of reported symptoms seems to decline with age. With percentages ranging from 5.1% (PHQ1) to 7.3% (PHQ8), the 50+ age group, for example, consistently reports lower percentages of feeling symptoms "Nearly every day" across all PHQ items. The higher percentages in the "Nearly every day" category across most PHQ items indicate that younger age groups, especially those aged 17 to 25, are more likely to experience depressed symptoms more frequently, according to the research [5], [10], [35].

3.4.2. Chi-square tests

The chi-square tests for age across the nine PHQ items to determine whether there are any significant connections between age groups and the frequency of depressed symptoms. PHQ1 and PHQ2 had Pearson chi-square values of 8.522 and 8.322, respectively, with p-values of.202 and.215. These findings show no statistically significant relationship between age groups and responses to these items, as the p-values above the conventional significance level of 0.05. Pearson chi-square values for PHQ3 and PHQ4 are 11.923 and 11.366, with p-values of.064 and.078, respectively. These findings indicate a tendency toward significance, particularly for PHQ3, where the p-value is near 0.05. However, neither achieves conventional statistical significance. Pearson chi-square values for PHQs 5–9 range from 1.825 to 6.721, with p-values between 0.935 and 0.347. These findings repeatedly reveal no significant relationship between age and depression symptoms for these measures, with all p-values over 0.05. Overall, the findings indicate that age has no significant impact on the frequency of depressed symptoms as judged by the PHQ items in this population. The Pearson chi-square values for all nine PHQ items (PHQ1 through PHQ9) range from 0.194 to 2.184, with corresponding p-values of .335 to .908. These p-values are significantly above the standard significance level of 0.05, demonstrating that there is no statistically significant correlation between responses to any of the PHQ items and gender.

3.5. Pretrained model analysis

The proposed depression classification framework will be evaluated systematically against a group of pretrained models using a wide range of performance metrics including precision, recall, accuracy, F1-score, specificity, false positive rate (FPR), negative predictive value (NPV), false negative rate (FNR), area under the curve-receiver operating characteristic (AUC-ROC), and Matthew’s correlation coefficient (MCC). The models will be evaluated based on each model’s ability to help identify depressed symptoms, including experiencing sadness, diminished interest, sleep disturbances, and suicidal ideations in the urban population, as indicated in Table 1. The proposed model has substantially higher performance than compared architectures to the point that the proposed model had the highest performing scores across all metrics [11]–[15], [29]–[31].

Table 1 highlights the performance of the suggested model in comparison to the previously defined deep learning architectures like AlexNet, MobileNetv2, DenseNet121, and ResNet50 across multiple evaluation metrics. The suggested model (the new model) achieved the best overall prediction accuracy of 98.785%, an outstanding achievement. The MCC score of 0.994 provides strong evidence to support the agreement between predicted and actual outcomes and produced a very good AUC-ROC value of 0.992. In summary, the results provided good accuracy, generalizability, and reliability to classify levels of depression using the proposed model.

Table 1. Evaluation of performance with proposed and existing methods

Model	AlexNet	MobileNetv2	DenseNet121	ResNet50	Proposed
Accuracy	95.424	96.172	96.917	97.335	98.785
Precision	94.143	95.324	95.94	96.599	97.935
Recall	93.872	94.604	95.775	96.15	98.214
F1-score	93.977	94.933	95.866	96.346	98.056
Specificity	96.65	97.128	97.578	98.118	99.04
NPV	95.765	96.349	96.89	97.459	98.896
FPR	2.843	2.134	1.882	1.458	0.983
FNR	6.191	5.418	4.27	3.898	1.796
MCC	0.941	0.94	0.966	0.977	0.994
AUC-ROC	0.972	0.978	0.981	0.987	0.992

Figure 4 shows the confusion matrix for depression severity classification using the proposed multimodal deep transfer learning framework. It highlights the model's ability to accurately classify individuals across five severity levels: minimal, mild, moderate, moderately severe, and severe. The diagonal cells indicate high correct classification rates, ranging from 88.0% to 98.0%, while off-diagonal cells show minimal misclassifications, mostly between adjacent severity levels. This demonstrates the robustness and reliability of the model in distinguishing varying degrees of depression severity. Table 2 presents how tuning key hyperparameters epochs, learning rate, and batch size impacts the performance for the proposed model for depression detection. As the epochs are increased from 20 to 100 the performance for all metrics is improved. At 60 epochs (learning rate = 0.003, batch size = 64) this model passed above the 96% accuracy mark with a balanced precision and recall meaning the model had stabilized its learning. By 100 epochs the model using a smaller batch size of 16 and somewhat higher learning rate of 0.005, achieved the best performance: 98.785% accuracy, 98.214% recall, 97.935% precision, and 98.056 F1-score to accompany high specificity of 99.040%. This demonstrates that hyperparameter tuning did not just augment the effectiveness of classification, it also reduced false positives and false negatives indicating much larger true value for the maintaining reliable depression detection in real world mental health assessments [11]–[15], [29].

