Received: January 31, 2024. Revised: March 19, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Predicting Postpartum Depression with Aid of Social Media Texts Using Optimized Machine Learning Model

Devasagayam Suganthi[1]* **Arumugam Geetha[2]**

[1]*Department of Computer Science, Chikkanna Government Arts College, Tiruppur - 641602, Tamil Nadu, India,*
[2]*Department of Computer Science, Government Arts and Science College, Avinashi - 641654, Tamil Nadu, India*
* Corresponding author's Email: suga07bca@gmail.com

Abstract: Post-Partum Depression (PPD) is a significant medical condition that occurs in some women after childbirth as a consequence of physiological behavioral and mental alterations. The complicated symptoms of this condition make it difficult to diagnose and differentiate from other conditions. Timely detection and diagnosis of PPD are crucial for controlling morality rates and ensuring effective treatment. Various Machine Learning (ML) models was developed to predict PPD based on the patients demographic status, mental health history and vital signs. But, additional psychological attributes are needed for predicting mental status and identifying individuals with PPD risk. Also, cost-effective and innovative methods are needed to identify individuals with PPD and detect potential development tendencies. Hence, in this paper, Osprey Parameter Optimized MLP (OPOMLP) is developed to address the above issues for efficient PPD. Initially, the Application Programming Interface (API) function of online social network like Twitter (tweets) and Instagram (comments) are used to collect the data posted by health care professionals. Then, Natural Language Processing (NLP) is utilized to pre-process the collected data \and extract relevant text of Twitter users to estimate PPD phases. The additional psychological attributes like mental health and behavioural changes attributes of women are extracted using Linguistic Inquiry Word Count (LIWC) and Latent Semantic Analysis (LSA) methods. Next, MLP network is trained using the extracted attributes along with the attributes of demographic status, mental health history and vital signs. Since, the parameters of MLP were not optimized properly which leads to computational complexities in the PPD prediction. So, the weights initialization and the hyper-parameters of MLP is optimized simultaneously by using an Osprey Optimization Algorithm (OOA). OOA is a metaheuristic optimization algorithm derived from osprey bird hunting behavior which aims to find the global optimum solutions and reduces the complex optimization issues. The relationship between hyperparameters and classification performance will identify an optimal hyperparameter space regions for optimal classification with less computational time and resources. Finally, the OPOMLP is employed for the final prediction of PPD. The test outcomes reveal that the OPOMLP model achieves an accuracy of 96.12% on the collected dataset compared to the classical PPD detection models.

Keywords: Postpartum depression, Machine learning, Electronic health records, Application programming interface, Osprey optimization.

1. Introduction

PPD is a severe mental condition that can develop within a year before childbirth, leading to maternal postnatal death [1]. It negatively impacts the physiological and emotional wellbeing of new mothers and hinders the growth and progression of newborns [2]. PPD symptoms are linked to inadequate mother-baby relationships, neonatal physiological and psychological analysis, linguistic growth, infant reactions and napping coordination, making it the most common cause of postpartum hemorrhage in women postpartum [3, 4]. New mothers may experience symptoms such as difficulty dropping or remaining drowsy, extended slumbering, mood fluctuations, starvation, fear of harming others, extreme anxiety, melancholy and self-harm [5, 6]. Early prevention can reduce mortality rates, improve the physiological and mental health of new mothers and enhance the infant's well-being.

Received: January 31, 2024. Revised: March 19, 2024.

Pharmacological or psychological interventions can reduce perinatal anxiety and improve maternal and infant outcomes. However, determining antidepressant drug exposure during pregnancy and breastfeeding remains a concern [7, 8]. Despite awareness of PPD-associated variables, no statistical sensitivity evaluation techniques are available for prenatal monitoring or medical care. Primary care EHRs can help in risk assessment and illness prediction, combining patient information with socio-demographic details [9, 10].

In order to diagnose PPD, psychiatrists and clinicians often employ the Edinburgh Postnatal Depression Scale (EPDS) [11]. Medical professionals reviewed the EPDS scores after one week of data collection from new moms and those exhibiting PDD symptoms were identified. Patient Health Questionnaire-9 (PHQ-9) and the Postpartum Depressive Surveillance Scale (PDSS) surveys were collected from the aforementioned subjects up to six weeks as part of a multistage progress follow-up [12]. EPDS testing acknowledges postpartum attitude disturbance but overlooks prevalent signs in women during reproductive periods like impatience and anxiousness. [13]. Nevertheless, The EPDS assessment, which does not include demographic data or social support information, may not be sensitive enough to identify various conditions before and after giving birth and it also performs worse in normal populations compared to validation cohorts in terms of positively forecasted value.

To address the above-mentioned issues, ML models have been developed to predict PPD early, efficiently categorize large EHR data, reduce computational issues and improve decision-making in psychiatry [15]. For instances, ML algorithm was developed [16] to predict the risk of preeclampsia in pregnant women by pre-processing and normalizing questionnaire EHR data, including socioeconomic health variables, treatments, medications, procedures, laboratory measures and demographic status. The basic characteristics were selected using structured queries and NLP [16] Finally, various ML models, such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), XGBoost (XGB) and Multilayer Perceptron (MLP) were used to identify shared patterns among characteristics that provide the greatest degree of result class discrimination for PPD identification.

However, this model considers common attributes of PPD patients, including intellectual well-being, personality characteristics, information positions, wellness assistance, cognitive wellness background, recently treated mental disorders during pregnancy, additional obstetrics or health-related illnesses and other essential symptoms. However, it needs to consider additional physiological attributes like mental health and behavior changes to predict the mental status of women with PPD risk. Additionally, cost-effective and innovative methods are needed to identify individuals with PPD and detect potential development tendencies.

