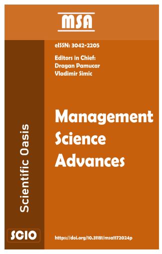
Management Science Advances

Journal homepage: www.msa-journal.org
ISSN: 3042-2205



Machine Learning in the Service of Mental Health Management: Optimization Metaheuristics for Improving Disorder Classification

Lepa Babic[1], Stanislava Kozakijevic[2], Luka Jovanovic[1], Branislav Radomirovic[1], Ivana Strumberger[3], Milos Antonijevic[3], Ninoslava Jankovic[1], Smiljana Tedic[1]

[1] Faculty of Technical Sciences, Singidunum University, Danijelova 32, 11000 Belgrade, Serbia
[2] Faculty Media and Communication, Singidunum University, Karadjordjeva 65, 11000 Belgrade, Serbia
[3] Faculty of Informatics and Computiong, Singidunum University, Danijelova 32, 11000 Belgrade, Serbia

ARTICLE INFO

Article history:
Received 15 February 2025
Received in revised form 17 July 2025
Accepted 5 August 2025
Available online 7 August 2025

Keywords:
Mental Health Management; AdaBoost; Firefly Algorithm; Hyperparameter Tuning; Hybridization.

ABSTRACT

Mental health awareness has become an increasingly pressing societal issue in recent years. Many individuals struggle with undiagnosed disorders, leading to a significantly reduced quality of life. Consequently, timely diagnosis and effective treatment have become essential. Improving mental health care benefits not only individuals but also society as a whole by promoting overall well-being and productivity. However, limited resources and infrastructure often constrain patients' access to mental health professionals. This work seeks to explore the use of advanced machine learning-powered classification algorithms for detecting and identifying mental health disorders with higher accuracy. Since the performance of classification algorithms depends heavily on proper parameter selection, a modified metaheuristic optimization algorithm, based on the firefly algorithm, is introduced to enhance performance and reliability. The proposed approach was evaluated on a publicly available real-world dataset, and a detailed comparative analysis with several contemporary algorithms was conducted. The best-performing models achieved an accuracy exceeding 94%, suggesting the approach's strong viability for real-world assistive applications in mental health care.

*Corresponding author.
E-mail address: lbabic@singidunum.ac.rs

https://doi.org/10.31181/msa21202522

1. Introduction

As the concept of mental health gains limelight in the public eye, the scrutiny it is under gets intensified. The complexity of mental health disorders makes it hard to create strict categories. A disorder might present with a certain set of symptoms in one person and a significantly different set of symptoms in another. This happens due to the influence of personality, past experiences, genetics, as well as social context. A notable example of this issue is the presentation of Attention-deficit hyperactivity disorder (ADHD) in male and female patients. In this case, there is a higher likelihood of diagnosing a patient with ADHD if they are male [1].

The steps taken to tackle this difficulty in diagnosis, besides continuing extensive research of the causes and manifestations of disorders, are varied. Most of the world uses standardized classification procedures suggested by the World health organization (WHO) in the International statistical classification of diseases and related health problems (ICD 10, and 11) [2]. Additionally, patients are assessed and treated by a team, most often including psychiatrists, psychologists, nurses, and context required experts. The initial diagnosis made are also often adjusted after learning more about the patient and their coping capabilities.

Yet, even with these precautions, wrong diagnosis happen in over a third of patients with severe illnesses, studies show [3]. Getting a wrong diagnosis can be a significant setback in therapy choice and execution, while also making it harder for the patient to process and accept their condition. The trust between the therapy-seeker and therapy-provider is worn down by misdiagnosis, thus making the patient potentially less receptive and involved in treatment. This alarmingly high rate alongside the obvious importance of the topic justifies the resuming search for the improvement of the diagnosis procedure.

The development and gaining popularity of artificial intelligence (AI) should come as no surprise when the versatility efficacy of this tool is taken into account. The application of AI allows for quick processing of huge chunks of data, while also being more flexible and capable of adjustment than regular statistical analysis. Thus, the suggestion that it be used as a tool to help mental health professionals should come as no surprise. Of course, diagnosis of any illness comes with great responsibility, and should consequently never be fully dependent on a machine. However, experts can benefit from utilizing the algorithms as tools, similarly to psychological test and biochemical analysis.

In order to gain the most benefit, the AI model must be trained for the specific problem. When it comes to the accuracy of prediction, a key factor are the parameters taken into account when classifying. If the parameters aren't relevant enough to the predict the correct category data belongs to, the model will lose on accuracy. While manual choice of the parameters is possible, it is also time consuming and requires a highly skilled researcher. Using metaheuristic algorithms, however, has become the norm in recent years. These algorithms show great success in hyperparameter tuning, as they can quickly and accurately choose the best fitting parameters for a given model. They are well suited to non-deterministic polynomial-time hard (NP-hard) problems such as mental health diagnosis.

One potential downside of AI models is that, unlike the human brain, the models tend to be very narrowly applicable. That is to say, a model might perform outstandingly on one sort of an assignment, but less than ideal on another. The concept affirming this idea is called the No free lunch theorem (NFL)[4]. Considering the consequences of diagnosis, the search for the optimal model for classification is required. With this in mind, the presenting paper shall offer a model trained on genuine medical data, with a hybridized metaheuristic algorithm specifically suited to the task. In addition, to ensure the relevance of hybridization, this algorithm will be compared with other models using the same base AdaBoos [5] but tuned by varying metaheuristic algorithms already available.

This paper's remaining sections are organized as follows: An overview of earlier studies that inform this research can be found in section 2. The presented approach is covered in Section 3. The experimental setup and results are then displayed in Section 4 and Section 5, respectively. Last but not least, Section 6 offers concluding thoughts on the study, its shortcomings, and prospects for further research in this area.