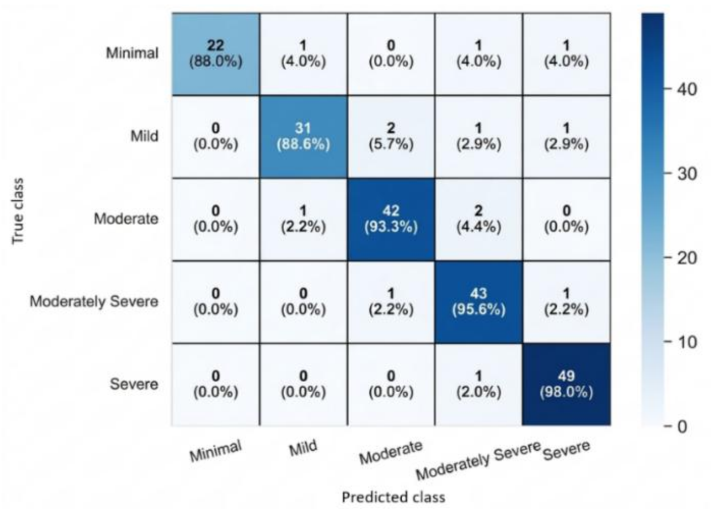


Figure 4. Confusion matrix-classifying depression severity

Table 2. Hyperparameter tuning results: classification performance metrics

Epoch	Learning rate	Batch size	Accuracy	Precision	Recall	F1-score	Specificity
20	0.001	256	94.230	92.876	92.145	92.508	95.098
40	0.002	128	95.617	94.008	93.761	93.884	96.084
60	0.003	64	96.728	95.476	95.051	95.262	97.012
80	0.004	32	97.885	96.743	96.432	96.587	98.574
100	0.005	16	98.785	97.935	98.214	98.056	99.040

The evidence in Table 3 demonstrates the effect of tuning hyperparameters—batch size, epochs, learning rate, on the suggested depression detection model's advanced performance metrics. Proposed model demonstrated progressive improvement across all evaluation metrics as the number of training epochs increased, as the training time increases from 20 to 100 epochs demonstrated with an overall improvement in both detecting depressed, and non-depressed cases. Figure 5 shows a comparative analysis of key performance metrics precision, F1-score, recall, and accuracy across different text-only BERT-based sentiment classifiers (DistilBERT, BioBERT, and DeBERTa) and proposed multimodal deep learning model. The proposed multimodal model outperforms all text-only classifiers across all metrics, demonstrating the added value of integrating multimodal features for more accurate depression severity classification.

Table 3. Hyperparameter tuning results: diagnostic performance metrics

Epoch	Learning rate	Batch size	NPV	FPR	FNR	MCC	AUC-ROC
20	0.001	256	94.345	4.902	7.855	0.926	0.961
40	0.002	128	95.554	3.916	6.239	0.936	0.971
60	0.003	64	96.812	2.988	4.949	0.961	0.982
80	0.004	32	97.865	1.426	3.568	0.982	0.989
100	0.005	16	98.896	0.983	1.796	0.994	0.992