In this paper, the cost-effective model OPOMLP is developed for the PPD prediction using diverse PPD attributes including physiological attributes (mental health and mood changes from the initial day), patients demographic status, mental health history and vital signs. Initially, the API services of Twitter and Instagram is used to get tweets and comments shared by health care professionals. Then, the feature extraction is performed using the LIWC and LDA schemes to extract the attributes regarding mental health and mood changes of the women's with PPD. The extracted features are trained using MLP for PPD prediction. But, the parameters of MLP were not optimized properly which leads to computational complexities in the PPD prediction. The weights initialization and the hyper-parameters of MLP is optimized simultaneously by using an OOA.

Some classical metaheuristic optimization models like Four Directed Search [17], Total Interaction [18] Walk-Spread [19] and AttackLeave [20] struggles with non-linear, discontinuous, non-differentiable and high-dimensional optimization problems, leading to unfavorable solutions due to poor local optimality. Also, the performance of these models does not guarantee its similar performance in solving other optimization problems.

OOA is also metaheuristic optimization algorithm derived from osprey bird hunting behavior, offering unique advantages. It mimics the behavior of osprey birds, aiming to find the global optimum while minimizing local optima. This makes OOA suitable for complex optimization issues. It is flexible, converges quickly and simplifies implementation and tuning for specific problems. OOA is robust, withstands noise and uncertainty and doesn't require the gradient of the objective function, making it suitable for non-differentiable or expensive optimization problems.

By considering the advantages of OOA method will identify the ideal hyperparameter space regions for classification with less time and computational resources by analyzing the relationship between each hyperparameters of MLP. Finally, the layers of MLP are optimized using OOA (OPOMLP) which is employed for the final PPD prediction. Thus, this OPOMLP algorithm enhances the efficiency of predicting women's who are at high risk towards the PPD with low cost resources.

The remaining portions of this paper are prepared as, various existing methods used to identify and categorize PPD. The suggested PPD model is presented in Section III. The performance assessment of the existing and suggested models is given in Section IV. The complete investigation is summed up in Section V which also recommends the future scope.

2. Literature survey

A predictive model was developed [21] for PPD depression using ML model. The model used pregnancy risk assessment monitoring system data to analyze imbalances between groups, predicting PPD using RF, SVM and LR. But, accuracy of this model was low as it trained with limited dataset.

An ML model for PPD detection was created [22]. The gradient-boosted DT algorithm was utilized to extract clinical characteristics from EHR data, which were then integrated into an ML model for automated risk analysis for early PPD prediction tasks. However, this accuracy was inefficient due to inappropriate bias selection.

A women with depressive postpartum symptoms was predicted [23] with ML using clinical, demographic and psychometric data from postpartum questionnaires. The collected dataset was pre-processed and normalized using K-Nearest Neighbour (KNN) model. Then the obtained data was given as input to the ML model to predict the depressive postpartum depression. However, this model results with lower F1-Score evaluation.

ML approach was constructed [24] for early detection of PPD in Bangladesh. The survey on socio-demographic questions and EPDS data was collected, pre-processed and augmented using SMOTE and PPD detection was performed using SVM, RF, LR and XGBoost. But, unbalanced sample distribution in the dataset may affect the accuracy results.

In order to predict PPD, a ML approach was constructed [25]. The RF model was used to select relevant features, while the Extremely Randomized Trees (XRT) model was trained to anticipate PPD symptoms and identify risk variables. However, high class disparity in training models hinders precise prediction due to reduced accuracy.

ML model predicted [26] women who are suffering with PPD. The data, including mother's relatives and information-related positions was pre-processed and standardized using Min-Max normalization. It was then used to predict PPD risk levels using Feed-Forward Neural Network (FFANN), RF and SVM. However, insufficient data samples leads to degrade the accuracy performance.

A DT model was developed [27] to predict the likelihood of recurrent post-traumatic stress disorder in pregnant women. The model divided women into two categories: Stable-High-PTS-FC and Stable-Low-PTS-FC and used implicit class assessment to identify women susceptible to Stable-High PTS-FC for early prognosis. However, when the data was increased, the performance accuracy decreased.

A prenatal depression assessment model was developed [28] using ML model, pre-processed and normalized EHR data from a large urban hospital. The data was analyzed using Shapley Addition Elucidation, Diversed Impression and Equivalent Opportunity Difference and fed into an elastic net to classify and predict PPD stages. But, this model failed to identify the optimal features subset which lowers the accuracy rate.

A LR model was employed [29] to detect PPD using HER data. Four models were developed using distributed RF and LR models, including socio-demographic data, pre-pregnancy mental health data, recursive feature removal and simplified pre-pregnancy mental health factors. But, the model's accuracy were significantly lower due to its training on a limited dataset.

An optimization ML model was developed [30] for assessing PPD risk and implementing preventive interventions. It collected and pre-processed EHR data from caesarean delivery patients, used SHAP for data interpretation, developed Propensity Score Matching for PPD incidence comparison and employed XGB for early intervention. However, this accuracy results were reduced due to train the model with predefined parameters.

3. Proposed methodology

In this section, the complete illustration of OPOMLP model for early PPD prediction. The Fig. 1 depicts the outline of the proposed model. Table 1 lists the notations used in this study.

3.1 Dataset description

For this experiment, a Google form is used to deliver a questionnaire that collects 1503 records from a medical institution [31]. Out of the fifteen characteristics in the dataset, ten were chosen, nine of which were utilized for analysis and one of which was the objective feature. The choice for a PPD predictor was "Feeling Anxious," the target features. Also, the Disease Control and Prevention website provides access to relevant keywords for extracting related terms from social media (instagram and twitter) [32], including terms for Symptoms of Postpartum Depression and Risk Factors for

Received: January 31, 2024. Revised: March 19, 2024.

