

Interpretable Deep Learning for Psychosocial Assessment of Women's Mental Health and Empowerment

Hanan Muhajab

hmuha.jab@jazanu.edu.sa

Jazan University

Saahira banu Ahamed

Jazan University

Article

Keywords: Women Mental Health, Empowerment Assessment, Psychosocial Factors, Explainability Artificial Intelligence, Deep Learning

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10963194/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Interpretable Deep Learning for Psychosocial Assessment of Women's Mental Health and Empowerment

Author-1 :

¹Hanan Muhajab,

Department of Computer Science, College of Engineering and Computer Sciences,
Jazan University, Jazan, Saudi Araiba

²Engineering and Technology Research Center, Jazan University, P.O.Box 114,Jazan 82817,
Saudi Arabia

hmuhajab@jazanu.edu.sa

<https://orcid.org/0009-0000-3472-3723>

Author-2 :

¹Saahira banu Ahamed

Department of Computer Science, College of Engineering and Computer Sciences,
Jazan University, Jazan, Saudi Araiba

²Engineering and Technology Research Center, Jazan University, P.O.Box 114,
Jazan 82817, Saudi Arabia

sahamed@jazanu.edu.sa

<https://orcid.org/0009-0000-3472-3723>

Abstract

Women's mental health is an important public health concern that significantly impacts their overall well-being and quality of life. Among several mental health challenges, depression is one of the most common conditions affecting women, while maternal depression presents additional psychological challenges during pregnancy and the postpartum period. These conditions are influenced by multiple social, economic, psychological, and empowerment-related factors, making their accurate assessment challenging. Machine learning (ML) and deep learning (DL) approaches are offers effective means of identifying hidden patterns and predicting women's mental health outcomes. Moreover, existing methodologies have limited capability to explain how empowerment and psychosocial factors influence women's mental health, highlighting the need for an explainable framework. To overcome these challenges, this research work develops an Explainable Women's Mental Health and Empowerment Assessment (X-WMHEA) model using psychosocial factors for effective analysis. The

proposed model utilizes HSIC-LASSO to identify informative and non-redundant psychosocial factors by recognizing their dependencies with the target outcomes. Moreover, the TabM model learns complex relationships among the selected psychosocial factors and classify women's mental health conditions. In addition, the Adan optimizer is utilized to refine the model parameters and enhance the convergence of the learning process. To enhance model transparency, the SHAP is integrated to quantify the contribution of individual psychosocial factors and provide meaningful insights into factors associated with women's mental health and empowerment. The performance of the proposed X-WMHEA model demonstrates superior performance, achieving accuracy of 95.53% and 96.75% on the Women Depression Dataset and Maternal Depression Dataset. Additionally, an ablation study is conducted to examine the contribution of each component in the proposed framework. Thus, the proposed model provides an effective and interpretable approach for analyzing psychosocial factors associated with women's mental health and empowerment, supporting reliable assessment and informed decision-making.

Keywords: Women Mental Health; Empowerment Assessment; Psychosocial Factors; Explainability Artificial Intelligence; Deep Learning

1. Introduction

Depression is a serious and common mental illness that affects an individual's mood, including anxiety, lack of energy, trouble focusing, losing interest in daily activities or hobbies, and impaired memory [1]. It is a major mental condition that can severely impact individuals of any age or gender. Mostly, women experience mental health conditions with the highest frequency of 21.5% compared to 15.7% in men. Various factors impact the mental health of women, including hormonal, socioeconomic, trauma, biological, and other factors that affect their mental health [2]. In a developing country, discrimination based on gender, lack of educational opportunities and employment possibilities, and external demands regarding marriage and family roles lead to mental illness in women [3]. Mainly, Postpartum Depression Disorder (PPDD) is a common mental health disorder that occurs in women after childbirth, which particularly affects the mother's mental health, reduces the positive experience of motherhood, and negatively impacts a woman's well-being [4]. To address these risk factors a significant mental health treatment is needed to offer essential support. The risk factors for PPDD include physical, social, mental, and way of living. Existing postpartum mental health

analysis depends on standardized screening questionnaires and, most particularly, the Edinburgh Postnatal Depression Scale (EPDS) administered during postpartum visits [5]. These traditional methods are time-consuming, have inter-rater variability, are sensitive to self-report bias, and have poor availability for women in underprivileged communities. Routine screening mostly concentrates on depression, while post-delivery anxiety often ignored [6].

In recent times, Artificial intelligence (AI) and machine learning (ML) have been significantly emerging and providing various solutions for maternal mental health screening [7]. As ML-based methods can advance simple statistical tools to intricate, distributed systems, they must adhere to a multi-layered principle authorizing privacy, transparency, and non-maleficence, mainly when employed for subtle clinical diagnoses including PPD, since increasing model difficulty magnifies the risk of responsibility failures [8]. ML-based approaches can recognize non-linear patterns among demographic, emotional, complex, and clinical variables with reliability that matches healthcare professionals. Deep learning (DL), a subdivision of ML, has the ability to learn hierarchical data representations and has attained advanced performance in healthcare fields [9]. Multidimensional data integration has been demonstrated to effectively improve predictive accuracy compared to traditional models by influencing interactive terms that offer an overall view of patient condition. However, the major challenge for clinical adoption is the “black-box” nature of complex ML-based methods, because predictions will be generated without a proper explanation [10].

Healthcare professionals need an explainable model to understand the mathematical outputs and to recognize the errors. Explainable AI (XAI) overcomes these problems by offering transparent, reliable predictions and quantifying the involvement of individual features in those predictions, therefore enhancing the robustness and scalability of ML-based decision-making [11]. Despite several successful applications of deep learning (DL) in analyzing mental health in women, the performance often suffers from variability in psychosocial factors and empowerment-related features. While DL models are able to extract latent patterns from data, there remains a difficulty in clearly showing how an individual factor contributes to the outcome. Hence, the problem that arises in analyzing mental health accurately and explaining what contributing factors affect it motivates to develop a women’s mental health and empowerment assessment framework that is accurate and interpretable. To address these conditions, this study proposed an Explainable Women's Mental Health and Empowerment Assessment (X-WMHEA) model using psychosocial factors for effective analysis. The key contributions of this research are as follows:

- Employ a HSIC-LASSO feature selection strategy that combines the ability of HSIC (Hilbert-Schmidt Independence Criterion) to measure feature relevance with the sparsity of LASSO to select the most informative and non-redundant features.
- Design a TabM (Tabular Model) to model complex relationships and interactions among tabular features, learn meaningful feature representations, and improve the classification of women's mental health and empowerment assessment.
- Integrate an Adan (Adaptive Nesterov Momentum) optimizer for efficient model training by adapting the learning rate based on the gradients of the model parameters.
- Apply a SHAP-based explainability artificial intelligence (XAI) to assess the importance of individual features and explain their influence on the predicted mean health outcomes.
- Assess the performance of the proposed X-WMHEA methodology using the Women Depression Dataset and the Maternal Depression Dataset, and demonstrate its effectiveness compared with existing models under various evaluation measures.

2. Related Works

This section provides a concise overview of existing related works on women's mental health analysis and empowerment assessment. Akter et al. [12] explored the connection between women's empowerment and the incidence of selected non-communicable diseases among married, divorced, separated, and reproductive-aged women. Women's empowerment was evaluated by utilizing the survey-based women's empowerment index (SWPER), including three different domains: social freedom, attitudes toward violence, and making decisions. Kumari [13] explored the relation between economic participation, individual decision-making, education, and socio-economic independence in empowering women. Employing quantitative research methods, it examines the influence of key factors including employment status, level of education, and public consciousness on empowerment. The Statistical tools like mean, standard deviation, and correlation analysis are also employed to examine the key factors. The research determines that empowerment needs incorporated efforts in decision-making, socio-economic reforms, and education. Deroncele-Acosta et al. [14] explored the relation between psychological capital, educational motivation, and university academic performance (UAP) and classified the factors of women's mental health success. A cross-

sectional, mixed-methods design was used with a sample of 108 female doctoral candidates. At the qualitative level, meetings and interviews were carried out to collect information related to mental health conditions. Al-Mamary et al. [15] examined the influence of a number of challenges on women's empowerment to commence a new business. The study concentrated on different factors, including market circumstances, psychological, socio-economic culture, innovative mindset, and education and skill factors, to understand the impact of these intermediaries on women's empowerment in entrepreneurship.