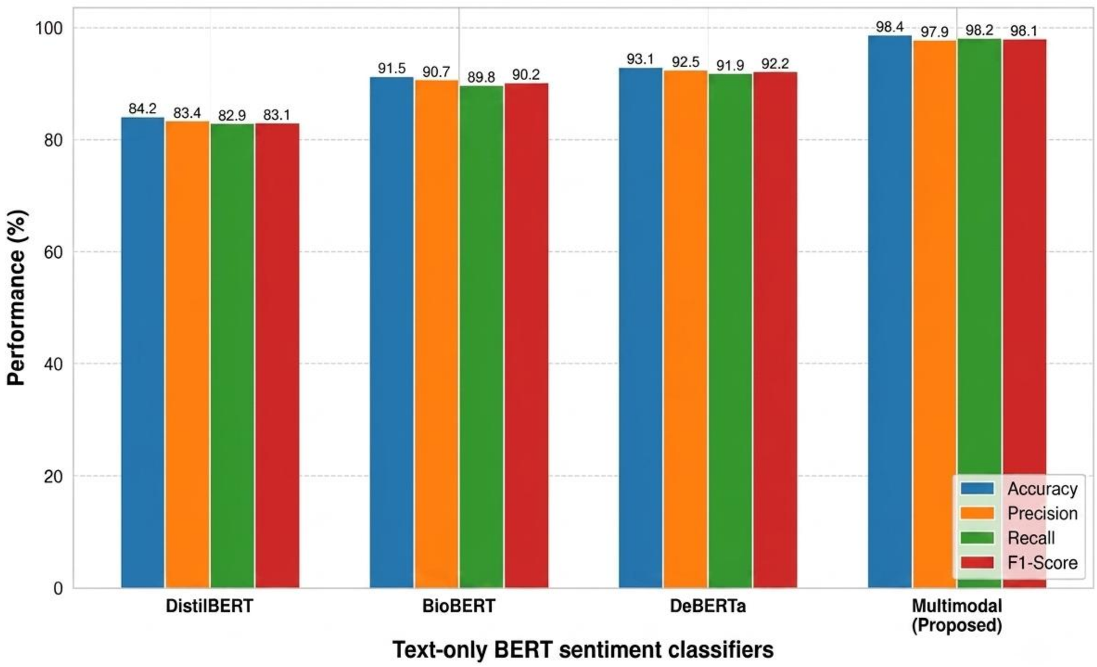
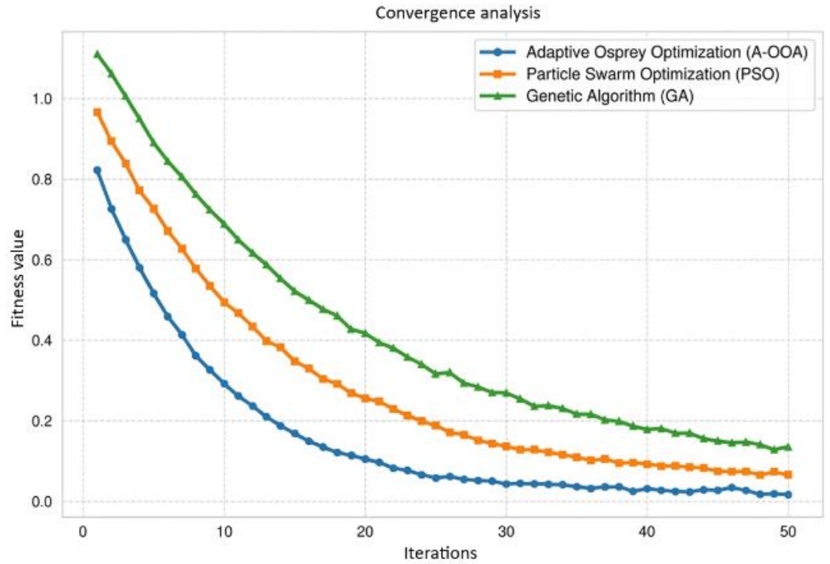
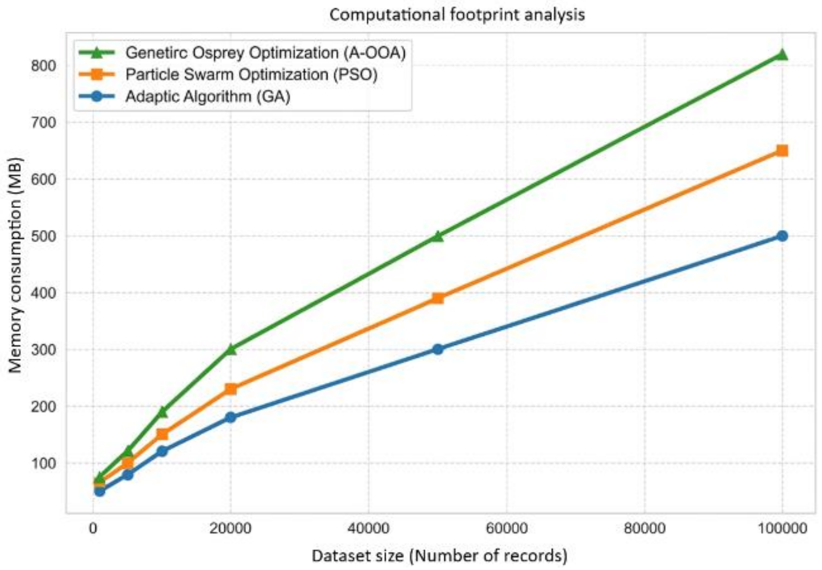