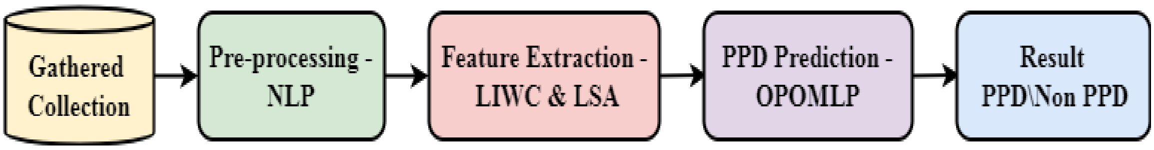

Figure. 1 Entire pipeline of the proposed study

Table 1. Lists of notations

Notations	Description
O_{C_k,R_W}	LIWC outcome
W_{C_k,R_w}	Weight of a Category
d_W	Words category in the database
S	Text topic generated from the actual data
Z	Input matrix column
$Z^\top$	transpose matrix of Z
L	Diagonal matrix with singular values
$\hat{L}$	Non-negative integer matrix with all diagonal members
V	Term-document matrix
$\hat{V}$	Mean Square Error (MSE) approximation of V
H_{1y}	Node y of the hidden layer H_1, is the and
W_{xy}	Input gate of the H_1
B_y	Bias
G	Population matrix of osprey's positions,
G_a	a^{th} osprey (i.e., a candidate solution)
$g_{a,b}$	b^{th} size (i.e., problem variable)
n	Number of ospreys
p	Number of MLP hyperparameters
$\mathcal{L}_b$ and $\mathcal{U}_b$	Lower and upper bound of the b^{th} problem variable
$r_{a,b}$	Random value ranging [0,1]
F	Fitness function values
F_a	F for the a^{th} osprey
T_a	Fish locations for the a^{th} osprey
$G_a^{T_1}$	Location of the a^{th} osprey based on the first phase of OOA
ST_a	Selected fish for a^{th} osprey, is it b^{th} dimension
$ST_{a,b}$	Selected fish for a^{th} osprey with b^{th} dimension
$J_{a,b}$	Random value in the range of {1,2}
e	Number of iterations
E	Greatest number of iterations

Depression. To extract the keywords tweets and comments from the twitter and instagram, this model adopts for the API function Posted by Health Care Professionals.

3.2 Data pre-processing

In this model, various NLP approaches are used to pre-process the data, integrating gathered tweets and comments to provide an extensive collection of characteristics. In this stage, emojis, punctuation, stopwords, distinctive typescripts, superfluous space and special characters in tweets are removed. Before preprocessing, the tweet content is tokenized, stemmed and lemmatized. This step processes redundant or undesired terms stated in tweets and comments, such as typographical mistakes or acronyms of common nouns. After pre-processing, a few online behaviors of Twitter and users, such as emotion, event, online activity, user-specific features and consumption of depression-related texts, posts posted by physicians or depressed people, are extracted to estimate the PPD phases. By deleting duplicate data, these pre-processing methods may lower the overall quantity of information acquired from online social networks concerning PPD instances. In addition to these variables, attributes from online social networks are acquired in order to train the MLP network model.

3.3 Feature extraction

Following the collection of the dataset, the feature extraction approaches listed below are used to extract the physiological aspects of the women's sufferings from PPD phases.

3.3.1 LIWC feature extraction

LIWC is a feature extraction software technique that gathers and classifies language features based on psychological factors. Based on the LIWC lexical dictionary, the proportion of all mental health classes for every woman afflicted with PPD conditions.
Consider, the weight of a category W_{C_k,R_w} determined by the cumulative amount of instances of the words in that category in the database d_W. The outcome of LIWC $\left(O_{C_k,R_W}\right)$ for all C_k classes is then normalized by dividing W_{C_k,R_W} in Eq. (1) by the overall number of data instances in the dataset R_w.

$$O_{C_k,R_W} = \frac{W_{C_k,R_W}}{d_W} \tag{1}$$

LIWC attributes are physiological words such as a woman's mental health behavior that varies considerably from mood swings in women the previous day. It identifies phrases that depressed patients often use and that are associated with their psychological actions as a result of their psychological backgrounds (e.g., sentiments, happiness, regrets and sorrow).

3.3.2 LSA Feature Extraction

It is also known as Latent Semantic Indexing (LSI), is an unsupervised indexing approach used in natural language processing that retrieves semantically linked terms from text sources. The co-occurrence of the approach demonstrates the relationships between the phrases used in the document's sentences to incorporate Singular Value Decomposition (SVD). SVD categorizes words and texts, preserving noiseless data for optimal retrieval. It emphasizes pattern relevance in content. LSI uses SVD to find semantically related terms hidden in a given material. The LSI approach develops a term document matrix and the words from the document generate the SVD to find the appropriate word.

Assume V is a vector space model-determined term-document matrix.

$$V = SLZ^{\top} \tag{2}$$

In above Eq. (2), S and Z are the orthogonal matrices. S specifies the topic of the text generated from the original with the most relevant word is put in the matrix's first column. As the extracted notion, Z represents the input matrix column and the transpose matrix is provided as $Z^{\top}$. L is a diagonal matrix with diagonal components representing singular values in decreasing order. The term-document matrix is correlated using LSI by choosing bases with the greatest singular values. If $\hat{L}$ is a matrix with all diagonal members except the N −largest set to zero, the term-document matrix (V) may be approximated as in Eq. (3)

$$\hat{V} = S\hat{L}Z^{\top} \tag{3}$$

The resulting matrix $\hat{V}$ is the best mean square error (MSE) approximation of V. According to LSI, reducing the tiny single values in L reveals a latent semantic pattern exhibited by word usage across texts. The LSA employs a database with an arbitrary mix of latent themes, where every topic is described by a distribution over terms such as physiological terminology frequently utilized by PPD patients. Thus, both the LIWC and LSA methods are used to extract physiological variables linked to women's mental health and behavioural changes from the online twitter tweets and instagram comments.

