

Enhancing Feature Selection Accuracy using Butterfly and Lion Optimization Algorithm with Specific Reference to Psychiatric Disorder Detection & Diagnosis

Chandrabhan Singh¹, Dr. Mohit Gangwar², Dr. Upendra Kumar³

¹Research Scholar, Computer Science & Engineering
Dr. A.P.J. Abdul Kalam Technical University
Lucknow, UP, India.

²Professor, Department of Computer Science & Engineering
BNCET, Lucknow, UP, India.
Email: mohitgangwar@gmail.com

³Assistant Professor, Department of Computer Science & Engineering
IET, Lucknow, UP, India.
Email: upendra.ietlko@gmail.com

Corresponding Author: Dr. Mohit Gangwar (mohitgangwar@gmail.com)

Abstract— As the complexity of medical computing increases the use of intelligent methods based on methods of soft computing also increases. During current decade this intelligent computing involves various meta-heuristic algorithms for Optimization. Many new meta-heuristic algorithms are proposed in last few years. The dimension of this data has also wide. Feature selection processes play an important role in these types of wide data. In intelligent computation feature selection is important phase after the pre-processing phase. The success of any model depends on how better optimization algorithms is used. Sometime single optimization algorithms are not enough in order to produce better result. In this paper meta-heuristic algorithm like butterfly optimization algorithm and enhanced lion optimization algorithm are used to show better accuracy in feature selection. The study focuses on nature based integrated meta-heuristic algorithm like Butterfly Optimization and lion-based optimization. Also, in this paper various other Optimization algorithms are analyzed. The study shows how integrated methods are useful to enhance the accuracy of any computing model to solve Complex problems. Here experimental result has shown by proposing and hybrid model for two major psychiatric disorders one is known as autism spectrum and second one is Parkinson's disease.

Keywords- Autism Spectrum Disorder (ASD), Parkinson's disorder (PD), Butterfly optimization algorithm (BOA), Enhanced Lion optimization algorithm (ELOA), Modified Adaptive Neuro-Fuzzy System (MANFIS)

I. INTRODUCTION

Computational intelligence becomes the core of artificial intelligence. Various research related to intelligent computing is performed to solve complex problems. Medical domain is a complex domain where decision making needs lot of experience, test an analysis. Correct diagnosis is the base of perfect treatment. Medical data is complex and analysis it is not easy task. For better diagnosis classifications is important phase of analysis. Detection and diagnosis process of mental health illness is very complex in nature. Health care data contains many dimensions the digenesis process for such mental illness contains multiple stages. Due to the complexity and multi dimensionality nature of such sector intelligent computing method is needed. Various Machine learning based approach

for diagnosis of psychiatric disorder has been presented by researchers [1,2]. The research focus of two major Psychiatric created disorder where one is Autism spectrum disorder and second is Parkinson's disorder. Various previous research using data mining was carry doubt to deal with mental health [3,4,5]. Due to change in behavior the symptoms change in mental disorder and so it needs intelligent mechanism to classify it [6,7].

Research carried out based on data mining approaches to check its relationship with unemployability [8]. Artificial neural network (ANN) and support vector machine (SVM) based classifiers are used by the researcher to detect Parkinson's disease [9]. In another research where, higher accuracy has been obtained using SUM for the same disease to have early

detection [10]. To identify ASD and its diagnosis [11] the author has used multi-dimensional scaling (MDS) along with PCA. Modified grasshopper-based nature optimization algorithm (MGOA) is used by another researcher [12] for ASD with maximum accuracy.

The computing method contain several stages likes preprocessing feature selection classification etc. the initial data may contain noise, inconsistency or it may be incomplete to meet the data quality elements like accuracy, consistency, completeness preprocessing play an important role. In preprocessing stage, we also handled the data set imbalancing to overcome with the difference between minority and majority and class of sample. SMOTE (Synthetic Minority over sampling technique) is used to overcome these class imbalancing.