Figure 5. Comparison of text-only BERT models and multimodal classifier performance

Figure 6 illustrates the comparative analysis of the proposed optimization algorithms. Where Figure 6(a) depicts the convergence analysis, and Figure 6(b) depicts memory consumption trends across increasing dataset sizes for A-OOA, particle swarm optimization (PSO), and genetic algorithm (GA). As shown, A-OOA consistently maintains the lowest memory usage, demonstrating superior efficiency and scalability compared to PSO and GA.



(a)



(b)

Figure 6. Comparative analysis of the proposed optimization algorithms (a) convergence analysis and (b) computational footprint analysis

The linear growth pattern observed in Figures 6(a) and 6(b) confirms that A-OOA achieves optimal resource utilization, making it suitable for large-scale urban mental-health datasets and real-time public-health deployment. Figure 7 shows the runtime performance (in seconds) of three optimization algorithms A-OOA, PSO, and GA as the dataset size increases. The plot demonstrates that A-OOA consistently requires less runtime compared to PSO and GA across all dataset sizes, indicating superior scalability and efficiency for hyperparameter optimization in the proposed depression classification model. The present research found a significant incidence of depression symptoms in the Metropolitan city community, indicated by moderate mean PHQ-9 scores. The very consistent distribution of depressive symptoms across all measures suggests a broad issue of emotional anguish. Although the overall model was found to be beneficial in the multivariate analysis, depression levels were not significantly impacted by gender, demonstrating that symptom experiences are similar across the sexes. Examining depressed symptoms by age, it is clear that younger participants, ages 17-25, have higher frequencies of symptoms on PHQ-9 than older groups. The symptoms are most commonly noticed by younger participants, particularly in categories such as "Nearly every day", and less frequently by older responders aged 50 and up. A hybrid deep transfer learning framework was proposed and incorporated into this study to autonomously classify the levels of depression in the datasets, which combined both ResBCNN (in one branch) with A-IRes (in another branch), and the feature maps from each of the branches were concatenated and entered into fully connected layers and a softmax classifier. The output measurements even when demonstrated in terms of the statistical scores of accuracy, recall, and MCC have shown expected performance levels worth noting for factional and technical framework demonstrating corresponding levels of effectiveness for these respective studies.

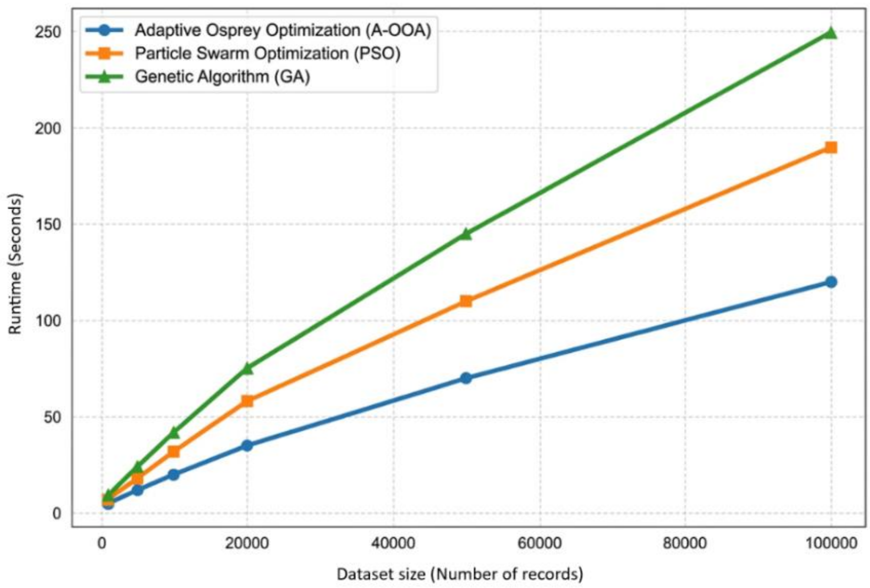


Figure 7. Scalability analysis of optimization algorithms

Figure 8 shows the attention scores assigned to each of the nine PHQ-9 questionnaire items by the proposed deep transfer learning model. These attention weights indicate the relative importance of each symptom in contributing to the classification of depression severity. PHQ3 (sleep disturbances), PHQ7 (psychomotor changes), PHQ8 (feelings of worthlessness), and PHQ2 (depressed mood) receive the highest attention, highlighting their significant influence in the model's decision.