3.4 Proposed PPD prediction model

For the final prediction result, the retrieved features from LIWC and LSA are input into the OPOMLP model.

3.4.1 MLP structure

MLP is a back-propagation-based artificial neural network (ANN) that is extensively used in supervised learning for data categorization and prediction. MLP is made up of numerous layers of neurons organized in a directed graph, with each layer completely linked to the next. A directed graph connects every layer of neurons in an MLP, creating an organized network of connections between them. Weights and polarization units are often modified throughout training. It should be emphasized that, with the exemption of the input nodes Q, each node in the network is a neuron with a complex activation function, as shown in Eq. (4).

$$H_{1y} = f\left(\sum_{x=1}^{n} W_{xy}\, Q_x + B_y\right) \tag{4}$$

Loss In Eq. (4), H_{1y} denotes the node y of the hidden layer H_1, W_{xy} is the input gate of the H_1 and B_y depicts the bias. functions play an important role in MLP network training since they represent feature vectors and are assessed based on the structure's ability to model them. MLPs do not need vast volumes of data, making them appropriate for a wide range of applications. Since the MLP parameters were not adequately tuned, this results in computational difficulties in the PPD prediction. The OOA is employed in this prediction model to fine-tune the parameters of MLP.

3.4.2 Osprey optimization for hyperparameter tuning of MLP

The capacity of the nocturnal bird of prey known as the osprey to grab fish and move them to a more suitable position might serve as inspiration for a novel optimization technique. The OPOMLP mathematical modelling is detailed further below.

A. Initialization

The suggested OOA is a population-based strategy that finds an appropriate solution based on the exploration capacity of its population members via a repetition-based procedure. Based on its location in the search space, each osprey in the OOA population determines issue parameters, allowing every osprey a possible solution. Every osprey in the OOA population is a possible solution because it identifies decisions about the problem's variables based on its position in the search space. First, a matrix is used to represent the OOA population as follows in Eq. (5).

$$G = \begin{bmatrix} G_1 \\ \vdots \\ G \\ \vdots \\ G_n \end{bmatrix}_{n\times p} = \begin{bmatrix} G_{1,1} & \cdots & G_{1,b} & \cdots & G_{1,p} \\ \vdots & \ddots & \vdots & \iddots & \vdots \\ G_{a,1} & \cdots & G_{a,b} & \cdots & G_{a,p} \\ \vdots & \iddots & \vdots & \ddots & \vdots \\ G_{n,1} & \cdots & G_{n,b} & \cdots & G_{n,p} \end{bmatrix}_{n\times p} \quad (5)$$

Then, the location of ospreys in the search space is initialized randomly by Eq. (6).

$$g_{a,b} = \mathcal{L}_b + r_{ab} \cdot (\mathcal{U}_b - \mathcal{L}_b), a = 1, \dots, n;\ b = 1, \dots, p \quad (6)$$

In Eqs. (5) and (6), G denotes the population matrix of osprey's positions, G_a indicates the a^{th} osprey (i.e., a candidate solution), $g_{a,b}$ indicates its b^{th} size (i.e., problem variable), n represents to the amount of ospreys, p denotes the numeral of problem variables (i.e., number of hyperparameters of the MLP model), $r_{a,b}$ is the random value ranging [0,1], $\mathcal{L}_b$ and $\mathcal{U}_b$ represents the lower and upper bound of the b^{th} problem variable, respectively. The objective (fitness) function of a problem such as prediction accuracy is computed by comparing each osprey as a candidate solution and the evaluated values can be depicted using a vector as:

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_a \\ \vdots \\ F_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} F(G_1) \\ \vdots \\ F(G_a) \\ \vdots \\ F(G_n) \end{bmatrix}_{n\times 1} \quad (7)$$

In Eq. (7), F is the vector of fitness function values and F_a means the derived fitness function value for the a^{th} osprey. The calculated values of the objective function are critical for obtaining the feasibility of proposed solutions. The finest value represents the best potential solution, while the worst value represents the poor solution. Updating the ospreys' location in the search space necessitates in changing the best candidate solution throughout each iteration of the search process.

B. Exploration - Searching Location and Hunting Prey

Because of their keen vision, ospreys are able to see fish underwater and attack them. This natural behavior serves as the basis for modeling the first phase of the OOA population update. The osprey's location in the search space is significantly altered by this simulation, improving OOA's search ability in locating ideal regions and avoiding local optima. Underwater fishes with higher fitness function values for each osprey's location in the exploration space are taken into account by the OOA design. Each osprey's group of fish is illustrated by

$$T_a = \{G_u | u \in \{1, \dots, n\} \wedge F_u < F_i\} \cup \{G_{best}\} \quad (8)$$

The collection of fish locations for the a^{th} osprey is denoted by T_a in Eq. (8) and the best osprey or optimum candidate solution is denoted by G_{best}. One of these fish is randomly located by the osprey which then strikes it. An osprey's new position is ascertained by tracking its movement in relation to the fish, which is specifically defined as

$$g_{a,b}^{T_1} = g_{a,b} + r_{a,b} \cdot \left(ST_{a,b} - J_{a,b} \cdot g_{a,b}\right) \quad (9)$$

$$g_{a,b}^{T_1} = \begin{cases} g_{a,b}^{T_1}, & \mathcal{L}_b \le g_{a,b}^{T_1} \le \mathcal{U}_b \\ \mathcal{L}_b, & g_{a,b}^{T_1} < \mathcal{L}_b \\ \mathcal{U}_b, & g_{a,b}^{T_1} > \mathcal{U}_b \end{cases} \quad (10)$$

The osprey's previous position is modified by this new location as the fitness function value increases:

$$G_a = \begin{cases} G_a^{T_1}, & F_a^{T_1} < F_a \\ G_a, & else \end{cases} \quad (11)$$

In Eqs. (9), (10) and (11), $G_a^{T_1}$ defines the subsequent location of the a^{th} osprey based on the first phase of OOA, $g_{a,b}^{T_1}$ indicates its b^{th} dimension, $F_a^{T_1}$ indicates its fitness function value, ST_a denotes the selected fish for a^{th} osprey, $ST_{a,b}$ is it b^{th} dimension and $J_{a,b}$ is random value in the range of {1,2}.