2. Related works

Disordered emotional functioning is referred to as an affective disorder. Manic behavior is characterized by extremely high, impulsive, energetic, and frequently optimistic (but not always) moods, while depressive behavior is marked with extremely low moods with strong negative feelings, little energy, and an inability to feel positive emotions. ICD-10 differentiates between manic episodes, bipolar affective disorder, depressive episodes, recurrent mood disorders, and chronic mood disorders based on the severity and pattern of symptoms [6]. A history of depressive episodes, ranging in severity from minor to severe, is the primary indicator of recurrent mood disorders, although separate manic episodes are also possible. Bipolar disorder is characterized by recurrent manic and depressive episodes [7].

While overdiagnosis has the potential for higher stigmatization it also makes the resources for therapy more available, with th opposite being true for underdiagnosis [8]. research shows that getting the right diagnosis can, unfortunately, significantly depend on the gender, age, and race of the patient [8, 9]. While the use of AI for research in the field of mental disorders has flourished, it also leaves much to be desired. Namely, the accuracy of models used varies greatly, as the meta-analysis written by Iyortsuun [10] shows. This variance highlights the need for an ongoing search for the best model.

Some notable examples of optimization algorithms include the genetic algorithm (GA) [11] and particle swarm optimization (PSO) [12]. Further algorithms included in the evaluation are nature inspired optimizers such as artificial bee colony (ABC) [13], Harris hawk optimization (HHO) [14] and whale optimization algorithm (WOA) [15]. Finally, more recent optimizer examples include including the reptile search algorithm (RSA) [16] and the COLSHADE [17] algorithms The tuning of a model is best done with metaheuristic algorithms. They are usually molded after the behaviors of creatures in natures, but also by the laws of social and natural science. By modeling these phenomena, the algorithms provide an optimization of the AI models.

As seen in preceding research, the application of AI to current world problems shows promise. Medicine has been a particularly interesting topic, with highly practical results [18–20]. The analysis of sentiment from human-written texts has also benefited [21–23]. Forecasting optimization may assist the economy, according to [24–26].

2.1 AdaBoost

The differences in success rates of different algorithms when applies to different tasks is noticeable, and the contributing factor to the number of algorithms available. AdaBoost was mainly motivated by the prospect of combining inferior learners. Freund and Schapire are credited with developing AdaBoost in 1995 [**ADA**]. An algorithm must categorize only somewhat better than guessing in order to be deemed an inferior learner. With each iteration, the AdaBoost adds new learners while balancing the classifier weights based on accuracy. When performed errors occur, the weight is reduced; when predictions are accurate, the weight is increased.

The error rate of the inferior classifiers is determined per Eq.(1).

$$\epsilon_t = \frac{\sum_{i=1}^{N} w_{i,t} \cdot \mathrm{I}(h_t(x_i) \neq y_i)}{\sum_{i=1}^{N} w_{i,t}} \tag{1}$$

where $w_{i,t}$ gives the weight of the i-th training sample in the t-th iteration, N gives the quantity of training samples, and ϵ_t gives the error weight within iteration t. The true label is y_i, whereas the projected label is $h_t(x_i)$. The function I$(\cdot)$ provides the false and true examples, which are denoted by 0 and 1, accordingly.

The procedure of modifying the weights for recently added inferior learners starts with the establishment of the weights. A sizable student body is required for precise forecasts. A linear model is produced by combining sub-models and their outcomes. Weights in an ensemble are calculated using Eq. (2).

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right) \tag{2}$$

where $alpha_t$ is the weight allocated to each weak learner on the final version of the scale. The weights are updated in accordance with Eq. (3).

$$w_{i,t+1} = w_{i,t} \cdot \exp\left(-\alpha_t \cdot y_i \cdot h_t(x_i)\right) \tag{3}$$

The primary advantages of AdaBoost are the minimization of bias and the decrease of variance through the ensemble technique, which also prevents overfitting. As a result, AdaBoost offers reliable prediction models but has trouble with noisy data.

3. Methodology

The following section presents the original and modified algorithm used for tuning the AdaBoost model.

3.1 *Original firefly algorithm*

The idea, for the Firefly algorithm (FA) as referenced in the study, by [27] was inspired by the way these insects emit light naturally. Fireflies utilize the timing and frequency of their emissions to attract prey and potential mates alike. Both male and female fireflies are captivated by the light emitted as it serves as a measure of appeal. Brighter fireflies are more appealing since they attract others towards them while dimmer ones tend to move to ones. From this observation emerge some principles;

- Unisex light signal: The light signal may attract individuals of both sexes.
- Signal strength and attraction: Signal attractiveness is proportional to light intensity. Between two individuals, the one with the weaker light will move toward the one with the stronger signal.
- Behavior of the brightest individual: The brightest individual moves randomly, as no other individual can attract it.

Optimizing the light's intensity is crucial since it should decrease with distance, in proportion to the landscape, and in reference to the target function. First, n fireflies are positioned at random. A brighter firefly is indicated by a larger objective function, which is used to measure brightness (Eq. 7). The solution for specimen i is indicated here by x_i.

The objective function is assessed to calculate the brightness after the scattering of the agent population, as referred to in Eq. 7.

$$F_i = f(X_i), \tag{4}$$

the X_i stands for the location of the agent, while $f(X_i)$ marks the measure of the objective function.

Following an enticement concept, each agent advances toward the brighter firefly. The following is the calculation for this movement:

$$X_i(t+1) \;=\; X_i(t) \;+\; \beta\, e^{-\gamma\, r_{ij}^2}\, \left(X_j(t) \;-\; X_i(t)\right) \;+\; \alpha\, \epsilon_i(t), \tag{5}$$

β_0 depicts the intensity of attraction while $r = 0$. The Eq. (5) is usually replaced by Eq. (6):

$$\beta(r) = \beta_0 \;/\; \left(1 + \gamma \times r^2\right) \tag{6}$$

The present location of agent j is given by r_{ij} in iteration t, whereas the current position of the firefly is indicated by $X_i(t)$. Based on their distance from one another, β calculates the attraction between i and j. γ represents the light absorption coefficient, α indicates the degree of unpredictability, and $\epsilon_i(t)$ represents the stochastic vector.