This visualization enhances the interpretability of the model by identifying key depressive symptoms driving classification. Figure 9 shows the Shapley additive explanations (SHAP) feature importance scores for each PHQ-9 questionnaire item across different depression severity levels (minimal, mild, moderate, moderately severe, and severe). The mean absolute SHAP values are shown by the horizontal bars, indicating the contribution of each PHQ-9 item to the model's classification decisions. This visualization helps identify which symptoms have the greatest influence on predicting depression severity, enhancing model interpretability and clinical trust.

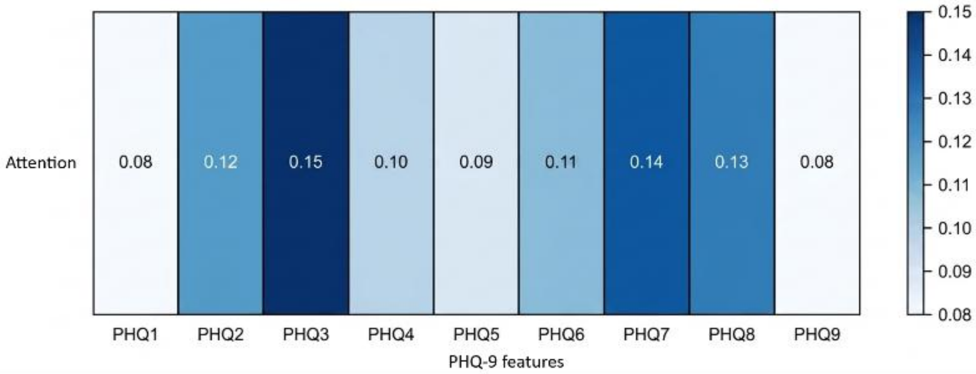


Figure 8. Attention weights for PHQ-9 features in depression classification

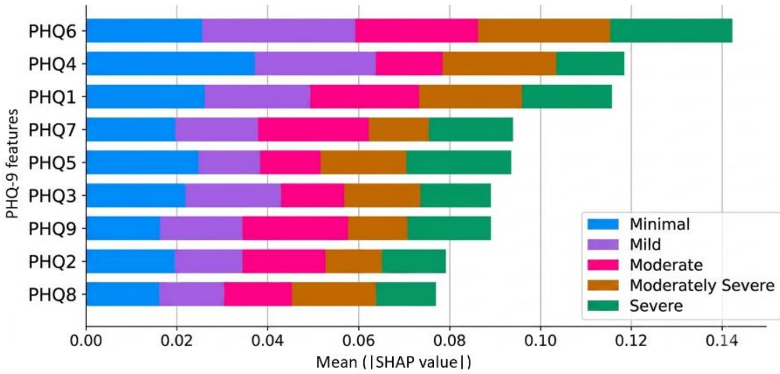


Figure 9. SHAP feature importance for PHQ-9 classification

4. CONCLUSION

The PHQ-9 was used to investigate the incidence of depression in metropolitan cities to better understand its scope and demographic correlates of depressed symptoms in the population. The investigation revealed a significant prevalence of depression, with a modest mean score for persons experiencing emotional discomfort. Second, demographic characteristics such as gender, age, marital status, income and educational level have been investigated for their correlations with the severity of depression. Though young persons exhibited greater rates of depressive symptoms, age, gender, and demographic variables had no effect on the estimated overall frequency. This homogeneity shows that depressive symptoms are rather common and affect all demographic groups. The current study shows the importance of both broad-based public health programs and targeted therapies in response to frequent depressed symptoms across all demographic profiles. These data can be utilized to guide public health policies and mental health initiatives in metropolitan cities, hence enhancing mental health outcomes and services. Using pretrained models like Inception V3 and ResNet50 coupled with different data mining methods create better classifications of subtypes of depression and contribute to a stronger assessment of mental health, particularly in an urban context. The model developed in the proposed framework performed best with an accuracy (98.78%), recall (98.21%), precision (97.93%), F1-score (98.05%), specificity (99.04%), and AUC-ROC (0.992), while achieving the highest performance of existing pretrained models like Inception V3, MobileNet, AlexNet, VGG16, and ResNet50 across all metrics that were performed in this study. These findings have significant consequences for Metropolitan city public and mental health policy. First, the consistent and high incidence of depressive symptoms across demographic characteristics highlights the need for intervention in the vast range of treatments available to patients with depression. The findings also demonstrate a strong need for mental health education and awareness initiatives that reduce the stigma associated with mental diseases and enhance help-seeking behavior.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
U. Ananthanagu	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Pooja Agarwal		✓		✓	✓	✓		✓	✓	✓	✓	✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