C. Exploitation - Moving the Fish to the Safe Location

Once the osprey has caught a fish, it brings it to a secure area so it may be eaten. This natural behavior serves as the basis for the second step of updating the OOA population. As a result, the osprey's location in the search space is somewhat altered, which strengthens the OOA's ability to utilize local search and causes it to converge toward better solutions that are closer to found solutions.

$$g_{a,b}^{T_2} = g_{a,b} + \frac{\mathcal{L}_b + r_{ab} \cdot (\mathcal{U}_b - \mathcal{L}_b)}{e}, a = 1, \dots, n; b = 1, \dots, p; e = 1, \dots, E \quad (12)$$

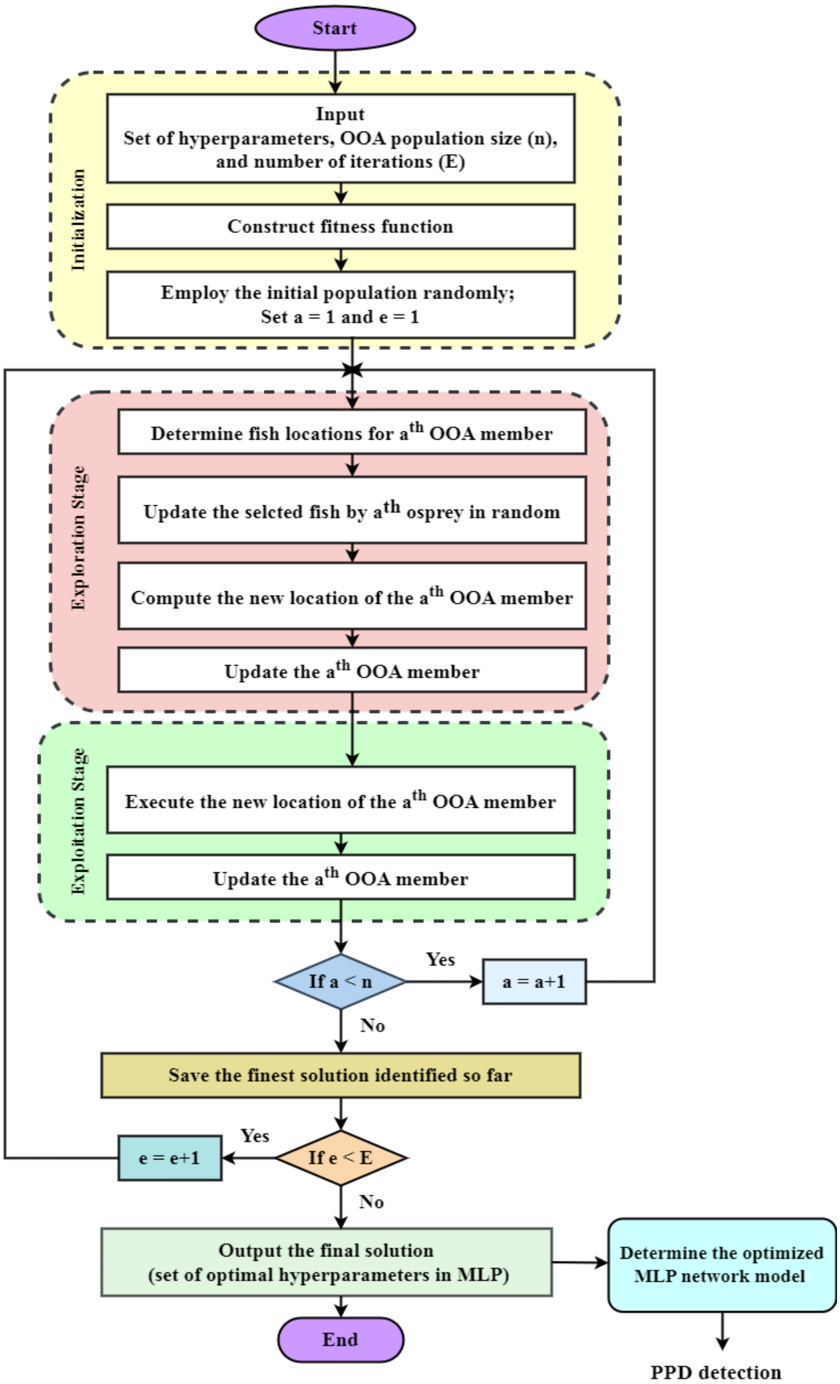

Figure. 2 Block structure of OOA model for MLP optimization

$$g_{a,b}^{T_2} = \begin{cases} g_{a,b}^{T_2}, & \mathcal{L}_b \leq g_{a,b}^{T_2} \leq \mathcal{U}_b \\ \mathcal{L}_b, & g_{a,b}^{T_2} < \mathcal{L}_b \\ \mathcal{U}_b, & g_{a,b}^{T_2} > \mathcal{U}_b \end{cases} \tag{13}$$

$$G_a = \begin{cases} G_a^{T_2}, & F_a^2 < F_a \\ G_a, & else \end{cases} \tag{14}$$

It is modelled by finding a new random place which is safe for each member of the group to eat the fish in Eqs. (12) and (13) Subsequently, the previous position of the related osprey is altered if the fitness function value is increased in this new place as follows in Eq. (14). The new position of the a^{th} osprey corresponding to the second stage of OOA is defined by $g_a^{T_2}$ in Eqs. (12), (13) and (14). $g_{a,b}^{T_2}$ denotes its b^{th} dimension, fitness function value is depicted by $F_a^{T_2}$, the number of iterations is indicated by e and the greatest number of iterations is denoted by E.