Vicent et al. [16] explored the relationship between social relations and women's empowerment on the 'constant utilization of clean cooking fuels and technologies' (CCF&Ts) across Ugandan households. The research suggested policies that enable women and assist in sharing energy-based information in social media or groups to improve household proficiency in adopting the utilization of CCF&Ts. Kumari et al. [17] examined the significant role of women empowerment and gender-based digital financial inclusion (GDFI) in reducing poverty. Therefore, the GDFI index includes different indicators and was constructed using principal component analysis (PCA). It also utilizes econometric methods like the cross-sectional dependence (CSD) test, second-generation unit root tests such as CADF and CIPS, and existing cointegration tests. Betrand et al. [18] developed a multimodal, gender-specific DL-based method introduced to predict depression and mental health conditions in women. The research uses a late-fusion architecture that incorporates various data streams such as: visual facial cues, textual transcripts, tabular clinical data (PHQ-8 scores), and acoustic features. The method uses particular neural network branches for each model- a Bidirectional LSTM (BiLSTM) for audio, a Temporal CNN for visual sequences, a Transformer (DistilBERT) for text, and a Multi-Layer Perceptron (MLP) for tabular data before concatenating their outputs for a final prediction. Qi et al. [19] developed an ML-based model by utilizing the biopsychosocial predictors for predicting risk raised from PPD for post-delivery and to offer various risk appraisal tools for timely detection of PPD. Six ML-based approaches and seven feature selection models were utilized to advance the methods and their execution of prediction performance. Furthermore, postnatal predictors significantly enhanced the predictive performance of the postnatal models.

Several studies have targeted women's health in relation to individual psychological or social elements; nevertheless, there has been little attention paid to the integration of psychosocial factors and empowerment-related attributes. Moreover, a variety of ML and DL algorithms have been applied for mental health prediction and classification; still, there has been inadequate attention to extracting relationships among various factors, and few studies have

clearly identified which attributes are important when making predictions. Thus, there arises the need to consider a robust and explainable mechanism that can contribute to performing women's mental health and empowerment assessment. Thus, to deal with those shortcomings, this research paper proposes X-WMHEA, an effective model which analyzes and predicts mental health and empowerment among women, with greater emphasis on using psychosocial attributes for the analysis.

3. Materials and Models

The proposed X-WMHEA technique follows a systematic workflow for effective analysis using psychosocial factors and obtaining interpretable outcomes. The complete workflow of the proposed X-WMHEA methodology is depicted in Fig. 1. Initially, the preprocessing is performed to improve data quality using missing value handling with multiple imputation by chained equations (MICE), duplicate removal, target encoding, IQR-based outlier detection, and Robustscaler. Next, the HSIC-LASSO utilized to identify the most relevant and informative attributes. TabM subsequently learns the complex relationships and interactions among the selected features to improve the classification performance. In addition, the Adan optimizer is used to improve the training performance and achieve stable convergence, while SHAP is applied to explain the model predictions and identify the contributions of individual features.

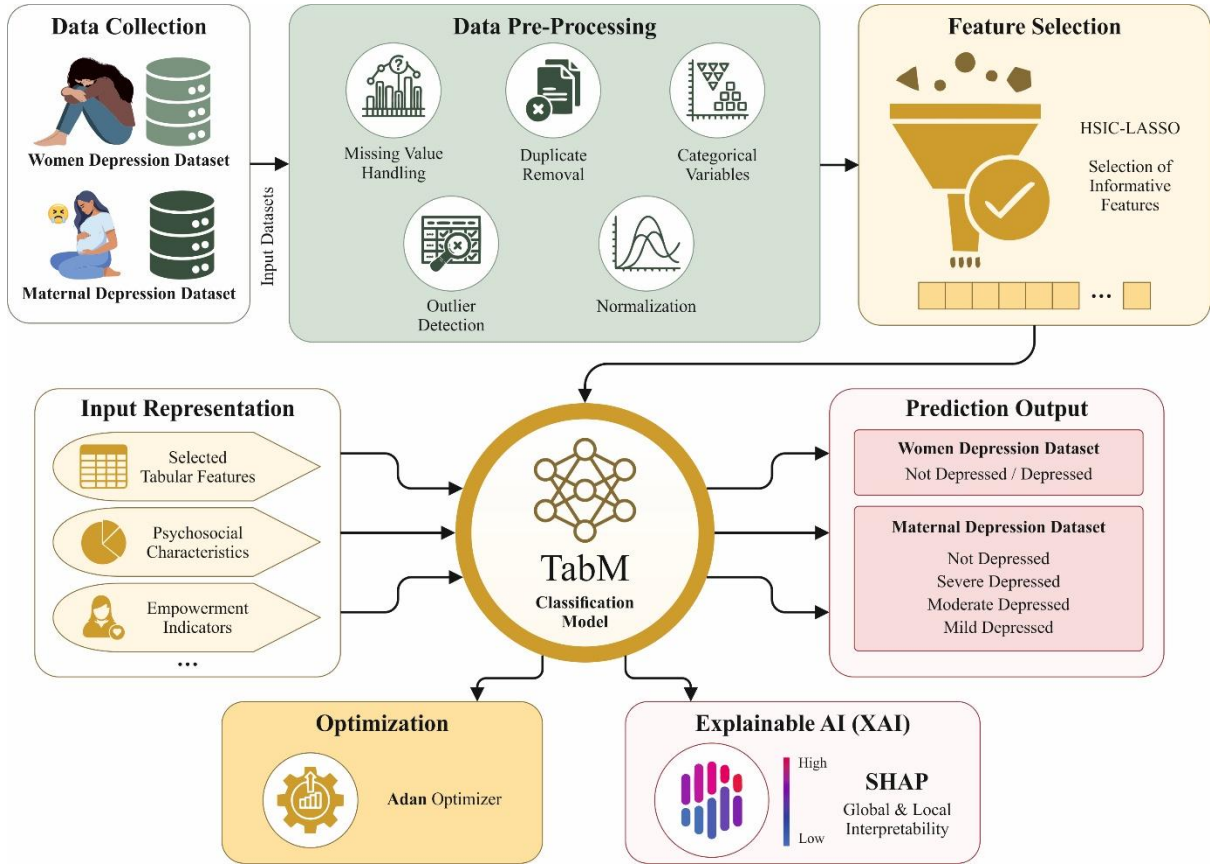


Fig. 1. Complete workflow of the proposed X-WMHEA model

3.1. Dataset Used

The proposed X-WMHEA is examined using Women Depression Dataset (Dataset 1) [20] and Maternal Depression Dataset (Dataset 2) [21].

Dataset 1: This dataset comprises factors that impact the mental state of women in India. The data has been gathered using a survey among different women impacted psychological issues over the country. These factors diverge with states, culture and financial status. Depression prediction can be outperformed by offering the factors as input to the system. The factors involved in the dataset are considered as the attributes for deploying the presented DL methodologies.

Dataset 2: This dataset consists of socio-demographic data and depression-based responses gathered by the Zung Self-Rating Depression Scale (SDS). It concentrates on mothers of school-going children from chosen public and private schools. Participants were selected at random from 12 schools and incorporated mothers of children studying in classes V to VIII. The participants had to have their children registered in the similar school for almost six months. Additional information from the Daffodil Institute of Social Sciences (DISS) was

involved to deliver wider representation in the dataset. Table 1 summarizes the class distribution of the two-depression dataset. In women depression dataset consist of two categorizes namely not depressed (2994) and depressed (3156) with total of 6150 instances. The maternal depression dataset includes four types of depressed states such as not depressed (200), severe depressed (140), moderate depressed (133), and mild depressed (121) with a total of 594 samples are depicted.

Table 1 Details of Dataset 1 and 2

Class Label	Class Count
Women Depression Dataset	
Not Depressed	2994
Depressed	3156
Total	6150
Maternal Depression Dataset	
Not Depressed	200
Severe Depressed	140
Moderate Depressed	133
Mild Depressed	121
Total	594

3.2. Data Preprocessing

The data preprocessing involves handling missing values using MICE, removing duplicate records, applying target encoding to categorical variables, detecting outliers using the IQR method, and scaling the features using RobustScaler [22].

Missing Value Handling using MICE: To evaluate missing values MICE is applied by considering the correlations among diverse variables in the dataset. By using the observed values from other relevant features, MICE iteratively estimates the missing entries.

Duplicate Removal: To detect and remove repeated records, duplicate removal is performed. This avoids similar observations being calculated many times and helps to retain the reliability of the data.

Target Encoding: It transforms categorical variables into numerical values according to their association with the target variable. Every class is exchanged with a statistical representation of the corresponding target values, which makes the non-numeric data suitable for further consideration.

Outlier Detection using IQR: IQR is employed to classify observations that unusually fall away from the central range of the data. The variation between the third quartile (Q3) and first quartile (Q1) is evaluated by the IQR, and values outside the determined range are considered as potential outliers.

$$IQR = Q_3 - Q_1 \quad (1)$$

Normalization using RobustScaler: RobustScaler converts numerical features utilizing the median and interquartile range. This creates a scaling procedure less sensitive to extreme values and delivers a more consistent feature representation for skewed distributions.

$$X_{scaled} = \frac{X - \tilde{X}}{Q_3 - Q_1} \quad (2)$$

$\tilde{X}$ implies the median of the feature; X denotes the original feature value; Q_3 represents the third quartile; Q_1 refers to the first quartile.