As we know that is dataset these are many features and if the feature reaches the threshold value it decreases the accuracy of the model. Because with that data there may be chances that model get confused and produce in accurate result. So, from all set of feature only good subset is selected and then based on these feature learning is performed. The traditional optimization techniques are easy to implement in comparison with Meta-heuristic algorithms, but they are not easy to find global optima Meta heuristic optimization approach is used in may design problem of engineering data clustering or say feature selection. As global optimization become a new way deal with feature selection. The benefit of opposition-based learning is applied on Meta heuristic algorithm. In this paper the benefit of OBL is shown on butterfly optimization algorithms and enhanced optimization algorithms. In this research integrated Neuro-fuzzy with hybrid optimization approach is illustrated. This composite model of optimization is based on Butterfly and lion-based nature inspired optimization algorithm. The purpose of this research is to show how effectively, and hybrid model of intelligent computing can be used for diagnosis of psychiatric disorders.

II. MOTIVATION

As we know that optimization is nothing but a point where we find the best and most desirable solution to apply any optimization algorithms first we apply some constraints that is a condition which must be satisfied by the solution. Optimization methods have two main types one is exact and second one is approximate method. Meta heuristic algorithm comes under approximate methods of optimization. Single solution based, and population based is two main type of heuristic algorithm.

The design of Meta heuristic algorithm depends on two criteria the first one is where we explore the search space and in second, we find the best solution. Sometime single algorithm does not find the best solution in complex problems so two are

more Meta heuristic approaches of optimization is combined to find the best solution.

III. WORKING OF META HEURISTIC METHODS

As optimization is performed randomly. It means that we start with a random solution after it we compute the fitness number or say value of all solution.

The main role of any Meta-heuristic algorithm is in combing, moving the initial population in iterative manner as defined. In the process of Meta-heuristic method, firstly initial population is generated randomly and then afterwards fitness values are calculated for each solution and later on the solutions are updated on basics of fitness value and this process it repeated until best solution is found. Figure 1 illustrates the working of Meta-heuristic algorithm-based optimization.

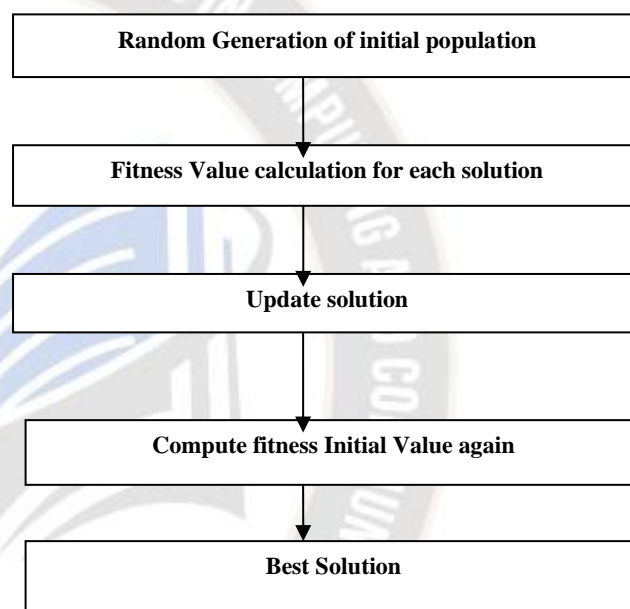


Figure 1: Stages in Optimization Algorithms

IV. LITERATURE Review

Feature selection (FS) is one of the responsible approaches which is deployed to construct an accurate model for prediction of problem related to search space. Searching methods might be concise by using heuristic search, probabilistic search, and spontaneous mixed enhancement algorithms. Meta-heuristic procedures strategy is quicker, time devastating and solitary search a specific pathway to find the best required solution [13]. Like in the fields of pattern recognition, machine learning, and data mining, feature selection is an active topic of research. The goal of feature selection is to locate the most pertinent features that will improve classification accuracy while reducing irrelevant data. It has been demonstrated to be effective in raising learning effectiveness, raising prediction accuracy, and decreasing complexity of learnt results, both in theory and in practice.

The process of displaying two unpredictable parameters are enhancing the space of searching as well as finding the most required outcome by using meta-heuristic algorithms. In such scenario it is not compulsory that the algorithms based on meta-heuristic approach are able to optimize the attributes of applied input data in each circumstance. Whenever the alterations or augmentations are possible to implement a new scope of meta-heuristic approaches for enhancing and manipulating the space of searching. This inducement motivates highest level of our efforts to generate a prototype of predictive model based on combination of tactics for solving subsections of complications, by dropping down the quantity of unrequired attributes, weak related and unrelated features. In practical a sub sequence of strongly related features is considered responsible for producing the best outcomes for the research. These techniques are used to introduce some enhancement for better classification by making comparisons among different techniques with one another. This is also required to mention the improvements. Various optimization algorithms are also applied in the task of searching the natural food of animals based on their inherited behavior.