Fig. 2 depicts the OOA's general procedure and Algorithm 1 describes the OOA's pseudocode for hyperparameter tweaking. As a result, the OOA is an iteration-based method that adjusts the osprey locations in the first iteration, which compares the values of the fitness function in the second iteration and modifies the optimal candidate solution in the third. In the last iteration, the algorithm changes the locations of the ospreys. The optimal candidate solution (best hyperparameters) captured throughout the iterations is taken into consideration.

Algorithm 1: Hyperparameter tuning using OOA

Input: Set of hyperparameters for the MLP model
Output: Optimal hyperparameters
Begin
//Initialization stage
Initialize the OOA population size n and the total number of iterations E;
Define the fitness function (prediction accuracy)
Create the initial population matrix randomly using Eqs. (5) and (6)
Determine the fitness function using Eq. (7)
$for\ (e = 1{:}E)$
$for(a = 1{:}n)$
//Exploration Stage
Update fish locations for the a^{th} OOA member using Eq. (8)
Determine the chosen fish by the a^{th} osprey randomly
Determine the new location of the a^{th} OOA member using Eq. (9)
Verify the boundary criteria for the new location of OOA members using Eq. (10)
Update the a^{th} OOA member using Eq. (11)
//Exploitation Stage:
Compute the new location of the a^{th} OOA member of OOA using Eq. (12)
Verify the boundary criteria for new location of OOA members using Eq. (13)
Update the a^{th} OOA member using Eq. (8)
$end\ for$
$end\ for$
Return the best solution (i.e., optimal hyperparameters)
End

3.4.3 Model Training

The OPOMLP model for PPD prediction is trained with a set of optimal hyperparameters listed in Table 1. Moreover, the trained model can be applied to accurately predict the women's depression from the tweets and comments shared by them gathered from the online media. Thus the proposed OPOMLP model helps to identify the individuals suffering from PPD with low computational cost and time.

4. Experimental results

This section evaluates the performance of OPOMLP in comparison to other existing models, namely MLP [16], XRT [25], DT [27], LR [29] and XGB [30].

4.1 Experimental setup and performance metrics

The implementation of both proposed and existing model is executed on a system with an Intel® CoreTM i5-4210 CPU @ 3GHz, 4GB RAM and a 1TB HDD running on Windows 10 64-bit which is carried out in Python 3.11 language. From the collected dataset, 1503 instances have been obtained which is divided into 70% for training (1052) and 30% for testing (451). Table 2 presents the parameters and their values utilized for simulating both existing and proposed OPOMLP model to measure performance. The model's ability to predict PPD is assessed using the performance indicators listed below.

- **Accuracy:** It is computed as the proportion of accurately forecasted samples to overall amount of samples.

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{TP + TN + False\ Positive\ (FP) + False\ Negative\ (FN)} \quad (15)$$

Table 2. Experiment parameters for existing and proposed model

Parameters	Search Space	Optimal Range
No. of. Hidden layers	[1, 3, 5, 7]	3
Word embedding size	[128, 256, 520]	256
Learning rate	[0.0001, 0.1]	0.001
Optimizer	[Stochastic gradient descent, Adam, RMSProp Adagrad]	Adam
Dropout rate	[0.1, 0.15, 0.2]	0.15
Weight decay	[0.0001, 0.001]	0.0001
Batch size	[20, 40, 60, 80]	80
Momentum	[0.4, 0.8, 0.12, 0.16]	0.12
Activation function	[linear, ReLU, tan-sigmoid, swish]	ReLU
Loss function	[MSE, Cross-entropy]	MSE

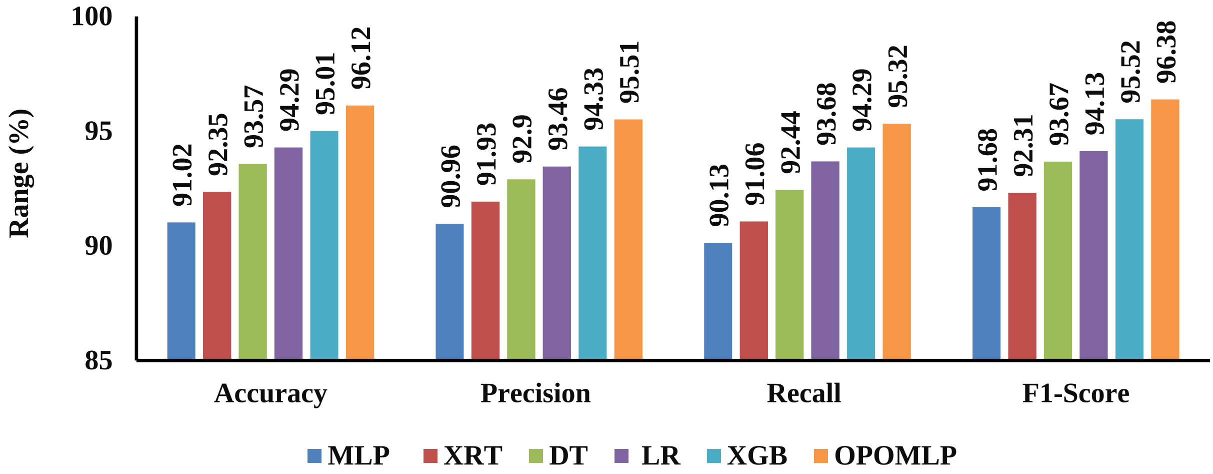

Figure. 3 Performance Comparison of Proposed and Existing for PPD

In Eq. (15), TP refers to the scenario in which the model accurately predicted a positive PPD cases when the actual case was indeed positive. TN alludes to the circumstances in which the classifier properly predicted a negative PPD when the actual class was negative. FP represents a situation in which the model detected a positive PPD case while the actual case was negative. FN denotes a condition in which the model anticipated a negative PPD case when the actual case was positive.