3.3. Feature Selection

The HSIC-LASSO is a feature selection approach that incorporates the HSIC with LASSO regularization for identifying informative and non-redundant attributes. HSIC-LASSO is employed to choose the most relevant psychosocial factors' relationship with women's mental health and empowerment, while minimizing redundant and low-informative variables [23]. This helps to simplify the input information and ensure the classification strategy concentrates on the psychosocial factors that contribute most to the valuation results. Let $y = (y_1, y_2, \dots, y_m)^T \in \mathbb{R}^m$ is the output label vector, a weight vector $w = (w_1, \dots, w_d)^T$ be the weight vector. Let $X = [x_1, \dots, x_m]^T \in \mathbb{R}^{m \times d}$ represent the input data, which has m instances with d attributes. K and L have kernel matrices that are determined on the input data X and output y , respectively. HSIC LASSO is handled by the LASSO regularization term, which upholds a sparse selection of psychosocial factors, though the HSIC component quantifies their reliance on the target output, hence helps to hold relevant factors and decrease repetition.

$$w^* = \arg \min \frac{1}{2} \left\| \bar{L} - \sum_{i=1}^d w_j \bar{K} \right\|_{Frob}^2 + \lambda \|w\|_1 \quad (3)$$

$$s. t. w_j \geq 0, i = 1, \dots, d$$

here w_j implies the regression co-efficient of the k_{th} feature, λ indicates the regularization parameter, and $\|\cdot\|_{Frob}$ depicts the Frobenius norm, $w = [w_1, \dots, w_d]$ signifies a regression coefficient vector, and K_i implies the kernel matrix on the i -th feature for all samples. Unlike the traditional mRMR methodology, HSIC LASSO utilizes HSIC to compute feature dependence/relevance along with an L_1 regularized feature selection system.

$$\begin{aligned} & \frac{1}{2} \|\bar{L} - \sum_{i=1}^d w_j \bar{K}\|_{Frob}^2 \\ &= \frac{1}{2} \text{HSIC}(y, y) - \sum_{i=1}^d w_j \text{HSIC}(x^{(i)}, y) + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d w_j w_j \text{HSIC}(x^{(i)}, x^{(j)}) \end{aligned} \quad (4)$$

Here, $\text{HSIC}(x^{(i)}, y) = \text{tr}(\bar{K}\bar{L})$ indicates a kernel-based independence measure of HSIC, $x^{(i)}$ specifies the i_{th} feature vector for each instance, and tr refers to the trace. HSIC always gets a non-negative value and becomes 0 if and only if dual random variables were statistically independent. Accordingly, HSIC-LASSO maintains the most informative psychosocial factors while decreasing redundant variables, subsequently offering a relevant attribute set for outcomes of women's mental health and empowerment.

The selected psychosocial factors attained by HSIC-LASSO are then delivered as input to the TabM classification algorithm. TabM utilizes these relevant attributes to learn their associations and assess women's mental health and empowerment.

3.4. TabM-based Classification Module

TabM is utilized to learn complex relationships efficiently from tabular psychosocial data though preserving a compact model structure. A parameter-effective deep method called TabM is especially developed for tabular data [24]. Its fundamental idea is to build different basic structures within a shared MLP architecture and produce distinct feature representations using lightweight parameter modulation. When compared to conventional ensemble techniques that need several independently trained systems, TABM shares backbone-network parameters and utilizes a lesser vector counts for each subsection. This strategy delivers stable learning with lower computational complexity. Consider the input sample to be $x \in \mathbb{R}^p$; here, p represents the input-feature dimension. The standard MLP l_{th} layer can be defined by

$$h^{(0)} = x, h^{(l)} = \phi(W^{(l)}h^{(l-1)} + b^{(l)}), l = 1, \dots, L \quad (5)$$

here $W^{(l)}$ and $b^{(l)}$ depicts the weight matrix and bias term, respectively, $h^{(l)}$ implies the hidden module at layer l , and $\phi(\cdot)$ indicates the nonlinear activation function. This shared MLP backbone is introduced by TabM for K implicit submodels. For the k_{th} submodel, the l_{th} layer representation is determined by:

$$h_k^{(l)} = \phi \left[s_k^{(l)} \odot W^{(l)} \left(r_k^{(l)} \odot h_k^{(l-1)} \right) + b_k^{(l)} \right] \quad (6)$$

Here, $r_k^{(l)}$ and $s_k^{(l)}$ signify the input and output vectors of the k_{th} vector; $W^{(l)}$ demonstrates the shared main weight matrix; $\odot$ represents element-wise multiplication; and $b_k^{(l)}$ refers to the submodel-specific bias term. This method is approximately equivalent to a shared weight matrix low-rank modulation:

$$W_k^{(l)} = W^{(l)} \odot [s_k^{(l)} (r_k^{(l)})^T] \quad (7)$$

Therefore, an entire network for every sub-model does not duplicate by TabM. Rather, it produces multiple mappings using lightweight adaptation vectors and a shared backbone. This framework minimizes the parameter scale whereas improving the capability to depicts nonlinear associations, which makes it appropriate for small-instance, multi-source tabular-data modeling.

During this research, TabM is utilize to learn the correlations among psychosocial factors and assigns women's mental health and empowerment according to the learned pattern features.

$$\hat{y}_k = g(h_k^{(L)}) \quad (8)$$

Here $g(\cdot)$ specifies the mapping that maps the learned representation of the classification head to the corresponding mental health and empowerment categories.

The implicit submodels outputs are integrated to attain the final class decision:

$$\hat{y} = \frac{1}{K} \sum_{k=1}^K \hat{y}_k \quad (9)$$

The training objective is stated by:

$$L_{TabM} = L_{cls}(y, \hat{y}) + \Omega(\theta) \quad (10)$$

Next, to learn the relevant psychosocial patterns in TabM, the model parameter Adan optimizer is applied to update the model and enhance the classification of women's mental health and empowerment performance. Fig. 2 indicates the structure of TabM approach.

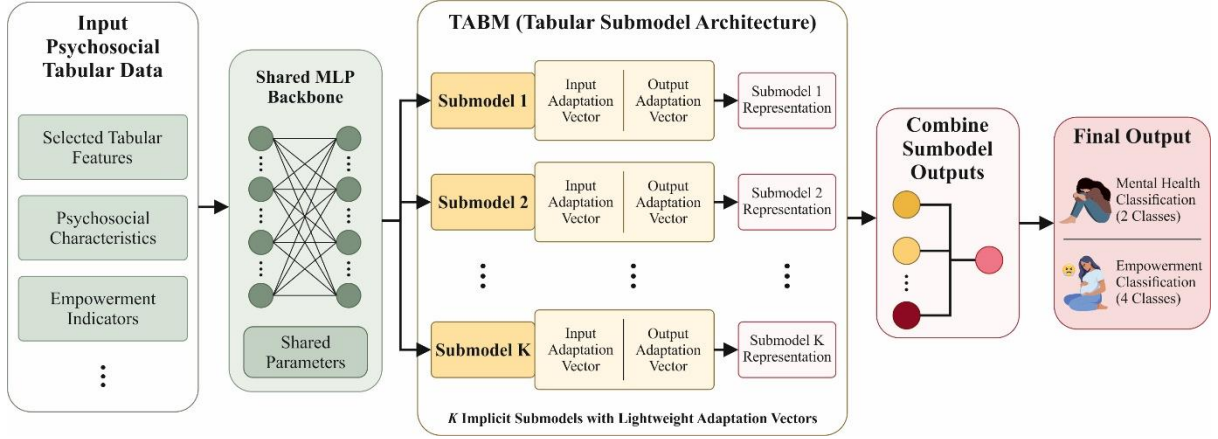


Fig. 2. Structure of TabM approach

3.5. Adan Optimization Strategy

This study employed Adan to effectively optimize the intricate association among psychosocial factors, including women's mental health and empowerment. Its adaptive momentum and Nesterov-driven upgrade system enable the model to effectively learn distinct patterns in psychosocial factors [25]. An adaptive optimization strategy named Adan (Adaptive Nesterov Momentum) that integrates momentum estimation with Nesterov acceleration. In contrast to conventional optimizers that are heavily dependent on first-order gradient momentum, Adan further examines the difference among consecutive gradients, allowing parameter gradients to adapt to variations in the optimization direction. The general Adan update process can be defined by:

$$g_t = \nabla_{\theta} L(\theta_t) \quad (11)$$

$$m_k = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (12)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)(g_t - g_{t-1}) \quad (13)$$

$$\theta_{t+1} = \theta_t - \eta \frac{m_t + \beta_1 v_t}{\sqrt{n_t} + \epsilon} \quad (14)$$

Here m_t refers to the first-order momentum, g_t signifies the current gradient, n_t represents the adaptive second-order term, v_t denotes the gradient variation, η depicts the learning rate, ϵ

indicates a small constant used for numerical stability, and β_1 and β_2 imply momentum coefficients.

Subsequently the model parameter is optimized by Adan, SHAP is used for interpreting the contribution of individual psychosocial factors in predicting mental health and empowerment performances. Fig. 3 represents a flowchart of Adan optimizer.