Brightness is adjusted according to the fitness function, increasing for higher function values and decreasing for lower ones.

$$I(r) = I_0 \cdot e^{-\gamma \times r^2} \tag{7}$$

3.2 Hybrid FA

The baseline FA optimizer showcases impressive performance and is widely used, known for its remarkable intensification mechanism. However, its fast convergence rate can limit outcomes in creative executions. Specifically, by converging quickly, the FA algorithm can miss promising regions in the search space due to local optima, leading to less favorable overall results. This work seeks to hybridize the baseline FA algorithm with the depletion mechanism from the artificial bee colony (ABC) [13] algorithm to enhance diversification and prevent premature convergence. The modified optimizer is referred to as the hybrid FA (HFA).

To integrate the depletion mechanism into the baseline algorithm, each agent is augmented with an additional parameter, D. The value of D is initially set to 0. However, following an iteration in which the quality of an agent does not improve, the value of D is incremented. Once D exceeds a predetermined threshold, the agent is removed from the population. In its place, a new agent is generated pseudorandomly. The threshold depends on the experimental configuration; for the simulations conducted in this work, it is defined as $\frac{T}{2N}$, where T denotes the maximum number of iterations, and N defines the maximum population size. The pseudocode for the proposed optimizer is presented in algorithm.

4. Simulation configuration

A dataset of genuine information was retrieved from Kaggle [*] in order to evaluate the proposed approach. The dataset contains 17 indicators that are used to distinguish between healthy individuals and those who suffer from affective disorders. Some of the symptoms include the degree of depression, exhaustion, difficulty sleeping, frequent sudden changes in mood, euphoria, contemplation of suicide, eating disorders, anxiety, alexia, nervous breakdowns, avoidance behaviors, mistake recognition, overthinking, aggressive reactions, optimism, sexual behavior, and cognitive focus. These parameters are then categorized and statistically encoded in order to adjust the data to the model analysis. The model is trained using 70% of the data, with the rest of it being used for evaluation.

[*] https://www.kaggle.com/datasets/cid007/mental-disorder-classification

noend 1 Modified HFA optimizer pseudocode

```
Create population P of agents A
Evaluate each A based on objective function
Assign depletion parameter D to each A
while T > t do
    for Each A in P do
        Update A positions in accordance to FA rules
    end for
    for Each A in P do
        Evaluate A quality based on objective function
        if A did not improve then
            Increment D
        end if
        if D of A is > T/2N then
            Replace A with new pseudo random solution
        end if
    end for
end while
return  Best performing A as solution
```

Table 1
Adaboost hyperparameters and ranges.

Adaboost parameter	range
Count of estimators	$[5, 10]$
Depth	$[1, 5]$
Learning rate	$[0.01, 2.00]$

An analysis of the performance of the improved optimizer performs in contrast to other algorithms is conducted. The modified algorithm and the original FA are two of the algorithms used in the analysis. The GA [11] and PSO [12] are also evaluated as well established optimization techniques. Further algorithms included in the evaluation are nature inspired optimizers such as ABC [13], HHO [14] and WOA [15]. Finally, more modern optimizers are also tested including the RSA [16] and COLSHADE [17] algorithms. The parameter settings from the source works are used to initialize each optimizer in its most basic form. Eight iterations are permitted to enhance performance, and population sizes are restricted to eight agents. The goal assigned to optimizers is to choose the best Adaboost parameters from the empirically determined sub-range shown in Table 1.

Models undergo training using a randomly chosen 70% of the available data after parameters are chosen, and they are assessed using an established set of classification metrics [28]. Because the cases in the used dataset are naturally unbalanced, the Cohen's kappa metrics are monitored as the goal. The metric can be calculated using the formula shown in: (8):

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e} \tag{8}$$

where the expected values are indicated by p_e and the observed values by p_o.

5. Simulation outcomes

Simulation scores in terms of the best, worst, mean and median objective evaluations are provided in Table 2. The introduced optimizer manages to attain the most favorable outcomes in the best case execution scenario with an objective function score of 0.925926 outperforming the baseline FA. The RSA, a relatively recent optimizer with a powerful multi-stage optimization scheme showcases favorable outcomes in other cases, including worst and median. The RSA also showcases the highest rate of stability, with low STD and VAR scores.

Table 2
Objective function outcomes across evaluations.

Method	Best	Worst	Mean	Median	Std	Var
AB-HFA	**0.925926**	0.777778	0.840741	0.851852	0.043979	0.001934
AB-FA	0.888889	0.740741	0.818519	0.814815	0.042066	0.001770
AB-GA	0.851852	0.777778	0.811111	0.814815	0.025926	0.000672
AB-PSO	0.888889	0.777778	0.822222	0.814815	0.036289	0.001317
AB-ABC	0.851852	0.777778	0.822222	0.814815	0.022222	0.000494
AB-HHO	0.851852	0.740741	0.796296	0.796296	0.037952	0.001440
AB-WOA	0.814815	0.740741	0.777778	0.777778	0.028689	0.000823
AB-RSA	0.888889	**0.814815**	**0.844444**	0.851852	**0.022222**	**0.000494**
AB-COLSHADE	0.888889	0.740741	0.818519	0.814815	0.038668	0.001495

Objective function outcome distribution and swarm diagrams for each of the algorithms included in the comparative analysis are provided in Figure 1. While the stability of the modified algorithm is somewhat digressed in comparison to the baseline optimization approach, this is to be expected when boosting population diversity, and can be considered a worth lie trade off for performance improvement.