In this section the study of various meta-heuristic algorithms of optimization are done. BOA is a nature based meta-heuristic algorithm that works on two main behavior of butterfly. The first one is foraging and the second one is mating. In past one or say two decades many natures inspired meta-heuristic algorithms are designed and developed. Artificial Bee colony is inspired by honey bee [14]. In artificial Bee colony the intelligence of bee is used. There are used three phases to divide the movement of Bee. In the first phase for instance employed phase new solutions are generated, fitness values are calculated then greedy based selection is applied. While in next phase for example onlooker phase based on probability new solutions are generated. Then again fitness is calculated for applying greedy selection. In the last phase i.e. scout phase random solutions are generated and if best one is found previous are replaced.

Cuckoo search algorithm [15] is based on the obligate brood parasitism of cuckoo. In this algorithm three difference operations are defined to optimize the problem. The first phase that is Levy flight phase which is like the natural process of egg generation by cuckoo is treated as generation of new candidate solution. In next phase some nests are replaced with the new solutions. There may be a probability that the cuckoo bird egg is discovered by host bird. Host bird will either build new nest or may throw that egg. Best solution is i.e. egg will move to the new nest.

In firefly algorithm [16] three rules are used. As we know that firefly's unisex in nature i.e. any firefly can be attracted by any other butterfly. Objective function is used in

this algorithm to find the brightness of firefly. Thus, this algorithm is based on movement of firefly which is decided by the brightness value if no value is found then random movement will be performed.

Particle swarm optimization [17] is also a nature-based algorithm which is inspired by the social behavior of the world. Cognitive and social behavior learning of swarm is used to find the best solution.

Butterfly optimization is based on the behavior of the butterflies [18]. As we know that butterfly emit fragrance and the movement is decided by the fragrance. Particle swarm optimization is mainly inspired by animal like fishes or say flock and their coordination whereas BOA is based on foraging nature of butterfly. In Particle swarm optimization all particles are aware with other particles where as in BOA such information sharing is based on the fragrance they emit. In compare to firefly algorithm where the movement is decided by brightness of firefly, in BOA movement is done towards best butterfly based on fragrance. In BOA switching probability is used.

LION optimization also a meta-heuristic algorithm which work on social behavior of lion. Hunting, moving and mating nature of lion is treated as generating and replacing solutions [19]. Grasshopper optimization algorithm [20] is also nature based meta-heuristic algorithm inspired by the behavior (long range and movement) of grasshopper. This is also known as population-based algorithm. Foraging is the behavior which is used in this algorithm and the other behavior of jump are also used to find the best solution.

There are also many other feature selection methods which commonly used by various researchers. In general, many researchers apply traditional methods like gain and chi-square for which has been mentioned in [21,22,23]. In many other findings genetic algorithms such as [24] and ACO and memetic feature optimization process [25,26] are used to remove noisy data from the used input data to enhance the accuracy of feature selection and as well as for better result production.

Nakamura has previously put out a wrapper feature selection strategy based on BA and Optimum Path Forest [27]. With this method, an attribute selection problem was described as a binary oriented optimization problem. Six datasets were utilized in trials to show that the suggested strategy produces statistically significant more compacted sets and, in some circumstances, improves the efficiency of classification.

One of the inspired binary versions feature selection algorithms was the Binary Bat Algorithm (BBA), which claimed to locate the most important feature in a search space [28]. The process of combining the PSO (particle swarm optimization) and SSA (slap swarm optimization) produces a hybridized method considered as (PSOSSA), in this approach the competency of making subsequence of irrelevant and

relevant attributes are improved [29]. SSA is a useful and well adapted optimization algorithm for feature selection applied basically for two problems: one applied for the selection of feature using global optimization [30] and applied to mark statistically the relevant attributes for improving the capability of prediction of classification models [31].