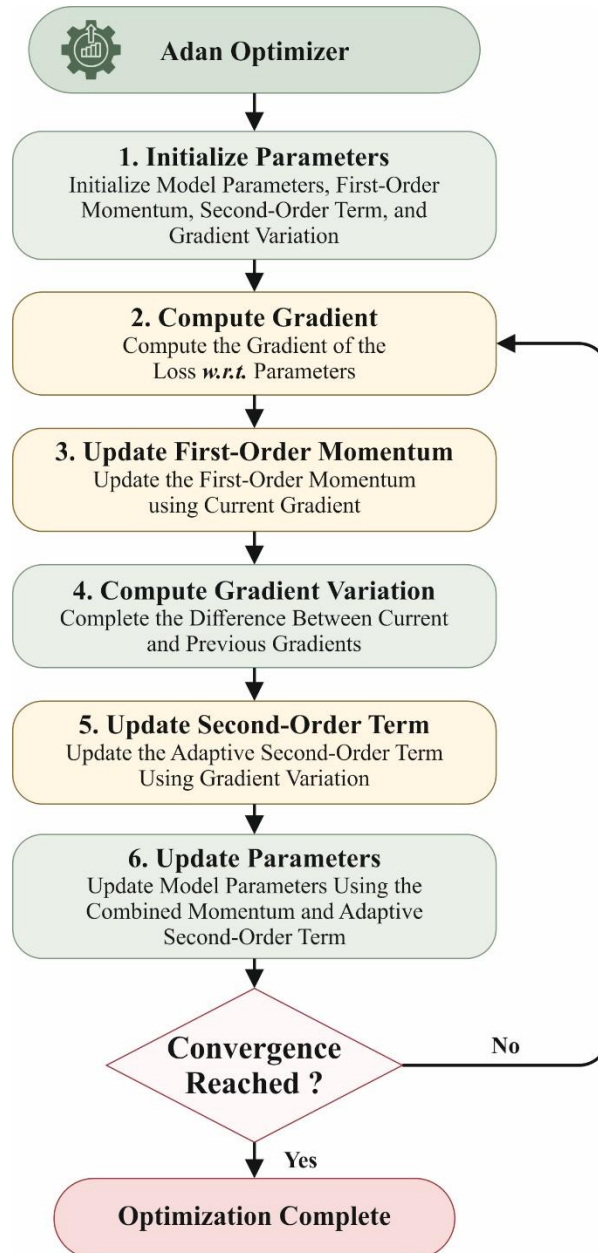


Fig. 3. Flowchart of Adan optimizer

3.6. SHAP-based Explainability Analysis

An explainable algorithm called SHAP, driven by cooperative game theory, computes the contribution of every attribute to a model prediction. By evaluating its contribution, SHAP

assigns a Shapley value to every attribute across diverse incorporations of features, ensuring a clear explanation of the model's decision-making process [26]. To identify and interpret the contribution of individual psychosocial factors, this paper utilized SHAP to assess the impact on women's mental health and empowerment. This helps to identify which factors have a higher effect in predicting performance and offers a clearer interpretation of the deployed predictive method. The contribution of feature (i) through SHAP is expressed by:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! (|N| - |S| - 1)!}{|N|!} |f(S \cup \{i\}) - f(S)| \quad (15)$$

Here, S denotes the feature subset of i , N signifies the set of all attributes, $f(S \cup \{i\})$ implies the prediction when adding i to the subset, ϕ_i implies the contribution of feature i , and $f(S)$ represents the model prediction when only features in S were included.

4. Experimental Analysis

The empirical analysis of the proposed X-WMHEA model is conducted using the Women Depression Dataset and Maternal Depression Dataset, and this section discusses the classification results, comparative analysis, ablation study, and interpretability analysis.

4.1. Result Analysis of the Dataset 1

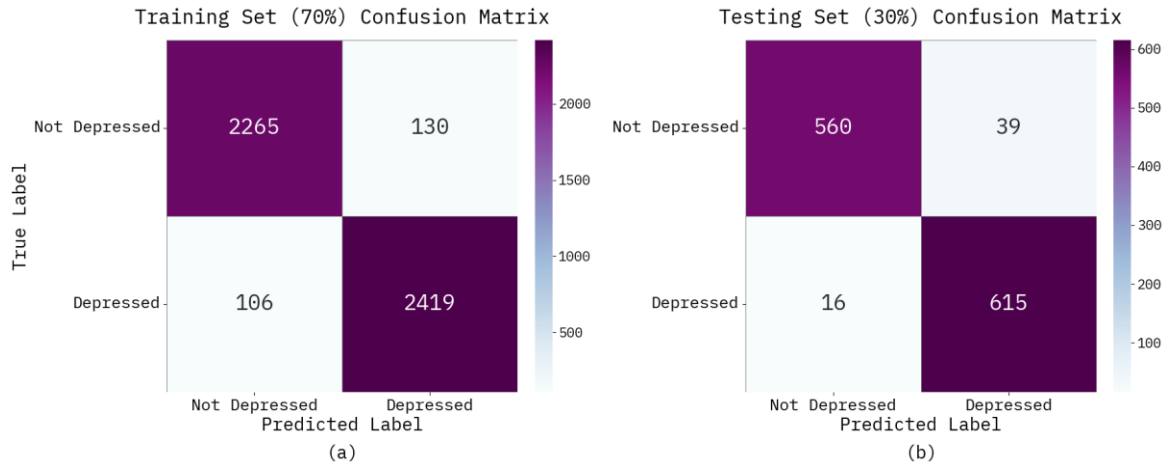


Fig. 4. Classification outcome of (a-b) confusion matrix of the X-WMHEA system using 70:30 for Dataset 1

Fig. 4 implies the confusion matrices of the X-WMHEA method with 70:30 using Dataset 1. Fig. 4a demonstrates a 70% training set; most of the records are properly categorized as not depressed (2,265) and depressed (2,419) samples. Fig. 4b represents the 30% testing set

correctly identifies 560 not-depressed and 615 depressed records, indicating a reliable classification outcome, with some samples misclassified in both phases.

Table 2 and Fig. 5 exhibit the classification results of the X-WMHEA system with a 70:30 for Dataset 1. The proposed X-WMHEA algorithm achieved an average accuracy of 95.20% in training and 95.53% in testing, while the MCC was 90.40% and 91.11% for both phases, representing its consistent classification capability and robust generalization performance for women's depression prediction.

Table 2 Classification results of the X-WMHEA algorithm with 70:30 for Dataset 1

Classes	Accuracy	Precision	Recall	F1-Score	MCC
Training Set (70%)	95.20	95.22	95.19	95.20	90.40
Testing Set (30%)	95.53	95.63	95.48	95.52	91.11

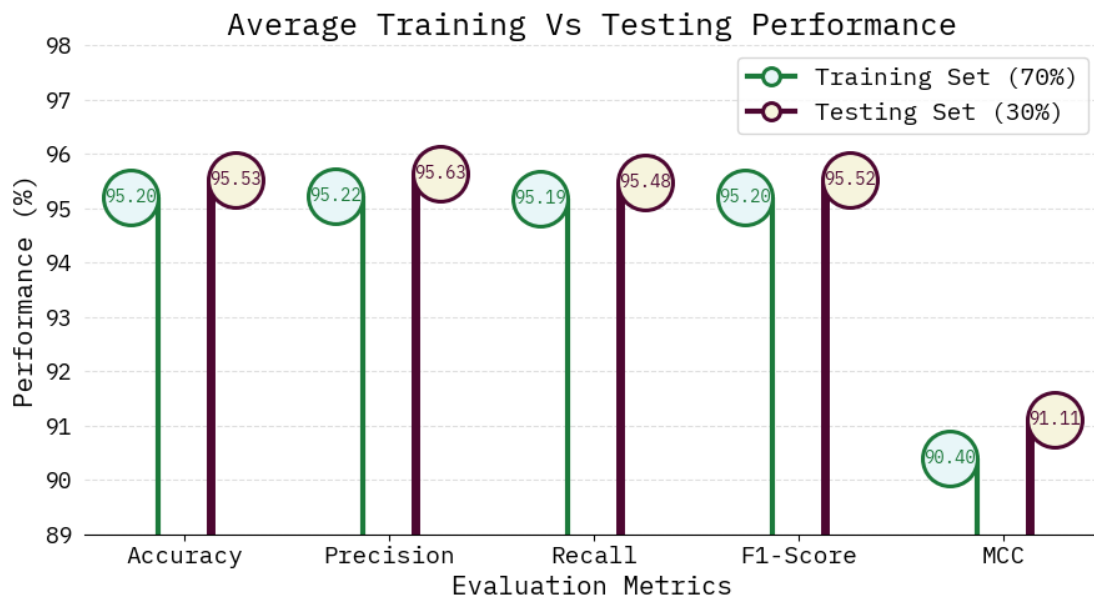


Fig. 5. Average training and testing performance of the X-WMHEA method for Dataset 1

Fig. 6 illustrates the ROC and PR curve of the X-WMHEA methodology with Dataset 1. Fig. 6a shows the ROC curve maintains an AUC score of 0.990 for both classes, signifying robust classification results. Fig. 6b portrays the PR curve remaining higher for the outcome, with AP values of 0.989 for not depressed and 0.992 for depressed types.