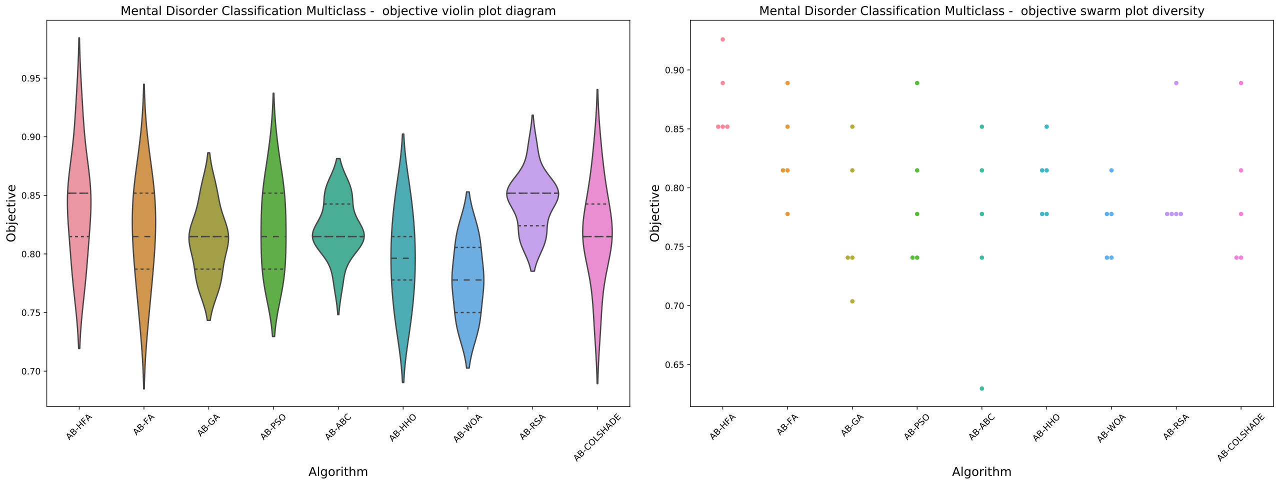


Fig. 1. Objective function distribution and swarm diagrams.

Comparisons in terms of objective convergence rates between each of the optimizers is provided in Figure 2. The improvements in diversification can be observed in the convergences of the proposed algorithm. The baseline FA stalls in a sub-optimal range, the introduced algorithm improves on the attained outcomes attaining the best outcomes in iteration four avoiding a local optimum in favor of a global better solution.

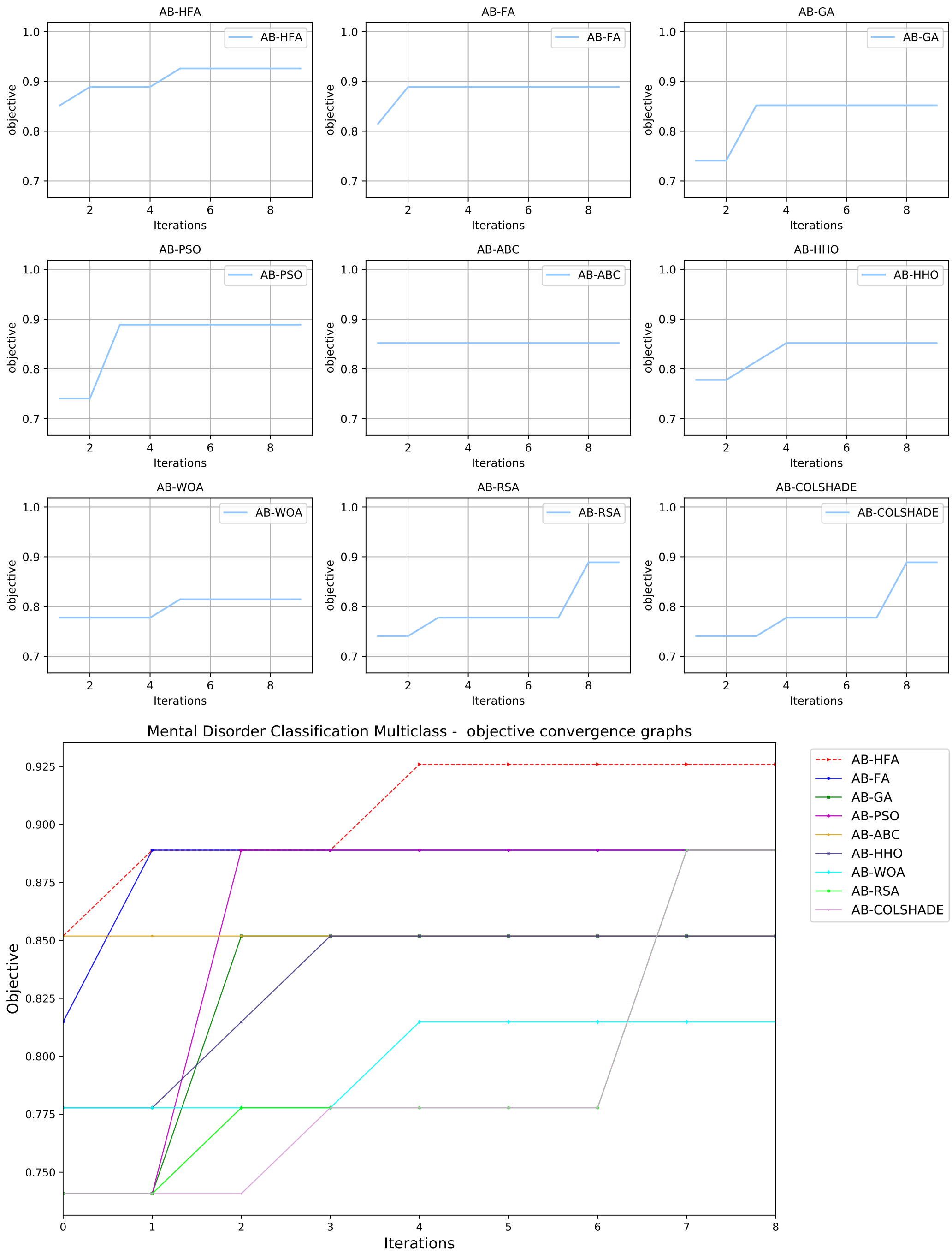


Fig. 2. Objective function convergence diagrams.

Simulation scores in terms of the best, worst, mean and median indicator evaluations are provided in Table 3. The introduced optimizer manages to attain the most favorable outcomes in the best case execution scenario with an objective function score of 0.055556 outperforming the baseline FA. The RSA, a relatively recent optimizer with a powerful multi-stage optimization scheme showcases favorable outcomes in other cases. The RSA also showcases the highest rate of stability, with low STD and VAR scores.