An approach of feature selection to improve the clustering by applying particle swarm optimization in which the main objective is to enhance the present deployment of Bayesian calibration for producing energy simulation [32]. A method for FS to enhance clustering of a document by particle swarm optimization where this approach focuses on enhancements to the current implementation of Bayesian calibration to building power simulation [32].

V. PROPOSED METHODOLOGY

The working of proposed methodology depends on the hybrid nature of two main nature-based optimization algorithm. The BOA and enhanced LOA is used for feature selection in this study and they are applied for two major psychiatric disorders. The working of both algorithms is discussed in following subsections.

A. Butterfly optimization algorithm

This is a nature-based population-based algorithm it focuses on ignoring and mating nature of butterfly. As we know that butterfly uses their ability to smell taste touch hearing to find the food and mating. In this butterfly is nature algorithm for Optimization the role of search agent is played by butterfly like fragrance the fitness value is generated in natural way [18].

- Movement of butterfly: In butterfly optimization algorithm the initial assumption is that each butterfly will produce some fragrance. This fragrance will be the reason to attract other one. The movement of butterfly will be randomly towards the butterfly with more1. fragrance. The movement of butterfly depends on certain stages and based on that stages the steps of algorithm is defined as follows:

- 1). Objective function
- 2). Generation of initial population
- 3). Fitness value calculation
- 4). Define sensor modality, power exponent and switch probability
- 5). Check stopping condition
- 6). For each butterfly calculate fragrance

$$\text{Here Function } f = CI^a \quad (1)$$

Where I is stimulus intensity.
- 7). Then for eat butterfly in the population random moment will be done according to the best solution.

$$X_i^{t+1} = X_i^t + (r^2 * g^x - X_i^t) * f_i \quad (2)$$

Random value r will be in between [0-1].

Here f_i is Butterfly fragrance.

I move from 1 to n where n is total number of butterflies.

B. Lion optimization algorithm

Lion optimization algorithm is also another-Meta heuristic algorithm based on Nature. In this optimization Algorithm first, initial population is generated randomly. In the next step Random selection of Nomad lions are done and remain are partitioned as member of p-prides randomly. In each Pride the ratio of female is defined, and randomly selected female go for hunting and in remaining resident mating is performed [19].

C. BOA-ELOA (Enhanced Lion optimization algorithm)

A mixed optimization is useful to reduce the over fitting of classifier and it will improve the convergence. Here in proposed mixed algorithm of nature inspired Optimization opposition learning consent is also used to enhance convergence. The proposed BOA- ELOA algorithm contains following steps:

- 1). Population initialization
- 2). Modality of sensor, probability for switching and stimulus probability power exponent
- 3). Fragrance is calculated for each butterfly in population
- 4). For each butterfly if random generated number is less than switched probability and all will move towards the path of best butterfly accordingly to equation 2. Otherwise female lion update function is performed.
- 5). Value of power exponent is updated, and best solution is found.

VI. WORKING OF PROPOSED MODEL FOR DIAGNOSIS OF PSYCHIATRIC DISORDERS

This study focuses on two major psychiatric disorders, Parkinson's disease and Autism Spectrum disorder. Data set is taken from UCI Data repository [33, 34]. The stages of proposed model are as follows:

- 1). Preprocessing is done on dataset to remove noise and reduce imbalancing.
- 2). After preprocessing feature selection process is performed using BOA and LOA as follows:
 - **Butterfly optimization algorithm:**
 - a) Number of search agents.
 - b) Maximum number of iterations.
 - c) Load detail of selected benchmark function.
 - d) Initialize positions of butterfly.
 - e) Check if butterfly go out of the search space then bring it back.
 - f) Calculate fitness of butterfly.

- g) Sort the population of butterfly.
 - h) Update fragrances.
 - i) Update the position of butterfly with respect to its corresponding fragrance.
- **Enhanced Lion Algorithm:**
 - a) Initialize variables to save the position of resident sorted Pride lions.
 - b) Calculate the fitness of initial Pride lions and sort it.
 - c) Calculate fitness of pride lions.
 - d) Select Nomad lions based on their fitness.
 - e) Random walk around selected Pride lion.
 - f) Random walk around resident.
 - g) Update the position of resident.
 - h) Keep resident in population.
- 3). Classification
 - 4). Successful Identification

The above-mentioned steps are involved to find the best result. The initial preprocessing is done to handle the imbalancing in the data. This imbalancing of data need to be handled in situation where we have classes have unequal distribution in dataset. Means if we have to work of dataset where positive class i.e. minority class is less in compared to negative class that is majority class. The result of success of any model specially in medical computing depends on successful analysis of that minority class.