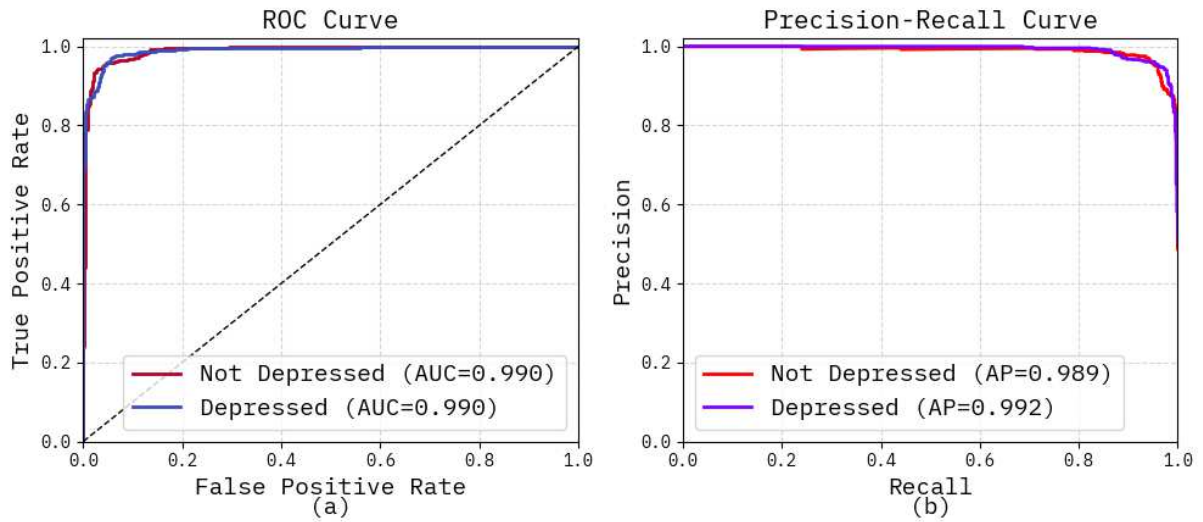


Fig. 6. (a) ROC and (b) PR curve of the X-WMHEA system for Dataset 1

Fig. 7 describes the accuracy and loss curve of the X-WMHEA system using Dataset 1. Fig. 7a demonstrates that the accuracy curve sustains steady improvement over the epochs, with both curves achieving their greatest values toward the end of the training phase. Fig. 7b portrays the loss curve gradually declining with increasing epochs, signifying that the model efficiently learns and obtains stable performance. The close alignment between training and testing outcomes suggests good generalization without overfitting.

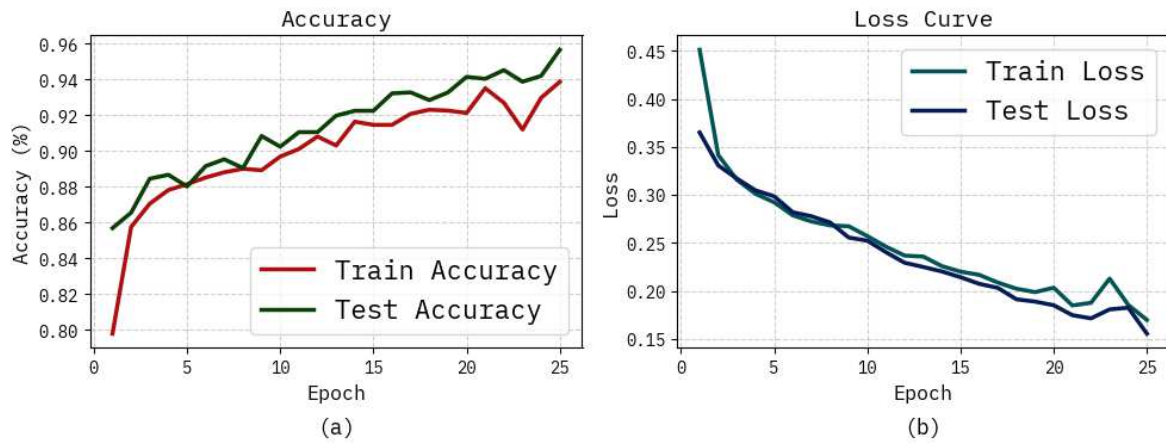


Fig. 7. (a) Accuracy and (b) loss curve of the X-WMHEA model for Dataset 1

4.2. Result Analysis of the Dataset 2

Fig. 8 represents the confusion matrix attained for the training and testing of the X-WMHEA method for Dataset 2. Most instances are properly classified along the diagonal, with Not Depressed having the maximum correct predictions in both phases, along with 138 samples in training and 59 samples in testing. Just a small number of cases are miscategorized, particularly

among the various depression levels, indicating good prediction performance as well as generalization results.

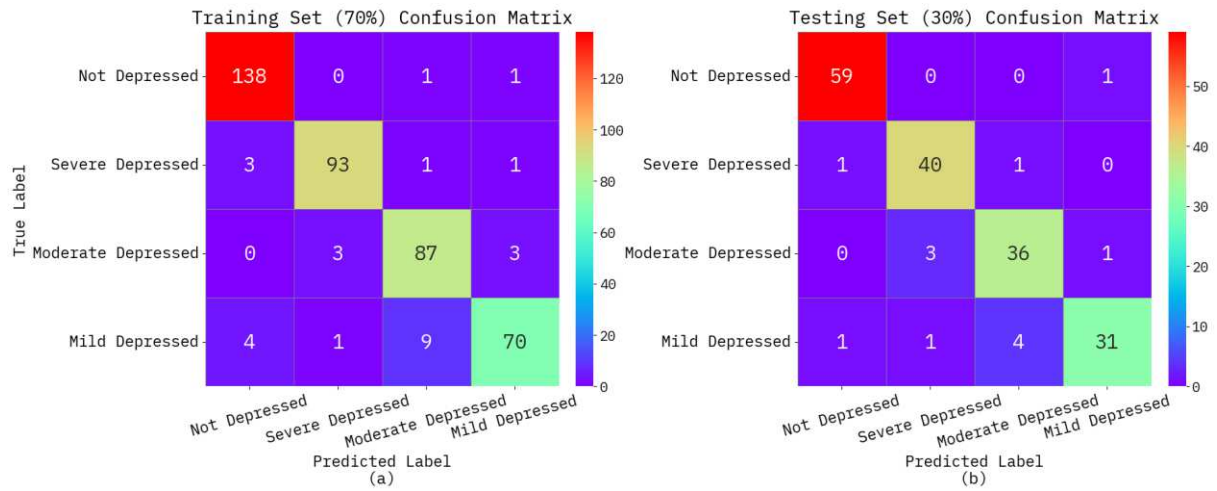


Fig. 8. Classification outcome of (a-b) confusion matrix of the X-WMHEA system using 70:30 for Dataset 2

Table 3 and Fig. 9 illustrate the classifier outcomes of the X-WMHEA method utilizing a 70:30 split with Dataset 2. The presented X-WMHEA method attained an accuracy of 96.75% with an SD of $\pm 1.27\%$ in the 70% training phase, whereas it obtained an accuracy of 96.37% with an SD of $\pm 1.48\%$ on the 30% testing set, among other evaluation metrics. The comparable results achieved in both phases validate that the presented X-WMHEA method effectively learns discriminative attributes and generalizes well to unnoticed maternal depression recognition.

Table 3 Classification results of the X-WMHEA system using 70:30 for Dataset 2

Classes	Accuracy	Precision	Recall	F1-Score	MCC
Training Set (70%)					
Not Depressed	97.83	95.17	98.57	96.84	95.23
Severe Depressed	97.83	95.88	94.90	95.38	93.97
Moderate Depressed	95.90	88.78	93.55	91.10	88.49
Mild Depressed	95.42	93.33	83.33	88.05	85.44
Average	96.75	93.29	92.59	92.84	90.78
Standard Deviation	± 1.27	± 3.19	± 6.53	± 4.02	± 4.61
Testing Set (30%)					
Not Depressed	98.32	96.72	98.33	97.52	96.26
Severe Depressed	96.65	90.91	95.24	93.02	90.86
Moderate Depressed	94.97	87.80	90.00	88.89	85.65
Mild Depressed	95.53	93.94	83.78	88.57	86.02
Average	96.37	92.34	91.84	92.00	89.70
Standard Deviation	± 1.48	± 3.85	± 6.38	± 4.20	± 4.98

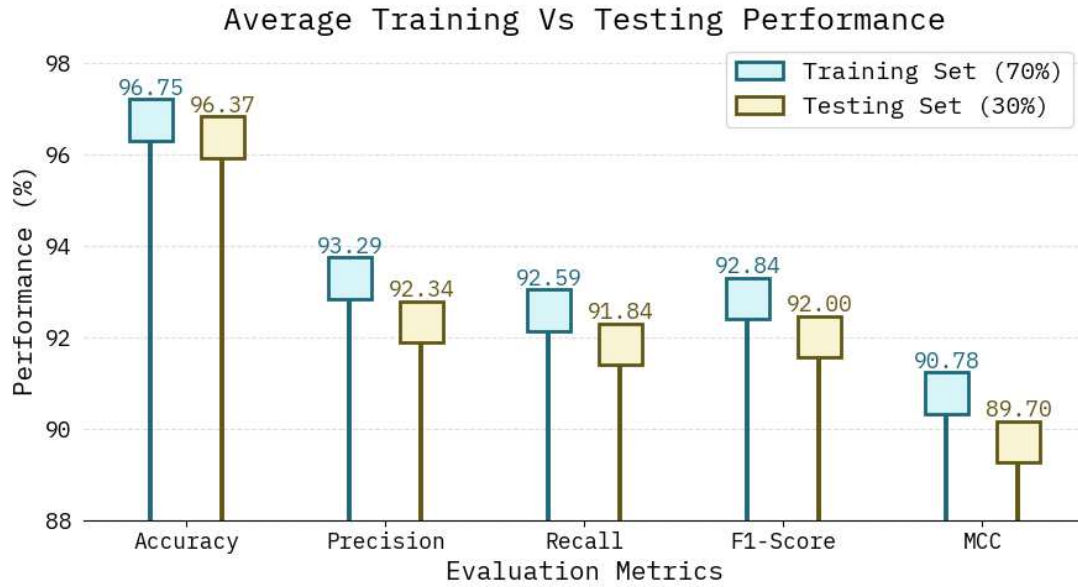


Fig. 9. Average training and testing performance of the X-WMHEA algorithm for Dataset 2