Indicator function outcome distribution and swarm diagrams for each of the algorithms included in the comparative analysis are provided in Figure 3. While the stability of the modified algorithm is somewhat digressed in comparison to the baseline optimization approach, this is to be expected when

Table 3
Indicator function outcomes across evaluations.

Method	Best	Worst	Mean	Median	Std	Var
AB-HFA	**0.055556**	0.166667	0.119444	0.111111	0.032984	0.001088
AB-FA	0.083333	0.194444	0.136111	0.138889	0.031549	0.000995
AB-GA	0.111111	0.166667	0.141667	0.138889	0.019444	0.000378
AB-PSO	0.083333	0.166667	0.133333	0.138889	0.027217	0.000741
AB-ABC	0.111111	0.166667	0.133333	0.138889	0.016667	0.000278
AB-HHO	0.111111	0.194444	0.152778	0.152778	0.028464	0.000810
AB-WOA	0.138889	0.194444	0.166667	0.166667	0.021517	0.000463
AB-RSA	0.083333	**0.138889**	**0.116667**	**0.111111**	**0.016667**	**0.000278**
AB-COLSHADE	0.083333	0.194444	0.136111	0.138889	0.029001	0.000841

boosting population diversity, and can be considered a worth lie trade off for performance improvement.

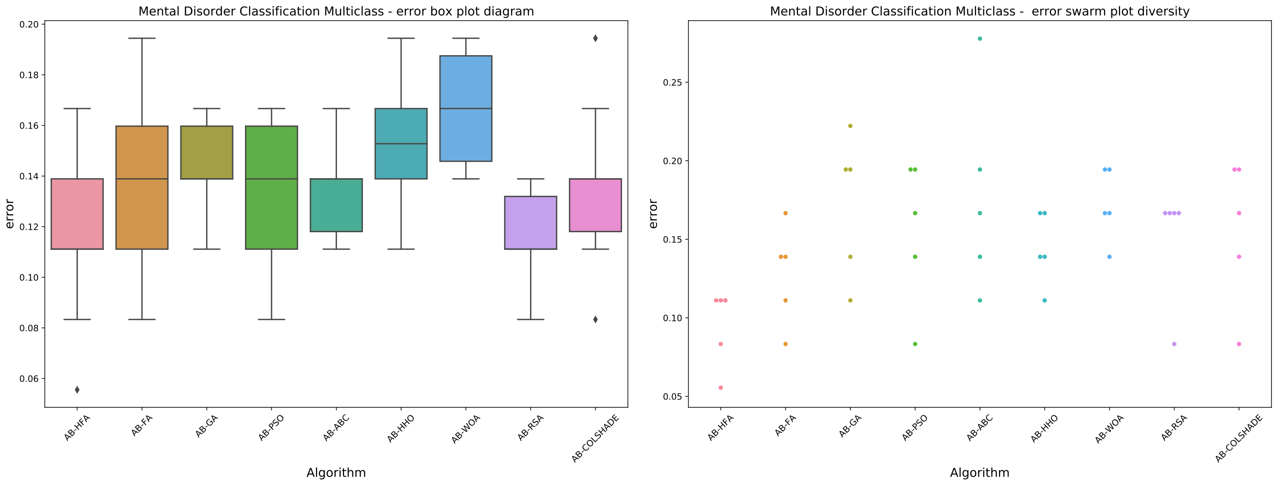


Fig. 3. Indicator function distribution and swarm diagrams.

Comparisons in terms of indicator convergence rates between each of the optimizers is provided in Figure 4. The improvements in diversification can be observed in the convergences of the proposed algorithm. The baseline FA stalls in a sub-optimal range, the introduced algorithm improves on the attained outcomes attaining the best outcomes in iteration four avoiding a local optimum in favor of a global better solution.

Detailed comparisons between the best performing models optimized by each respective algorithm are provided in Table 4. While several algorithms manage to match precision for disorder identification, the highest accuracy, as well as macro and weighted average scores are attained by the model optimized by the introduced HFA optimizer.

Further details for the best performing HFA optimized AB model and it's performance are given in the for of OvR curve in Figure 5. A confusion matrix and PR diagram are also given in Figure 6. Finally, to support experimental repeatability the parameter choices made by each optimizer for the respective best scoring models are provided in Table 5.