Without handling the class imbalancing we cannot better performance from classifier. The imbalancing also causes in false prediction or say biased model. SMOTE (Synthetic minority oversampling technique) is based on creating synthetic samples of positive class to handle the imbalancing [35]. In this research we have introduced SMOTE with modification by removing the samples that are near or say close to the majority class.

After this the feature selection process is involved. The performance of any predictive or diagnostic system depends on how the features are selected. Here two nature-based optimization methods are introduced in combination of each other. BOA is first introduced and to handle the problem of low convergence ratio the lion optimization algorithm is involved to select the best features. This combined approach will enhance in the accuracy of feature extraction.

In the next stage of classification first fuzzy C mean clustering is done to create the clusters and then Adaptive Neuro-fuzzy inference system is used for classification purpose. The Capability of Neural Network is used for learning purpose and for decision making capability FIS is introduced.

VII. EXPERIMENTAL RESULT

After applying synthetic minority over sampling technique on the data set feature selection is perform. Figure 2 shows the menu bar to import the data set and perform various operations. The main menu has option to load the data set for both Autism and Parkinson.

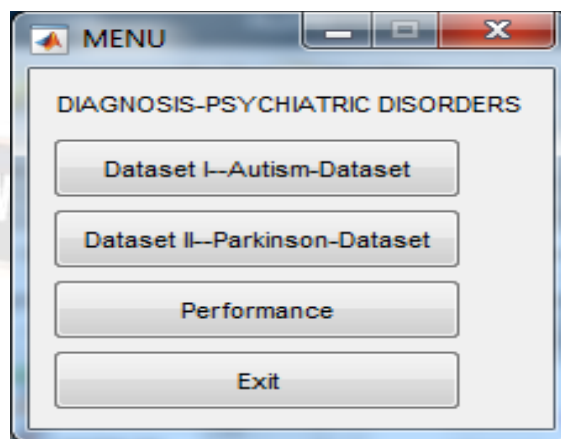


Figure 2: Menu Bar

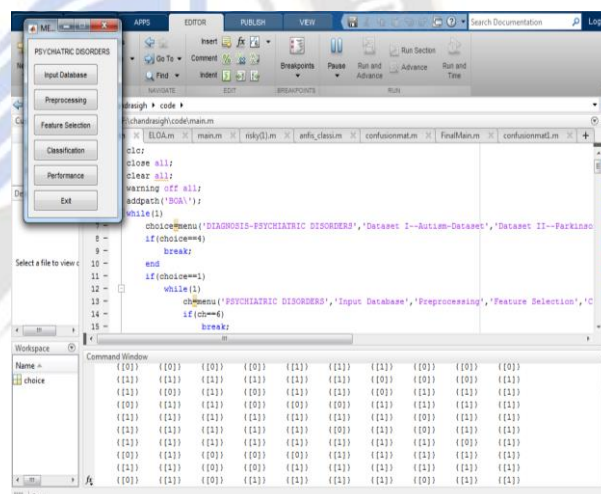


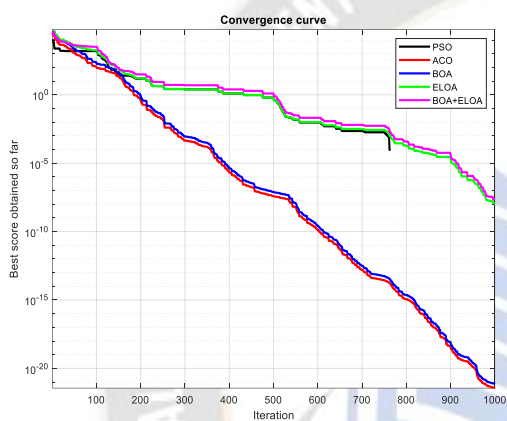
Figure 3: Input Database

Figure 3 shows how input database is loaded. After loading the input data set preprocessing is performed. For preprocessing enhanced SMOTE method is used to deal with class imbalancing.