Fig. 10 represents the ROC and PR curves of the X-WMHEA method with Dataset 2. Fig. 10a represents the ROC curve of four different thresholds. All curves maintain a top-left direction, representing remarkable prediction results. The AUC values are extremely high, as represented in the figure. Fig. 10b shows the ROC curve maintains high precision while identifying various classes. The curve remains close to 1.0 in precision, presenting robust results, and the AP values are also high.

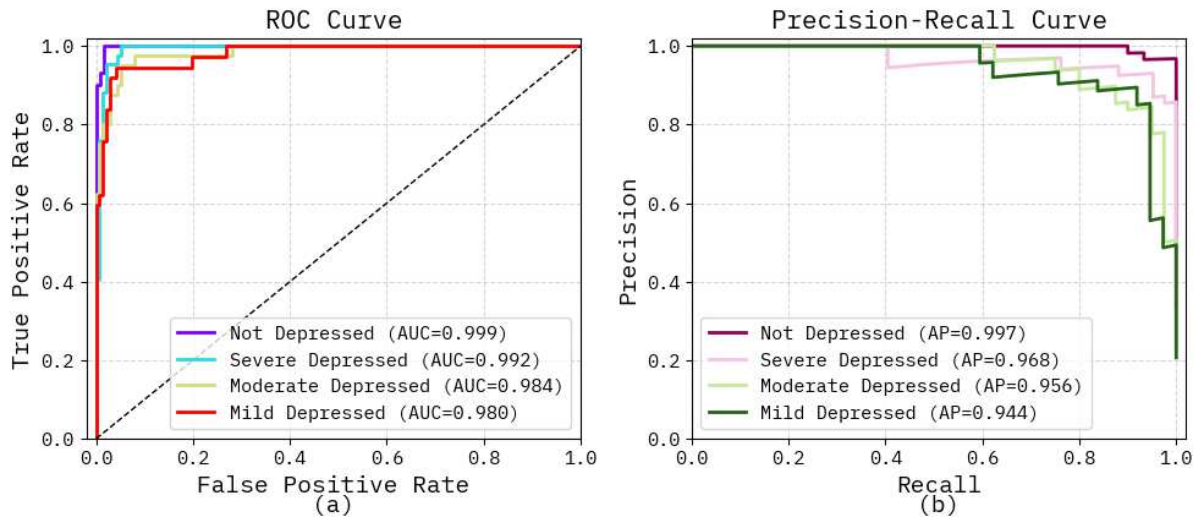


Fig. 10. (a) ROC and (b) PR curve of the X-WMHEA system for Dataset 2

Fig. 11 represents the accuracy and loss curve of the X-WMHEA method with Dataset 2. Fig 11a represents the accuracy curve, which increases gradually with each epoch. Both reach

around 90% or above, and the curves remain close, indicating that the model executes well on both training and unseen data. Fig 11b represents the loss curve, which decreases continuously as the epochs increase. This denotes that the method is learning efficiently, with the low and closely matched losses representing good generalization.

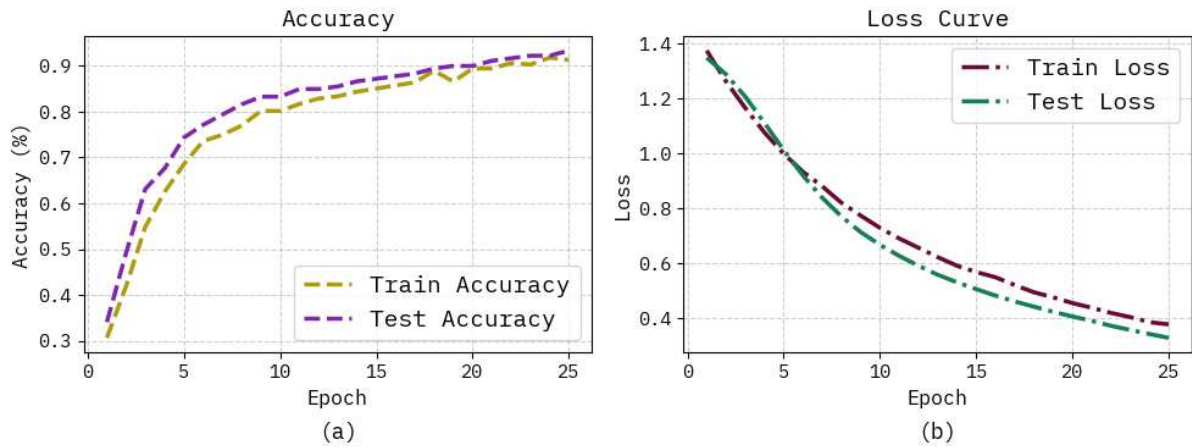


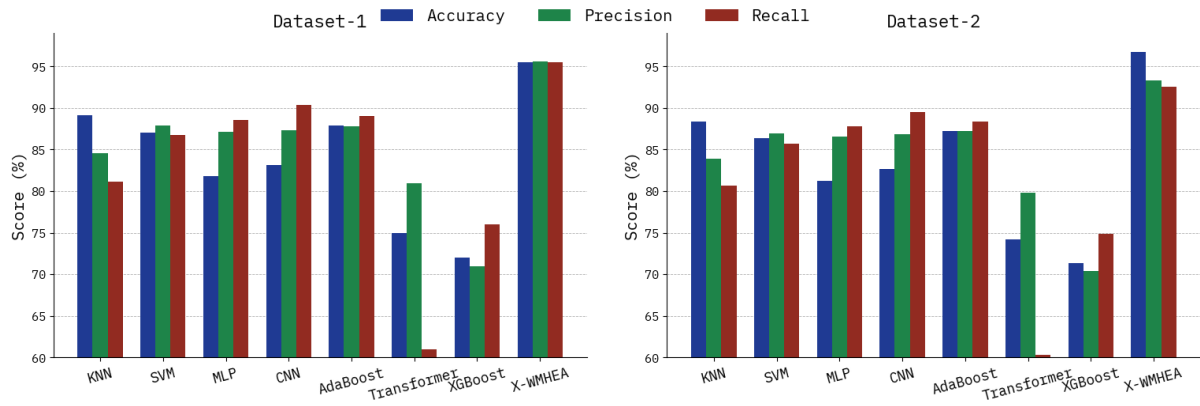
Fig. 11. (a) Accuracy and (b) loss curve of the X-WMHEA model for Dataset 2

4.3. Comparative Analysis and Ablation Study

Table 4 and Fig. 12 represent the comparison performance of the X-WMHEA method with current approaches utilizing various evaluation metrics [27, 28]. On dataset 1, the outcome implies that the XGBoost [28], Transformer [28], and MLP [27] models have attained a low accuracy value when compared with other techniques. However, MLP [27] and CNN [27] methods have moderately low outcomes, while the SVM [27], KNN [27], and AdaBoost [27] models have achieved significantly higher results. Concurrently, the presented X-WMHEA model has attained the highest result with an accuracy of 95.53%, a precision of 95.63%, and a recall of 95.48%, representing its efficiency as a potential algorithm for women's depression prediction. On dataset 2, the results suggest that the Transformer [28], XGBoost [28], and MLP [27] models have attained a low accuracy value. However, CNN [27] and MLP [27] methods have minimal to moderate outcomes, while the SVM [27], KNN [27], and AdaBoost [27] models have achieved significantly higher results. Concurrently, the presented X-WMHEA model has attained greater results with an accuracy of 96.75%, a precision of 93.29%, and a recall of 92.59%, demonstrating its efficiency as a potential algorithm for maternal depression prediction.