- **Precision:** The proportion of true positive predictions (properly predicted positive PPD instances) out of all positive predictions by the model is calculated in Eq. (16).

$$Precision = \frac{TP}{TP+FP} \quad (16)$$

- **Recall:** It calculates the proportion of genuine positive forecasts out of all positive occurrences in the dataset. It is mentioned in Eq. (17).

$$Recall = \frac{TP}{TP+FN} \quad (17)$$

- **F1-score:** It is the partials average of precision and recall. It is shown in Eq. (18).

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (18)$$

Fig. 3 displays the efficacy of various models on gather PPD dataset for PPD recognition. The OPOMLP model has a higher success rate in accuracy, precision, recall and F1-score when compared to all other previous models.

Accordingly, it is observed that the accuracy of the OPOMLP is 5.6% superior to MLP, 4.08% higher to XRT, 2.73% higher to DT, 1.94% higher to LR and 1.67% higher to XGB, respectively.

The precision of the OPOMLP is 5%, 3.81%, 2.81%, 2.19% and 1.25% higher when compared to the MLP, XRT, DT, LR and XGB models, respectively.

The recall of the OPOMLP model is 5.75%, 4.67%, 3.12%, 1.75% and 1.09% higher than the MLP, XRT, DT, LR and XGB models, respectively.

Similarly, the F-measure of the OPOMLP model is 5.13%, 4.41%, 2.89%, 2.39% and 0.90% higher than the MLP, XRT, DT, LR and XGB respectively.

4.2 Discussion

The existing deep learning models MLP [16], XRT [21], DT [23], LR [25] and XGB [26] can detect the PPD from the dataset, but they do not effectively optimize the parameters resulting in high computational complexity. In contrast, the proposed OPOMLP model specifically addresses this limitation and efficiently learns this information, resulting in high PPD detection ability. As because, the proposed model uses LIWC and LSA to learn complex features from input data, resulting in improved results with a large feature set. It also employs OOA for fine-tuning MLP hyperparameters, reducing computational time and resources while enhancing accuracy for effective global solutions for PPD recognition. From the above evaluation, it is clear that proposed OPOMLP model provides highest performances compared to other existing models. Thus, the proposed model can effectively address the limitations discussed in Section 2 and achieve higher accuracy in detecting PPD in contrasted to other exiting models.

Received: January 31, 2024. Revised: March 19, 2024.

5. Conclusion

In this paper, OPOMLP is proposed for PPD prediction using online social network data. The data is collected through the API function of social media platforms like Twitter and Instagram, Then, an additional psychological attributes like mental health and behavioural changes are extracted using LIWC and LSA methods. Finally, The OPOMLP network is trained using these extracted attributes, along with demographic status, mental health history and vital signs for the final prediction of PPD. The OPOMLP model achieves an accuracy of 96.12% compared to classical PPD detection models with less computational complexities in the prediction of PPD.

6. Main text

Type your main text in 11-point Times New Roman, single-spaced. Do not use double-spacing. All paragraphs should be indented 1.5 times character size. Be sure your text if fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, methodology, software, validation, Suganthi; formal analysis, investigation, Geetha; resources, data curation, writing—original draft preparation, Suganthi; writing—review and editing, Suganthi; visualization, supervision, Geetha.

References

[1] F. Zaidi, A. Nigam, R. Anjum, and R. Agarwalla, "Postpartum depression in women: a risk factor analysis", *Journal of clinical and diagnostic research: JCDR*, Vol. 11, No. 8, pp. QC13-QC16, 2017.

[2] B. F. Hutchens, and J. Kearney, "Risk factors for postpartum depression: an umbrella review", *Journal of midwifery & women's health*, Vol. 65, No. 1, pp. 96-108, 2020.

[3] M. W. O'Hara and J. E. McCabe, "Postpartum depression: current status and future directions", *Annual Review of Clinical Psychology*, Vol. 9, pp. 379-407, 2013.

[4] J. Slomian, G. Honvo, P. Emonts, J. Y. Reginster, and O. Bruyère, "Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes", *Women's Health*, Vol. 15, pp. 1-15, 2019.

[5] T. Beck and P. Indman, "The many faces of postpartum depression", *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, Vol. 34, No. 5, pp. 569-576, 2005.

[6] Z. Wang, J. Liu, H. Shuai, Z. Cai, X. Fu, Y. Liu, and B. X. Yang, "Mapping global prevalence of depression among postpartum women", *Translational psychiatry*, Vol. 11, No. 1, pp. 1-24, 2021.

[7] C. L. Dennis and E. D. Hodnett, "Psychosocial and psychological interventions for treating postpartum depression", *Cochrane database of systematic reviews*, No. 4, pp. 1-50, 2007.

[8] N. Valverde, E. Mollejo, L. Legarra, and M. Gómez-Gutiérrez, "Psychodynamic Psychotherapy for Postpartum Depression: A Systematic Review", *Maternal and Child Health Journal*, Vol. 27, No. 7, pp. 1-9, 2023.

[9] S. Wang, J. Pathak, and Y. Zhang, "Using Electronic Health Records and Machine Learning to Predict Postpartum Depression", *Studies in Health Technology and Informatics*, Vol. 264, No. 21, pp. 888-892, 2019.

[10] C. Wakefield and M. G. Frasch, "Identifying patients requiring treatment for depression in the postpartum period from common electronic medical record data available antepartum using machine learning", *medRxiv*, Vol. 11, pp. 1-17, 2022.