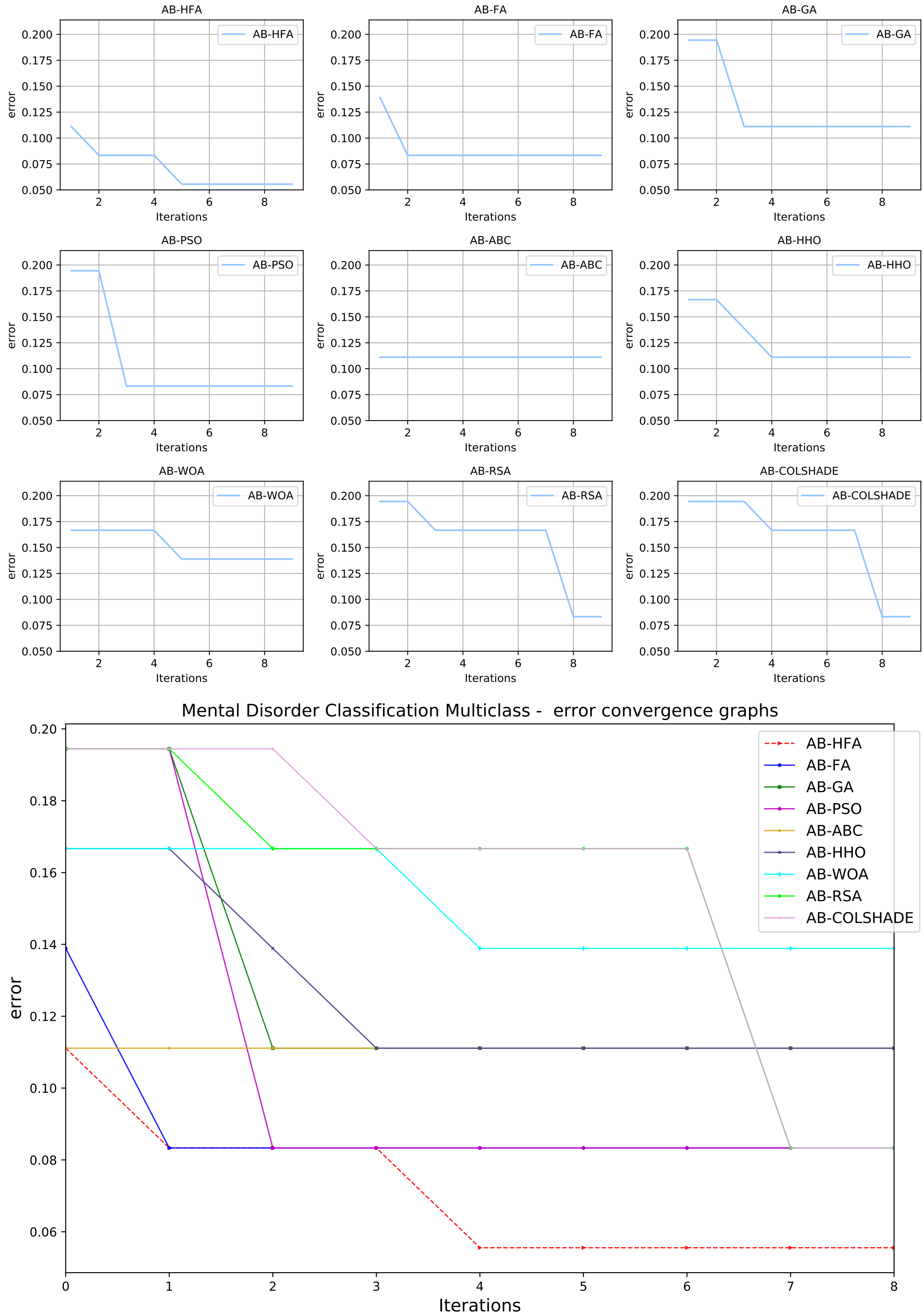


Fig. 4. Indicator function convergence diagrams.

6. Conclusion

Raising awareness about mental health has become a critical societal priority, as undiagnosed disorders continue to diminish the quality of life for many. Early diagnosis and effective treatment are vital not only for individual well-being but also for fostering a healthier and more productive society. However, access to mental health care is often hindered by limited resources and infrastructure. This

Table 4
Detailed metrics for the best performing optimized AdaBoost models.

Method	metric	normal	depression	BPD T1	BPD T2	accuracy	macro avg	weighted avg
AB-HFA	precision	0.888889	0.900000	1.000000	1.000000	**0.944444**	**0.947222**	**0.947222**
	recall	0.888889	1.000000	0.888889	1.000000	**0.944444**	**0.944444**	**0.944444**
	f1-score	0.888889	0.947368	0.941176	1.000000	**0.944444**	**0.944358**	**0.944358**
AB-FA	precision	0.800000	0.888889	1.000000	1.000000	0.916667	0.922222	0.922222
	recall	0.888889	0.888889	0.888889	1.000000	0.916667	0.916667	0.916667
	f1-score	0.842105	0.888889	0.941176	1.000000	0.916667	0.918043	0.918043
AB-GA	precision	0.875000	0.818182	0.888889	1.000000	0.888889	0.895518	0.895518
	recall	0.777778	1.000000	0.888889	0.888889	0.888889	0.888889	0.888889
	f1-score	0.823529	0.900000	0.888889	0.941176	0.888889	0.888399	0.888399
AB-PSO	precision	0.800000	0.888889	1.000000	1.000000	0.916667	0.922222	0.922222
	recall	0.888889	0.888889	0.888889	1.000000	0.916667	0.916667	0.916667
	f1-score	0.842105	0.888889	0.941176	1.000000	0.916667	0.918043	0.918043
AB-ABC	precision	0.800000	1.000000	0.777778	0.888889	0.861111	0.866667	0.866667
	recall	0.888889	0.888889	0.777778	0.888889	0.861111	0.861111	0.861111
	f1-score	0.842105	0.941176	0.777778	0.888889	0.861111	0.862487	0.862487
AB-HHO	precision	0.777778	0.800000	1.000000	1.000000	0.888889	0.894444	0.894444
	recall	0.777778	0.888889	0.888889	1.000000	0.888889	0.888889	0.888889
	f1-score	0.777778	0.842105	0.941176	1.000000	0.888889	0.890265	0.890265
AB-WOA	precision	0.777778	0.800000	0.888889	1.000000	0.861111	0.866667	0.866667
	recall	0.777778	0.888889	0.888889	0.888889	0.861111	0.861111	0.861111
	f1-score	0.777778	0.842105	0.888889	0.941176	0.861111	0.862487	0.862487
AB-RSA	precision	0.888889	0.888889	0.900000	1.000000	0.916667	0.919444	0.919444
	recall	0.888889	0.888889	1.000000	0.888889	0.916667	0.916667	0.916667
	f1-score	0.888889	0.888889	0.947368	0.941176	0.916667	0.916581	0.916581
AB-COLSHADE	precision	0.800000	0.888889	1.000000	1.000000	0.916667	0.922222	0.922222
	recall	0.888889	0.888889	0.888889	1.000000	0.916667	0.916667	0.916667
	f1-score	0.842105	0.888889	0.941176	1.000000	0.916667	0.918043	0.918043
	support	9	9	9	9			