BOA-ELOA best mixed Optimization method is deployed here to improve the accuracy. BOA [18] is used as explain in previous section and the result of this deployment along with ELOA is shown in figure 4. As you know that BOA has slow as well as low convergence rate and the above research combines LOA [19] with it to improve and enhance the convergence rate. The above integrated nature-based optimization algorithm will lower down the chances of over fitting.

The figure 4 mentioned below clearly shows that the convergence curve of BOA with ELOA with reference to both datasets that is Autism and Parkinson taken from UCI data repository have better rate in comparison to other meta heuristic optimization algorithms such that Particle swarm optimization (PSO), Ant Colony Optimization (ACO), Butterfly Optimization algorithm (BOA) and Lion optimization algorithm (LOA). It clearly indicated that the better the feature selection accuracy will lead to the success of the classification methods. Here the hybrid method is compared to both individuals along with the two others popular meta heuristic algorithms.

Autism Dataset:



Parkinson's Dataset:

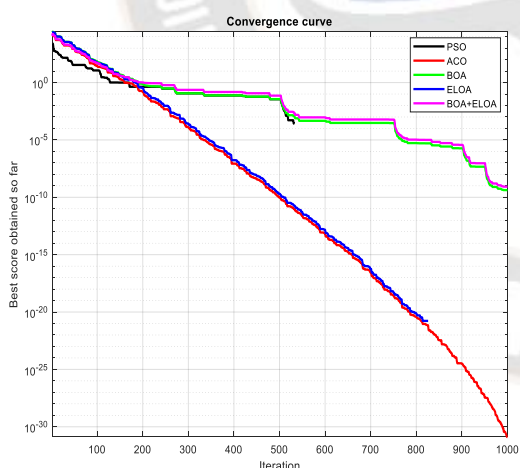


Figure 4: Comparison between Convergence Curve of various algorithms

After this modified and ANFIS algorithm is used for the classification purpose. Based on various statistical measuring approaches the result is compared and table 1 shows the comparison of the result obtained.

TABLE 1: COMPARISON TABLE AMONG VARIOUS PERFORMANCE MEASURES

Autism Dataset:

	MDS	MGOA	M-ANFIS
Accuracy	85.7988	88.7029	92.8082
Precision	84.4166	87.1429	92.1900
Recall	88.6552	91.6149	92.3541
F-Measure	86.4836	89.3229	92.2720
Error	14.2012	11.2971	7.1918

Parkinson's Dataset:

	MDS	MGOA	M-ANFIS
Accuracy	86.3905	89.0756	92.8571
Precision	84.8684	88.0734	92.9046
Recall	89.0862	91.6129	92.8739
F-Measure	86.9262	89.8083	92.8892
Error	13.6095	10.9244	7.1429

The result clearly indicates that there is no requirement to deal with over fitting if we use combination of these two nature-based algorithms. Here the comparison is done based on two major novel researches first one is based on MDS and other one is based on MGOA.

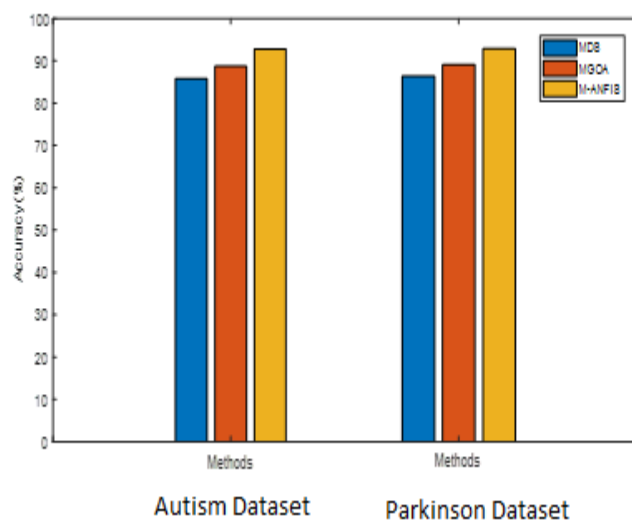


Figure 5: Comparative Analysis on Accuracy

The analysis clearly indicates that based on all statistical measurement parameters the M-ANFIS based approach where feature selection is done based on combined butterfly and lion-based algorithm is better. Figure 5 shows the comparative analysis based on accuracy. The analysis clearly indicates that the proposed approach is better.