All participants provided informed consent, and the study followed ethical guidelines to protect their confidentiality and rights.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, [UA].

REFERENCES

[1] R. K. Saini, S. Chaudhury, N. Singh, D. S. Chadha, and R. Kapoor, "Depression, anxiety, and quality of life after percuataneous coronary interventions," *Industrial Psychiatry Journal*, vol. 31, no. 1, pp. 6–18, Jan. 2022, doi: 10.4103/ipj.ipj_126_21.

[2] A. Kupferberg and G. Hasler, "The social cost of depression: Investigating the impact of impaired social emotion regulation, social cognition, and interpersonal behavior on social functioning," *Journal of Affective Disorders Reports*, vol. 14, Dec. 2023, doi: 10.1016/j.jadr.2023.100631.

[3] J. Purtle, K. L. Nelson, N. Z. Counts, and M. Yudell, "Population-based approaches to mental health: history, strategies, and evidence," *Annual Review of Public Health*, vol. 41, no. 1, pp. 201–221, Apr. 2020, doi: 10.1146/annurev-publhealth-040119-094247.

[4] J. Nübel, A. Guhn, S. Müllender, H. D. Le, C. Cohrdes, and S. Köhler, "Persistent depressive disorder across the adult lifespan: results from clinical and population-based surveys in Germany," *BMC Psychiatry*, vol. 20, no. 1, Dec. 2020, doi: 10.1186/s12888-020-2460-5.

[5] S. Krokstad *et al.*, "Divergent decennial trends in mental health according to age reveal poorer mental health for young people: repeated cross-sectional population-based surveys from the HUNT Study, Norway," *BMJ Open*, vol. 12, no. 5, May 2022, doi: 10.1136/bmjopen-2021-057654.

[6] L. G. Vu *et al.*, "Factor structures of patient health questionnaire-9 instruments in exploring depressive symptoms of suburban population," *Frontiers in Psychiatry*, vol. 13, Aug. 2022, doi: 10.3389/fpsyt.2022.838747.

[7] S. Cook *et al.*, "Socio-demographic, behavioural and psycho-social factors associated with depression in two Russian cities," *Journal of Affective Disorders*, vol. 290, pp. 202–210, Jul. 2021, doi: 10.1016/j.jad.2021.04.093.

[8] N. K. Gebregziabher, T. B. Netsereab, Y. G. Fessaha, F. A. Alaza, N. K. Ghebrehiwet, and A. H. Sium, "Prevalence and associated factors of postpartum depression among postpartum mothers in central region, Eritrea: a health facility based survey," *BMC Public Health*, vol. 20, no. 1, Dec. 2020, doi: 10.1186/s12889-020-09676-4.

[9] S. Verma and A. Mishra, "Depression, anxiety, and stress and socio-demographic correlates among general Indian public during COVID-19," *International Journal of Social Psychiatry*, vol. 66, no. 8, pp. 756–762, Dec. 2020, doi: 10.1177/0020764020934508.

[10] S. Sahay, C. F.-A. Vargas, S. C. Hoar, J. Choi, and S. M. O'Donovan, "The effect of the covid-19 pandemic on mental health among college students in the United States," *Translation: The University of Toledo Journal of Medical Sciences*, vol. 12, no. 5, 2024, doi: 10.46570/utjms.vol12-2024-1280.

[11] L. Luo *et al.*, "Predictors of depression among Chinese college students: a machine learning approach," *BMC Public Health*, vol. 25, no. 1, Feb. 2025, doi: 10.1186/s12889-025-21632-8.