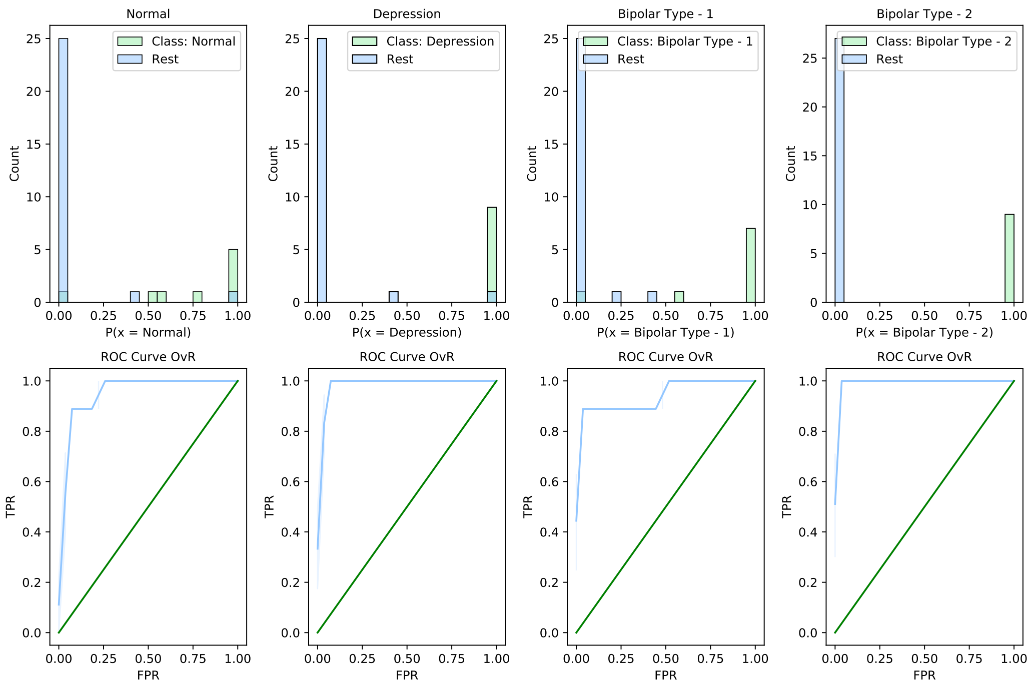


Fig. 5. Best AB-HFA model ROC OvR diagram.

study investigates the potential of advanced ML classification algorithms to improve the accuracy of mental health disorder detection. By introducing a modified metaheuristic optimization method based on the FA, this research enhances algorithm performance and reliability. Using a publicly available dataset, the proposed approach was rigorously evaluated and compared with current methodologies, achieving over 94% accuracy. These results underscore the approach's potential to support real-world

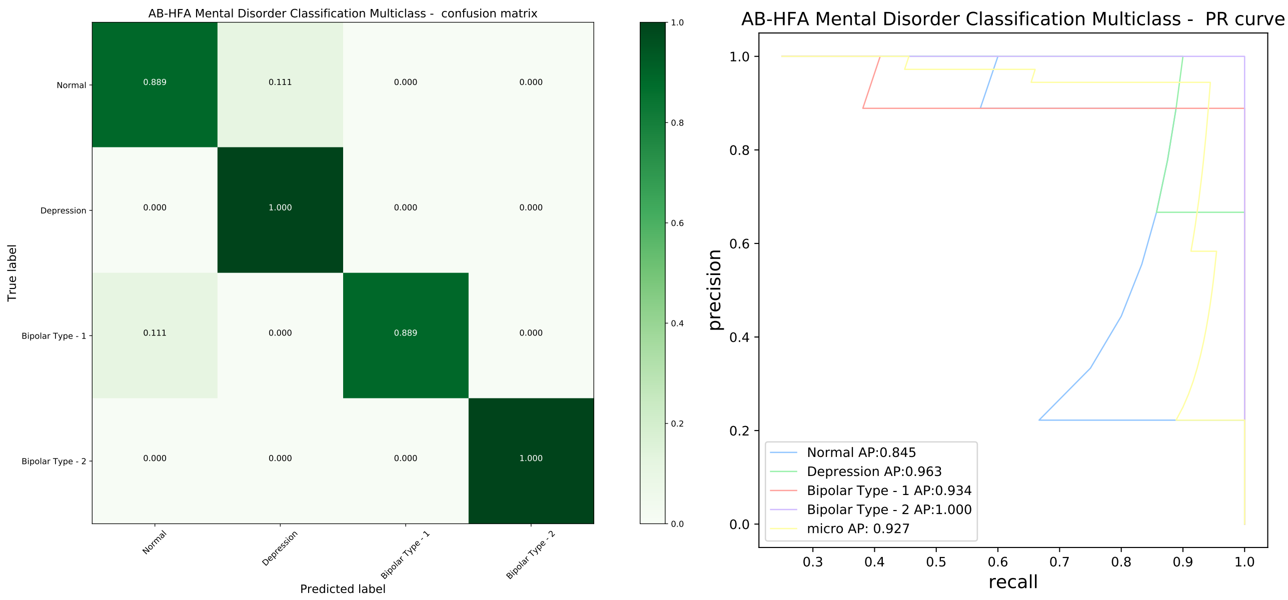


Fig. 6. Best AB-HFA model confusion matrix and PR diagram.

Table 5
Best constructed model parameters selected by each optimizer.

Methods	Number of estimators	Depth	Learning rate
AB-HFA	8	3	1.543823
AB-FA	10	4	2.000000
AB-GA	10	3	1.005382
AB-PSO	9	3	1.640232
AB-ABC	9	3	1.903635
AB-HHO	10	3	2.000000
AB-WOA	10	5	1.296480
AB-RSA	10	3	1.942897
AB-COLSHADE	6	4	1.408946

mental health care applications effectively.

High computing costs of optimization limit the utilized population sizes as well as duration of optimization procedures. These limitations hope to be addressed in future work and additional uses for the proposed optimziation metaheuristic explored in other pressing areas of research.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgement

This research was not funded by any grant.

References

[1] Katrin Bruchmüller, Jürgen Margraf, and Silvia Schneider. "Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis." In: *Journal of consulting and clinical psychology* 80.1 (2012), p. 128.

[2] James E Harrison et al. “ICD-11: an international classification of diseases for the twenty-first century”. In: *BMC medical informatics and decision making* 21 (2021), pp. 1–10.

[3] Getinet Ayano et al. “Misdiagnosis, detection rate, and associated factors of severe psychiatric disorders in specialized psychiatry centers in Ethiopia”. In: *Annals of general psychiatry* 20 (2021), pp. 1–10.