VIII. CONCLUSION

As we know that medical data has multidimensionality and for better classification there is need of proper preprocessing and

feature selection. This research focus on combination of two nature-based population algorithms. As we know that perfect identification will help in proper diagnosis. In compare to individual optimization algorithm the integrated methods will provide better result full stop the combined approach over comes the requirement of tuning that reduce the classification over fitting load and results in better accuracy.

The study shows how various population/nature-based algorithms works and what their social behaviors according to which they work are. Then in this study BOA with Enhanced LOA is used for feature selection which handles the problems of overfilling of classifier. This is causing no need for tuning. The experimental result shows how the combined approach is better in compare to other optimization algorithms. Comparative analysis based on two other novel approaches is shown in above section which clearly indicates that the proposed approach is better in all statistical measures.

REFERENCES

- [1]. Le Glaz, A., Haralambous, Y., Kim-Dufor, D. H., Lenca, P., Billot, R., Ryan, T. C., & Lemey, C. (2021). Machine learning and natural language processing in mental health: Systematic review. *Journal of Medical Internet Research*, 23(5), e15708.
- [2]. Scherer, S., Stratou, G., Mahmoud, M., Boberg, J., Gratch, J., Rizzo, A., & Morency, L. P. (2013, April). Automatic behavior descriptors for psychological disorder analysis. In 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG) (pp. 1-8). IEEE.
- [3]. Azar, G., Gloster, C., El-Bathy, N., Yu, S., Neela, R. H., & Alothman, I. (2015, May). Intelligent data mining and machine learning for mental health diagnosis using genetic algorithm. In 2015 IEEE International Conference on Electro/Information Technology (EIT) (pp. 201-206). IEEE.
- [4]. Khan, S. I., Islam, A., Hossen, A., Zahangir, T. I., & Hoque, A. S. M. L. (2018, October). Supporting the Treatment of Mental Diseases using Data Mining. In 2018 International Conference on Innovations in Science, Engineering and Technology (ICISSET) (pp. 339-344). IEEE.
- [5]. Alonso, S. G., de la Torre-Díez, I., Hamrioui, S., López-Coronado, M., Barreno, D. C., Nozaleda, L. M., & Franco, M. (2018). Data mining algorithms and techniques in mental health: a systematic review. *Journal of medical systems*, 42(9), 1-15.
- [6]. Lopez-Castroman, J., Abad-Tortosa, D., Aguilera, A. C., Courtet, P., Barrigón, M. L., Artés, A., & Baca-García, E. (2021). Psychiatric Profiles of eHealth Users Evaluated Using Data Mining Techniques: Cohort Study. *JMIR mental health*, 8(1).
- [7]. Tovar, D., Cornejo, E., Xanthopoulos, P., Guarracino, M. R., & Pardalos, P. M. (2012). Data mining in psychiatric research. In *Psychiatric Disorders* (pp. 593-603). Humana Press.
- [8]. Neto, C., Rodrigues, C., Mendonça, E., Sartori, L., de Pinho, R., Ferreira, D., ... & Machado, J. (2021, February). Data Mining Approach to Understand the Association Between Mental Disorders and Unemployment. In *International Conference on Information Technology & Systems* (pp. 70-79). Springer, Cham.
- [9]. Tahir, N. M., & Manap, H. H. (2012). Parkinson disease gait classification based on machine learning approach. *Journal of Applied Sciences (Faisalabad)*, 12(2), 180-185.
- [10]. Solana-Lavalle, G., Galán-Hernández, J. C., & Rosas-Romero, R. (2020). Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features. *Biocybernetics and Biomedical Engineering*, 40(1), 505-516.
- [11]. El-Ansary, A., Hassan, W. M., Daghestani, M., Al-Ayadhi, L., & Ben Bacha, A. (2020). Preliminary evaluation of a novel nine-biomarker profile for the prediction of autism spectrum disorder. *PLoS One*, 15(1), e0227626.
- [12]. Goel, N., Grover, B., Gupta, D., Khanna, A., & Sharma, M. (2020). Modified grasshopper optimization algorithm for detection of autism spectrum disorder. *Physical Communication*, 41, 101115.
- [13]. Liu, H., amdMotoda, H. (2012) "Feature selection for knowledge discovery and data mining", Springer Science & Business Media, Vol. 454.
- [14]. Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm. *J Glob Optim* 39(3):459–471.
- [15]. Yang X-S, Deb S (2009) Cuckoo search via lévy flights. In: *World congress on nature and biologically inspired computing, NaBIC 2009, IEEE*, pp 210–214.
- [16]. Yang X-S (2010b) Firefly algorithm, levy flights and global optimization. In: *Bramer M, Ellis R, Petridis M (eds) Research and development in intelligent systems XXVI. Springer, Berlin*, pp 209–218.
- [17]. Eberhart RC, Shi Y (2001) Particle swarm optimization: developments, applications and resources. In: *Proceedings of the 2001 congress on evolutionary computation, 2001, vol 1. IEEE*, pp 81–86.
- [18]. Arora S and Singh S 2018, Butterfly optimization algorithm: a novel approach for global optimization, *Soft Computing* 715–734.
- [19]. Boothalingam R 2018, Optimization using lion algorithm: a biological inspiration from lion's social behavior; *Evol. Intel.* 11 31–52.
- [20]. Shahrzad Saremi, Seyedali Mirjalili, Andrew Lewis 2017, Grasshopper Optimisation Algorithm: Theory and application, *Advances in Engineering Software*, Vol 105, Pages 30-47.
- [21]. Y. Sayes, I. Inza, , P. Larranaga, A review of feature selection techniques in bio-informatics, *Bioinformatics* vol. 23, no. 19, pp. 2507 – 2517, (2007)
- [22]. M. Y. Munirah, M. Rozlini, N. Wahid, A comparative analysis on feature selection techniques for medical datasets, *APRN Journal of Engineering and Applied Sciences*, vol 11, no 22, November 2016, (2016)