[12] Y. L. Q. Song, L. Chen, H. Liu, and Y. Liu, "Machine learning algorithms to predict depression in older adults in China: a cross-sectional study," *Frontiers in Public Health*, vol. 12, Jan. 2025, doi: 10.3389/fpubh.2024.1462387.

[13] J. W. Park, C. W. Ko, D. Y. Lee, and J. C. Kim, "Prediction of late-onset depression in the elderly Korean population using machine learning algorithms," *Scientific Reports*, vol. 15, no. 1, Jan. 2025, doi: 10.1038/s41598-025-85157-1.

[14] W. Qi *et al.*, "Prediction of postpartum depression in women: development and validation of multiple machine learning models," *Journal of Translational Medicine*, vol. 23, no. 1, Mar. 2025, doi: 10.1186/s12967-025-06289-6.

[15] C. Hu *et al.*, "Development and application of a machine learning-based antenatal depression prediction model," *Journal of Affective Disorders*, vol. 375, pp. 137–147, Apr. 2025, doi: 10.1016/j.jad.2025.01.099.

[16] S. Husain, S. Rao, S. Suresh, K. V. A. Jesudoss, B. Krishna, and J. P. Raj, "Screening for those at risk for anxiety and depression warranting further clinical evaluation among patients presenting to breast services: a single-centre, cross-sectional study," *BMJ Open*, vol. 14, no. 4, Apr. 2024, doi: 10.1136/bmjopen-2023-080216.

[17] K. T. Meitei and H. S. Singh, "Prevalence of depression among Indian population," *Online Journal of Health and Allied Sciences*, vol. 18, no. 4, pp. 1–10, 2019.

[18] J.-P. Ruben, M. L. Ekstrand, E. Heylen, and K. Srinivasan, "Correlates for the severity of suicidal risk in participants with common mental disorders with comorbid chronic medical conditions in rural primary healthcare settings in India," *Indian Journal of Psychiatry*, vol. 65, no. 6, pp. 687–693, Jun. 2023, doi: 10.4103/indianjpsychiatry.indianjpsychiatry_41_23.

[19] R. R. Shenoy, D. A. Prabhu, A. K. Pandey, S. K. Praharaj, and R. S. Shetty, "Common mental disorders and their correlates: a community-based survey among women in Southern Karnataka, India," *Indian Journal of Psychological Medicine*, vol. 46, no. 4, pp. 330–337, Jul. 2024, doi: 10.1177/02537176231220543.

[20] K. Srinivasan, E. Heylen, R. J. Pradeep, P. K. Mony, and M. L. Ekstrand, "Collaborative care compared to enhanced standard treatment of depression with co-morbid medical conditions among patients from rural South India: a cluster randomized controlled trial (HOPE Study)," *BMC Psychiatry*, vol. 22, no. 1, Jun. 2022, doi: 10.1186/s12888-022-04000-3.