[4] D.H. Wolpert and W.G. Macready. “No free lunch theorems for optimization”. In: *IEEE Transactions on Evolutionary Computation* 1.1 (1997), pp. 67–82. DOI: 10.1109/4235.585893.

[5] Trevor Hastie et al. “Multi-class adaboost”. In: *Statistics and its Interface* 2.3 (2009), pp. 349–360.

[6] World Health Organization. *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research*. Vol. 2. World Health Organization, 1993.

[7] Yi Hong and Marcia Lei Zeng. “International classification of diseases (ICD)”. In: *KO Knowledge Organization* 49.7 (2023), pp. 496–528.

[8] Linda A Teplin, Karen M Abram, and María José Luna. *Racial and ethnic biases and psychiatric Misdiagnoses: toward more equitable diagnosis and treatment*. 2023.

[9] Howard N Garb. “Race bias and gender bias in the diagnosis of psychological disorders”. In: *Clinical Psychology Review* 90 (2021), p. 102087.

[10] Ngumimi Karen Iyortsuun et al. “A review of machine learning and deep learning approaches on mental health diagnosis”. In: *Healthcare*. Vol. 11. 3. MDPI. 2023, p. 285.

[11] Seyedali Mirjalili. “Genetic Algorithm”. In: *Evolutionary Algorithms and Neural Networks: Theory and Applications*. Cham: Springer International Publishing, 2019, pp. 43–55. ISBN: 978-3-319-93025-1. DOI: 10.1007/978-3-319-93025-1_4. URL: https://doi.org/10.1007/978-3-319-93025-1%5C_4.

[12] J. Kennedy and R. Eberhart. “Particle swarm optimization”. In: *Proceedings of ICNN'95 - International Conference on Neural Networks*. Vol. 4. 1995, 1942–1948 vol.4. DOI: 10.1109/ICNN.1995.488968.

[13] Dervis Karaboga and Bahriye Basturk. “A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm”. In: *Journal of Global Optimization* 39.3 (Nov. 2007), pp. 459–471. ISSN: 1573-2916. DOI: 10.1007/s10898-007-9149-x. URL: https://doi.org/10.1007/s10898-007-9149-x.

[14] Ali Asghar Heidari et al. “Harris hawks optimization: Algorithm and applications”. In: *Future Generation Computer Systems* 97 (2019), pp. 849–872. ISSN: 0167-739X. DOI: https://doi.org/10.1016/j.future.2019.02.028. URL: https://www.sciencedirect.com/science/article/pii/S0167739X18313530.

[15] Seyedali Mirjalili and Andrew Lewis. “The Whale Optimization Algorithm”. In: *Advances in Engineering Software* 95 (2016), pp. 51–67. ISSN: 0965-9978. DOI: https://doi.org/10.1016/j.advengsoft.2016.01.008. URL: https://www.sciencedirect.com/science/article/pii/S0965997816300163.

[16] Laith Abualigah et al. “Reptile Search Algorithm (RSA): A nature-inspired meta-heuristic optimizer”. In: *Expert Systems with Applications* 191 (2022), p. 116158. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2021.116158. URL: https://www.sciencedirect.com/science/article/pii/S0957417421014810.

[17] Javier Gurrola-Ramos, Arturo Hernàndez-Aguirre, and Oscar Dalmau-Cedeño. "COLSHADE for Real-World Single-Objective Constrained optimization Problems". In: *2020 IEEE Congress on Evolutionary Computation (CEC)*. 2020, pp. 1–8. DOI: 10.1109/CEC48606.2020.9185583.

[18] Ana Toskovic et al. "Anomaly Detection in Electroencephalography Readings Using Long Short-Term Memory Tuned by Modified Metaheuristic". In: *International Conference on Sustainable and Innovative Solutions for Current Challenges in Engineering & Technology*. Springer. 2023, pp. 133–148.

[19] Katarina Kumpf et al. "A Two Layer Hybrid Approach for Parkinson's Disease Detection Optimized via Modified Metaheuristic Algorithm". In: *International Conference on Innovations and Advances in Cognitive Systems*. Springer. 2024, pp. 205–219.

[20] Stanislava Kozakijevic et al. "Seizure detection via time series classification using modified metaheuristic optimized recurrent networks". In: *Theoretical and Applied Computational Intelligence* 1.1 (2023), pp. 82–94.

[21] Stanislava Kozakijevic et al. "Consumer Feedback Sentiment Classification Improved Via Modified Metaheuristic Optimization Natural Language Processing". In: *International Journal of Robotics and Automation Technology* 11 (2024), pp. 81–95.

[22] Nebojsa Bacanin et al. "Leveraging metaheuristic optimization to enhance insider threat detection through email content natural language processing". In: *International Conference on Intelligent and Fuzzy Systems*. Springer. 2024, pp. 569–577.

[23] Stanislava Kozakijevic et al. "Machine Learning for Company Review Sentiment Analysis Interpretation". In: *International Conference on Multi-Strategy Learning Environment*. Springer. 2024, pp. 647–659.

[24] Marko Stankovic et al. "Forecasting ethereum price by tuned long short-term memory model". In: *2022 30th Telecommunications Forum (TELFOR)*. IEEE. 2022, pp. 1–4.

[25] Robertas Damaševičius et al. "Decomposition aided attention-based recurrent neural networks for multistep ahead time-series forecasting of renewable power generation". In: *PeerJ Computer Science* 10 (2024).

[26] Vule Mizdrakovic et al. "Forecasting bitcoin: Decomposition aided long short-term memory based time series modelling and its explanation with shapley values". In: *Knowledge-Based Systems* (2024), p. 112026.

[27] Xin-She Yang and Xingshi He. "Firefly algorithm: recent advances and applications". In: *International journal of swarm intelligence* 1.1 (2013), pp. 36–50.

[28] Mohammad Hossin and Md Nasir Sulaiman. "A review on evaluation metrics for data classification evaluations". In: *International journal of data mining & knowledge management process* 5.2 (2015), p. 1.