-
- [23]. C. Shang, M. Li, S. Feng, Q. Jiang, J. Fan, Feature Selection via Maximizing Global Information Gain for Text Classification, *Knowledge-Based Systems*, vol. 54, 298-309, (2013)
- [24]. J. H. Lee, J. R. Anaraki, C. W. Ahn, J. An, Efficient Classification System based on Fuzzy-Rough Feature Selection and Multitree Genetic Programming for intension Pattern Recognition using Brain Signal, *Expert Systems with Applications*, vol. 42, 1644-1651, (2015)
- [25]. S. Kashef, H. Nezamabadi-pour, An Advanced ACO Algorithm for Feature Subset Selection, *Neurocomputing* vol. 147, 271279, (2015).
- [26]. Y. Zhang, D. Gong, Y. Hu, W. Zhang., Feature Selection Algorithm based on Bare Bones Particle Swarm Optimazation, *Neurocomputing*, vol. 148, pp.150-157, (2015)
- [27]. Rodrigues, L. Pereira and R.Y.M Nakamura, "Wrapper approach for feature selection based on bat algorithm and optimum-path forest", *Expert Systems with Applications*, vol.41, pp.2250–2258, (2014)
- [28]. R.Y.M. Nakamura, L. Pereira, M. Acuckoo, K.A Costa, D. Rodrigues, J.P. Papa, X.S. Yang, "BBA: a binary bat algorithm for feature selection". in 25th, SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 22– 25 August, IEEE Publication, pp. 291-297, (2012)
- [29]. Ibrahim, A. and Abd Elaziz, S. "Improved salp swarm algorithm based on particle swarm optimization for feature selection," *Journal of Ambient Intelligence and Humanized Computing*, 2018.
- [30]. Sayed, and Ghada, M. "Anovel chaotic salp swarm algorithm for global optimization and feature selection," Springer, 2018.
- [31]. Mallenahalli and Sarma, "A Tunable Particle Swarm Size Optimization Algorithm for Feature Selection," IEEE, 2018.
- [32]. Abualigah and Khader, H. "A new feature selection method to improve the document clustering using particle swarm optimization algorithm," *Journal of computational science*, 2017.
- [33]. <https://archive.ics.uci.edu/ml/datasets/Parkinson>.
- [34]. <https://archive.ics.uci.edu/ml/datasets/AutisticSpectrumDisorder>.
- [35]. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.