[21] E.-H. Lee, E. H. Kang, H.-J. Kang, and H. Y. Lee, "Measurement invariance of the patient health questionnaire-9 depression scale in a nationally representative population-based sample," *Frontiers in Psychology*, vol. 14, Aug. 2023, doi: 10.3389/fpsyg.2023.1217038.
[22] E. F. -Pedrero, A. D. -Gómez, A. P. -Albéniz, S. Al-Halabí, B. L. -Molina, and M. Debbané, "Youth screening depression: validation of the patient health questionnaire-9 (PHQ-9) in a representative sample of adolescents," *Psychiatry Research*, vol. 328, Oct. 2023, doi: 10.1016/j.psychres.2023.115486.
[23] L. Mihić, G. Knežević, L. B. Lazarević, and N. P. Marić, "Screening for depression in the Serbian general population sample: an alternative to the traditional patient health questionnaire-9 cut-off score," *Journal of Public Health*, vol. 46, no. 1, pp. e15–e22, Feb. 2024, doi: 10.1093/pubmed/fdad204.
[24] C.-Y. Chen, J.-H. Chen, S.-C. Ree, C.-W. Chang, and S.-H. Yu, "Associations between estradiol and testosterone and depressive symptom scores of the patient health questionnaire-9 in ovariectomized women: a population-based analysis of NHANES data," *Annals of General Psychiatry*, vol. 19, no. 1, Dec. 2020, doi: 10.1186/s12991-020-00315-1.
[25] T. L. H. Ngoc *et al.*, "Patient health questionnaire (PHQ-9): a depression screening tool for people with epilepsy in Vietnam," *Epilepsy & Behavior*, vol. 125, Dec. 2021, doi: 10.1016/j.yebeh.2021.108446.
[26] U. Linnenkamp *et al.*, "PHQ-9, CES-D, health insurance data—who is identified with depression? A population-based study in persons with diabetes," *Diabetology & Metabolic Syndrome*, vol. 15, no. 1, Mar. 2023, doi: 10.1186/s13098-023-01028-7.
[27] V. F. J. Cumbe *et al.*, "Validity and item response theory properties of the patient health questionnaire-9 for primary care depression screening in Mozambique (PHQ-9-MZ)," *BMC Psychiatry*, vol. 20, no. 1, Dec. 2020, doi: 10.1186/s12888-020-02772-0.
[28] A. P. Miller *et al.*, "Construct validity and internal consistency of the patient health questionnaire-9 (PHQ-9) depression screening measure translated into two Ugandan languages," *Psychiatry Research Communications*, vol. 1, no. 2, Dec. 2021, doi: 10.1016/j.psycom.2021.100002.
[29] C. Uyulan *et al.*, "Major depressive disorder classification based on different convolutional neural network models: deep learning approach," *Clinical EEG and Neuroscience*, vol. 52, no. 1, pp. 38–51, Jan. 2021, doi: 10.1177/1550059420916634.
[30] Sofia, A. Malik, M. Shabaz, and E. Asenso, "Machine learning based model for detecting depression during the COVID-19 crisis," *Scientific African*, vol. 20, Jul. 2023, doi: 10.1016/j.sciaf.2023.e01716.
[31] R. Siddiqua, N. Islam, J. F. Bolaka, R. Khan, and S. Momen, "AIDA: artificial intelligence based depression assessment applied to Bangladeshi students," *Array*, vol. 18, Jul. 2023, doi: 10.1016/j.array.2023.100291.
[32] C. Li, Y. Xia, and Y. Zhang, "Relationship between subjective well-being and depressive disorders: novel findings of cohort variations and demographic heterogeneities," *Frontiers in Psychology*, vol. 13, Jan. 2023, doi: 10.3389/fpsyg.2022.1022643.
[33] C. Kim, "Exploring factors influencing depression: socioeconomic perspectives using machine learning analytics," *Electronics*, vol. 14, no. 3, Jan. 2025, doi: 10.3390/electronics14030487.
[34] B. A. Kohrt, N. P. Luitel, P. Acharya, and M. J. D. Jordans, "Detection of depression in low resource settings: validation of the patient health questionnaire (PHQ-9) and cultural concepts of distress in Nepal," *BMC Psychiatry*, vol. 16, no. 1, Dec. 2016, doi: 10.1186/s12888-016-0768-y.
[35] A. Mallya, R. Koppad, and P. Kumar, "Depression, anxiety, and stress among urban and rural adolescents in Shivamogga, Karnataka," *F1000Research*, vol. 12, Jul. 2024, doi: 10.12688/f1000research.139603.3.

BIOGRAPHIES OF AUTHORS

U. Ananthanagu currently pursuing a Ph.D. at PES University, India. He holds a master's degree in Computer Science and Engineering from Jawaharlal Nehru Technological University, a bachelor's degree in Computer Science and Engineering from Visvesvaraya Technological University, and a diploma in Computer Science and Engineering from the Technical Board of Education. With 15 years of experience in academics, his primary areas of research interest include data science, machine learning, NLP, social media analysis, and transfer learning. He has published over 30 papers in international journals and conferences. He is a life member of the Computer Society of India (CSI) and the Indian Society for Technical Education (ISTE). He can be contacted at email: pes1pd20cs006@pesu.pes.edu or nagu.anantha@gmail.com.

Pooja Agarwal completed her bachelor's degree in 1998 from CCS University, Meerut, and her master's in Computer Science in 2001 from Banasthali Vidyapith, Rajasthan, and was awarded the Ph.D. degree in Computer Science primarily focusing in the area of machine learning and deep learning from Visvesvaraya Technological University in 2018. Since July 2008, she has been associated with PES institutions at PESIT Bangalore South Campus (now PES University-EC Campus). She has worked with over a hundred students on various projects at the undergraduate and master levels in the application areas of data science, machine learning, soft computing, natural language processing, data analytics, and deep architectures. She has around 40 publications in various reputed international conferences and journals. She is a member of IEEE and member IAENG. She can be contacted at email: poojaagarwal@pes.edu.