Table 4 Comparative analysis of the X-WMHEA algorithm with various measures

Methods	Women Depression Dataset	Accuracy	Precision	Recall	Maternal Depression Dataset	Accuracy	Precision	Recall
KNN [27]		89.10	84.60	81.10		88.42	83.91	80.63
SVM [27]		87.00	87.90	86.80		86.35	86.92	85.74
MLP [27]		81.80	87.10	88.60		81.24	86.53	87.82
CNN [27]		83.10	87.30	90.40		82.67	86.84	89.51
AdaBoost [27]		87.90	87.80	89.00		87.26	87.21	88.37
Transformer [28]		75.00	81.00	61.00		74.18	79.84	60.27
XGBoost [28]		72.00	71.00	76.00		71.36	70.42	74.91
X-WMHEA		95.53	95.63	95.48		96.75	93.29	92.59

**Fig. 12.** Comparative study of the X-WMHEA methods using various metrics

The ablation study demonstrating how each component supports the performance of the X-WMHEA method is represented in Table 5. The baseline TabM method attains an accuracy of 92.84% with an F1-score of 89.14%. Incorporating Adan enhances the results to 93.76% accuracy and 90.18% F1-score, indicating that the optimizer supports the model to learn more efficiently. When HSIC-LASSO is integrated to TabM, the performance increases further, attaining 94.68% accuracy and 91.11% F1-score. Finally, the X-WMHEA (HSIC-LASSO + TabM + Adan) method attains higher outcomes with 96.75% accuracy, 93.29% precision, 92.59% recall, and 92.84% F1-score, representing that each component offers robust and most balanced performance.

Table 5 Ablation study of the X-WMHEA algorithm with existing methods

Methods	Accuracy	Precision	Recall	F1-Score
TabM	92.84	89.54	88.76	89.14
TabM + Adan	93.76	90.68	89.72	90.18
HSIC-LASSO + TabM	94.68	91.42	90.83	91.11
X-WMHEA (HSIC-LASSO + TabM + Adan)	96.75	93.29	92.59	92.84

4.4. Model Interpretability

Fig. 13 represents the factors have major contribution on the model's prediction Dataset 1. Medical Symptoms and Schooling Status are the most significant attributes, followed by Peer Pressure and Academic Performance. Features like Media Exposure and Sexual Abuse also have a noticeable contribution, but their influence is comparatively lower. The local explanation plot shows that higher or lower feature values can push the model's prediction in several directions. Thus, the outcome represents indicate that health, education, and social factors play a significant role in the model's decision-making.

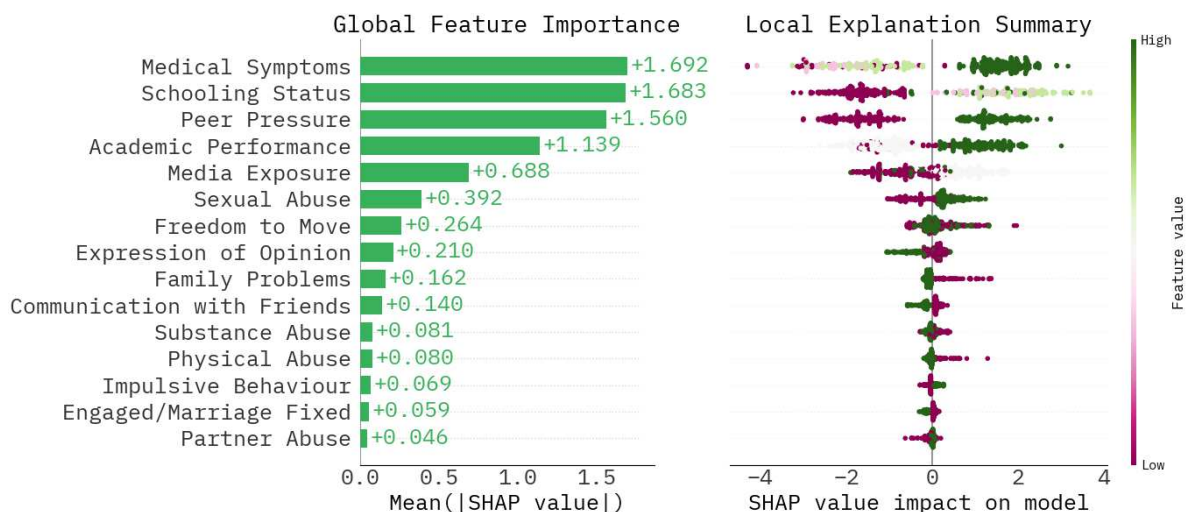


Fig. 13. Global feature importance and local SHAP-based explanation of the model prediction on Dataset 1

Fig. 14 compares the significant factors influencing predictions around 4 depression levels on Dataset 2. For not depressed, occupation, age range, and religion status have the strongest influence, whereas mild depression is mainly connected with mental clarity, heart condition, and age range. For severe depression, religion status, occupation, and marital status are the

most influential factors. In the moderate depression group, sleep problems, tiredness without a clear reason, and faster heartbeat show the highest importance. Therefore, the outcome exhibits that various social, personal, and physical factors contribute differently to every level of depressions.

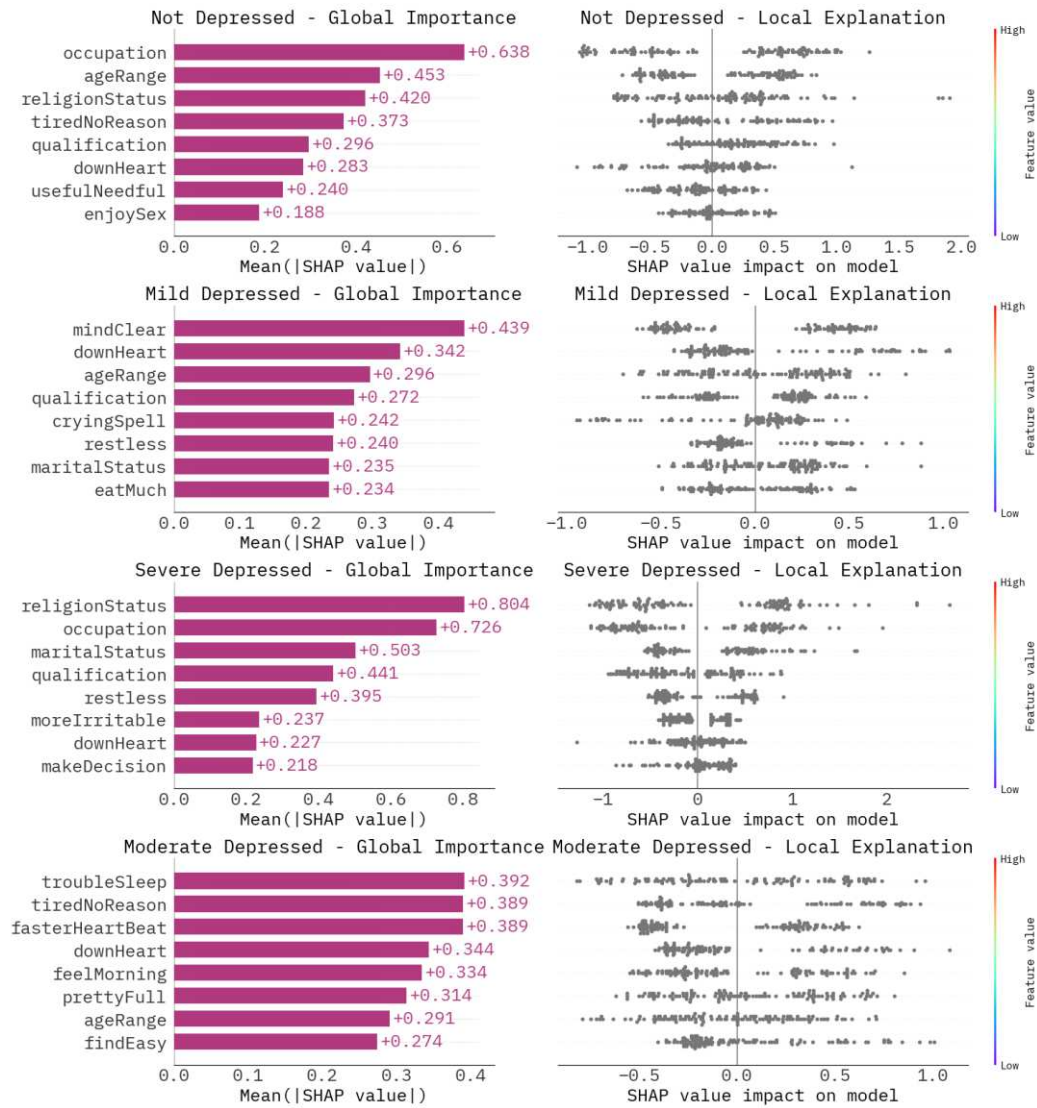


Fig. 14. Global feature importance and local SHAP explanations for different levels of depression Dataset 2

The X-WMHEA model in this study provides an efficient way to analyze women's mental health and empowerment using key psychosocial features. HSIC-LASSO selects relevant features and TabM efficiently models interactions between selected features for efficient classification, while the Adan optimizer and SHAP assist in efficiently training the model and illustrating feature interactions, respectively. The ablation analysis, as well as the results obtained on both datasets, highlight the contribution of individual modules.