[11] . Teissedre and H. Chabrol, "A study of the Edinburgh Postnatal Depression Scale (EPDS) on 859 mothers: detection of mothers at risk for postpartum depression", *L'encephale*, Vol. 30, No. 4, pp. 376-381, 2004.

[12] J. Pitanupong, T. Liabsuetrakul, and A. Vittayanont, "Validation of the Thai Edinburgh Postnatal Depression Scale for screening postpartum depression", *Psychiatry research*, Vol. 149, No. 1-3, pp. 253-259, 2007.

[13] A. Di Florio, K. Putnam, M. Altemus, G. Apter, V. Bergink, J. Bilszta, ... and S. Meltzer-Brody", The impact of education, country, race and ethnicity on the self-report of postpartum depression using the Edinburgh Postnatal Depression Scale", *Psychological medicine*, Vol. 47, No. 5, pp.787-799. 2017.

[14] P. Cellini, A. Pigoni, G. Delvecchio, C. Moltrasio and P. Brambilla, "Machine learning in the prediction of postpartum depression: A review", *Journal of Affective Disorders*, Vol. 309, pp. 350-357, 2022.

[15] K. Saqib, A. F. Khan, and Z. A. Butt, "Machine learning methods for predicting postpartum

depression: scoping review", *JMIR mental health*, Vol. 8, No. 11, pp. 1-14, 2021.

[16] Y. Zhang, S. Wang, A. Hermann, R. Joly, and J. Pathak, "Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women", *Journal of affective disorders*, Vol. 279, No. 1, pp. 1-8, 2021.

[17] P. Kusuma and A. Dinimaharawati, "Four Directed Search Algorithm: A New Optimization Method and Its Hyper Strategy Investigation", *International Journal of Intelligent Engineering and Systems*, Vol. 16, No. 5, pp. 598-611, 2023, doi: 10.22266/ijies2023.1031.51.

[18] P. D. Kusuma and A. Novianty, "Total Interaction Algorithm: A Metaheuristic in Which Each Agent Interacts with All Other Agents", *International Journal of Intelligent Engineering and Systems*, Vol. 16, No. 1, pp. 224-234, 2023, doi: 10.22266/ijies2023.0228.20.

[19] P. Kusuma and A. L. Prasasti, "Walk-Spread Algorithm: A Fast and Superior Stochastic Optimization", *International Journal of Intelligent Engineering and Systems*, Vol. 16, No. 5, pp. 275-288, 2023, doi: 10.22266/ijies2023.1031.24.

[20] P. D. Kusuma and F. C. Hasibuan, "Attack-Leave Optimizer: A New Metaheuristic that Focuses on The Guided Search and Performs Random Search as Alternative", *International Journal of Intelligent Engineering and Systems*, Vol. 16, No. 3, pp. 244-257, 2023, doi: 10.22266/ijies2023.0630.19.

[21] Shin, K. J. Lee, T. Adeluwa, and J. Hur, "Machine learning-based predictive modeling of postpartum depression", *Journal of Clinical Medicine*, Vol. 9, No. 9, pp. 1-14, 2020.

[22] Hochman, B. Feldman, A. Weizman, A. Krivoy, S. Gur, E. Barzilay, ... and G. Lawrence, "Development and validation of a machine learning-based postpartum depression prediction model: A nationwide cohort study", *Depression and anxiety*, Vol. 38, No. 4, pp. 400-411, 2021.

[23] S. Andersson, D. R. Bathula, S. I. Iliadis, M. Walter, and A. Skalkidou, "Predicting women with depressive symptoms postpartum with machine learning methods", *Scientific reports*, Vol. 11, No. 1, pp. 1-15, 2021.

[24] J. F. Raisa, M. S. Kaiser, and M. Mahmud, "A machine learning approach for early detection of postpartum depression in Bangladesh", In: *Proc. of International Conference on Brain Informatics Cham: Springer International Publishing*, pp. 241-252, 2022.

[25] A. Gopalakrishnan, R. Venkataraman, R. Gururajan, X. Zhou, and G. Zhu, "Predicting Women with Postpartum Depression Symptoms Using Machine Learning Techniques", *Mathematics*, Vol. 10, No. 23, pp. 1-26, 2022.

[26] T. H. K. R. Prabhashwaree, and N. M. Wagarachchi, "Predicting Mothers with Postpartum Depression using Machine Learning Approaches", In: *Proc. of 2022 International Research Conference on Smart Computing and Systems Engineering (SCSE), IEEE*, Vol. 5, pp. 28-34, 2022.

[27] S. Rousseau, I. S. Polachek, and T. I. Frenkel, "A machine learning approach to identifying pregnant women's risk for persistent post-traumatic stress following childbirth", *Journal of Affective Disorders*, Vol. 296, No. 1 pp. 136-149. 2022.

[28] Y. Huang, S. Alvernaz, S. J. Kim, P. Maki, Y. Dai, and B. Penalver Bernabe, "Predicting prenatal depression and assessing model bias using machine learning models", *medRxiv*, pp. 1-29, 2023.

[29] C. Wakefield and M. G. Frasch, "Predicting patients requiring treatment for depression in the postpartum period using common electronic medical record data available antepartum", *AJPM Focus*, Vol. 2, No. 2, pp. 1-8 2023.

[30] Liu, A. Dai, Z. Zhou, X. Xu, K. Gao, Q. Li and S. Wang, "An optimization for postpartum depression risk assessment and preventive intervention strategy based machine learning approaches", *Journal of Affective Disorders*, Vol. 328, pp. 163-174. 2023.

[31] https://www.kaggle.com/datasets/parvezalmuqtadir2348/postpartum-depression

[32] https://www.cdc.gov/reproductivehealth/depression/index.htm#Postpartum