5. Conclusion and Future Works

This study presented the X-WMHEA model using psychosocial factors for effective analysis. The proposed model employs a feature selection strategy that integrates HSIC and LASSO to identify the most relevant and informative features while reducing redundancy. Moreover, the TabM approach is utilized to model complex relationships among selected features and improve the classification performance. Furthermore, the Adan optimizer is employed to update the model parameters during training and achieve stable convergence, while SHAP is used to explain the model predictions and determine the contributions of the individual features. Comprehensive experimental analysis of the X-WMHEA model is performed using the Women Depression Dataset and Maternal Depression Dataset, achieving accuracy of 95.53% and 96.75%, outperforming existing models across various measures. These results highlight its effectiveness in women's mental health analysis and empowerment assessment. In future work, the model can analyze different and a greater number of datasets for enhancing the generalization capabilities over a wide range of the women population. Furthermore, a larger number of psychosocial and socio-economic attributes can be explored as factors of interest for a holistic assessment. Comprehensive evaluations can be performed on more sophisticated models along with different state-of-the-art explanation mechanisms for more robust outcomes.

Declarations

Data Availability Statement: The data that support the findings of this study are openly available at <https://www.kaggle.com/datasets/ucimachinelearning/women-depression-dataset> and <https://data.mendeley.com/datasets/f2r8whrtyx/2>, reference number [20, 21].

Conflict of Interest: The authors declare no conflict of interests.

Competing interests: The authors declare no competing interests.

Ethics Approval: Not Applicable

Consent to Participate: Not applicable.

Informed Consent: Not applicable.

Clinical Trial Registration: Not applicable.

Author Contribution: “Conceptualization, methodology, software, validation, writing—original draft preparation: H.M; formal analysis, investigation, writing—review and editing,

supervision, project administration: S.B.A. All authors have read and agreed to the published version of the manuscript.”

Funding Information: Funding will be provided based on the acceptance.

References

- [1] Esteban-Gonzalo, S., Khan, M.H., Hasan, M.M. and Díaz-Morales, J.F., 2026. Psychosocial predictors of women's mental health in Bangladesh: Evidence from BDHS-2022 data on fertility, empowerment, decision-making, child loss, and sex-affective behaviours. *Public Health*, 254, p.106222.
- [2] Alam, A., Prodhan, M.E., Akhter, T., Sagar, N.I. and Hasan, M.N., 2026. SWPER-Measured women's empowerment and mental health in Bangladesh: a cross-sectional study from 2022 demographic and health survey. *BMC Women's Health*, 26(1), p.119.
- [3] Mounika, C., Pradhan, K., Seepana, A.K., Chauhan, J.K. and Rajasri, S., 2026. Ecological Empowerment Through Nutritional Security: A Path Analysis of Farm Women's Socio-Economic Attributes. *Indian Research Journal of Extension Education*, 26(1), pp.101-106.
- [4] Sun, S. and An, S., 2026. How the partner's perceived psychological state affects perinatal mental health in Chinese women: multiple mediating effects of couple communication and perceived social support. *Psychology, Health & Medicine*, 31(2), pp.313-329.
- [5] Araby, O.A.W.A., El Desoky, N.A.E.I., Hendy, A., Aljaloud, T.R. and Afify, S.S.M., 2026. Empowering menopausal women to adopt climate-adaptive health behaviors and reduce climate-sensitive health outcomes: an IMB model-based quasi-experimental study. *Frontiers in Public Health*, 14, p.1908834.
- [6] Holod, A.F., Tornwall, J., Teall, A.M. and Overcash, J., 2025. Competency in trauma-informed care: Empowering advanced practice registered nursing students to adopt a trauma-informed approach during routine health assessment. *Nurse Education Today*, 144, p.106477.
- [7] Xiang, X., Turner, S., Ruiz-Sierra, S., Zheng, C., Ash, S., Kodkany, N., Dimoji, A., Gosdzinski, C. and Kubicek, G., 2026. Older adults experience with a layperson-supported digital mental health intervention for depression: qualitative insights on engagement. *Clinical gerontologist*, 49(2), pp.295-306.

- [8] Bourke-Taylor, H.M., Leo, M. and Tirlea, L., 2025. Health, wellbeing and empowerment e-workshops for mothers of children with disabilities: A non-randomised comparison study. *Journal of Autism and Developmental Disorders*, 55(4), pp.1465-1480.
- [9] Koly, K.N., Saba, J., Mallick, T., Rashid, F., Watson, J. and Neves, B.B., 2025. Exploring the pattern of mental health support-seeking behaviour and related barriers among women experiencing intimate partner violence in urban slums of Bangladesh: perspectives from multiple level stakeholders. *PLOS global public health*, 5(5), p.e0004568.
- [10] Wuryan, S., Ghofur, R.A., Jafar, M.M., Sanjaya, S. and Setiawati, R., 2025. Women's empowerment model in increasing economic independence at the community work training center (BLKK) of Lampung Province. *Journal of Community Service and Empowerment*, 6(1), pp.58-73.
- [11] Bageant, E., Lentz, E., Narayanan, S., Jensen, N. and Lepariyo, W., 2024. How do women's empowerment metrics measure up? A comparative analysis. *Food Policy*, 129, p.102764.
- [12] Akter, T., Hassan, R., Hossain, M.S. and Saha, S., 2026. Women's empowerment and health: nationwide insights on selected non-communicable conditions in Bangladesh. *BMC Public Health*.
- [13] Kumari, A., Women Empowerment: A Socio-Economic and Psychological Perspective. Future Fact Society Varanasi (UP) India, p.61.
- [14] Ferrando-Terradez, I., Núñez-Cortés, R., López-Bueno, L., Alcántara, E., Calatayud, J., Casaña, J., Parčina, I. and Dueñas, L., 2025. Relationship between psychological empowerment, physical activity enjoyment, and response to a HIIT Program in physically inactive young women: a prospective multicentre study. *BMC women's health*, 25(1), p.38.
- [15] Deroncel-Acosta, A., Norabuena-Figueroa, R.P. and Norabuena-Figueroa, E.D., 2025. Women's mental health in the doctoral context: Protective function of the psychological capital and academic motivation. *Women's Health*, 21, p.17455057251315318.
- [16] Al-Mamary, Y.H., Abubakar, A.A. and Abdulmohsen Alfalah, A., 2025. The effect of psychological, market, economic, socio-cultural, entrepreneurial orientation and educational and skills factors on empowering women for new venture creation. *Humanities and Social Sciences Communications*, 12(1), p.840.
- [17] Vicent, L., Senyonga, L., Namagembe, S. and Nantumbwe, S., 2025. Analysis of the impact of women's empowerment and social network connections on the adoption and

- sustained use of clean cooking fuels and technologies in Uganda. *Energy Policy*, 198, p.114435.
- [18] Kumari, D., Giri, A.K. and Saruparia, C., 2025. Role of gender-based digital financial inclusion and women empowerment in poverty reduction: evidence from Asian countries. *Discover Sustainability*, 6(1), p.283.
 - [19] Shawon, M.S.R., Hossain, F.B., Ahmed, R., Poly, I.J., Hasan, M. and Rahman, M.R., 2024. Role of women empowerment on mental health problems and care-seeking behavior among married women in Nepal: secondary analysis of nationally representative data: Shawon et al. *Archives of Women's Mental Health*, 27(4), pp.527-536.
 - [20] <https://www.kaggle.com/datasets/ucimachinelearning/women-depression-dataset>
 - [21] <https://data.mendeley.com/datasets/f2r8whrtyx/2>
 - [22] Salman, O.S., Latiff, N.M.A., Ariffin, S.H.S., Salman, O.H. and Abdalkareem, Z.A., 2026. Evaluation of triage performance in IoMT-based telemedicine using robust scaler and PCA preprocessing. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 15(1), p.63.
 - [23] Zhang, X., Wang, T. and Lai, Z., 2024. A feature-weighted support vector regression machine based on Hilbert–Schmidt independence criterion least absolute shrinkage and selection operator. *Information*, 15(10), p.639.
 - [24] Yuan, G., Ma, L., Zhang, P., Zhang, L., Lu, Z. and Cao, Y., 2026. Joint Prediction of Goaf Temperature and CO Concentration Using Multi-Source Monitoring Feature Fusion and CA-WOA-Optimized Models. *Applied Sciences*, 16(15), p.7422.
 - [25] Xie, C., Zhang, X., Li, L., Meng, H., Zhang, T., Li, T. and Zhao, X., 2023, June. Large kernel distillation network for efficient single image super-resolution. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 1283-1292). IEEE.
 - [26] Azlaoui, M., Karef, S., Foufou, A., Haied, N., Azlaoui, N., Rabehi, A., Habib, M. and Zeddouri, A., 2026. Machine Learning-Based Prediction of Irrigation Water Quality Index with SHAP Interpretability: Application to Groundwater Resources in the Semi-Arid Region, Algeria. *Water*, 18(8), p.959.
 - [27] Al Masud, G.H., Shanto, R.I., Sakin, I. and Kabir, M.R., 2025. Effective depression detection and interpretation: Integrating machine learning, deep learning, language models, and explainable AI. *Array*, 25, p.100375.

- [28] Jui, S.J.J., Deo, R.C., Acharya, R., Barua, P.D., Soar, J. and Devi, A., 2026. Explainable stress detection system using hybrid CNN-Transformer and uncertainty quantification with copula GAN-based data. *Biomedical Signal Processing and Control*, 120, p.109934.