International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Automated Detection of Autism Spectrum Disorder Using Bio-Inspired Swarm Intelligence Based Feature Selection and Classification Techniques

S.Saravana Kumar[1], K.SelvaKumar[2], V.Senthil Murugan[3]
[1]Research Scholar, Department of Information Technology
Faculty of Engineering and Technology, Annamalai University
Chidambaram, India
saravana.coolya@gmail.com
[2]Professor, Department of Information Technology
Faculty of Engineering and Technology, Annamalai University
Chidambaram, India
kskaucse@gmail.com
[3]Department of Networking and communications
Faculty of Engineering and Technology, SRM Institute of science and Technology
Kattankulathur, India
*Corresponding Author: senthilv5@srmist.edu.in

Abstract—Autism spectrum disorders, or ASDs, are neurological conditions that affect humans. ASDs typically come with sensory issues like sensitivity to touch or soundor odour. Though genetics are the main causes, their early discovery and treatments are imperative. In recent years, intelligent diagnosis using MLTs (Machine Learning Techniques) have been developed to support conventional clinical methods in the domain of healthcare. Feature selections from healthcare databases consume nondeterministic polynomial timesand are hard tasks where again MLTs have been of great use. AGWOs (Adaptive Grey Wolf Optimizations) were used in this study to determine most significant features and efficient classification strategies in datasets of ASDs. Initially, pre-processing strategies based on SMOTEs (Synthetic Minority Oversampling Techniques) removed extraneous data from ASD datasets and subsequently AGWOs repeat this procedure to find smallest features with maximum classifications values for recall and accuracy. Finally, KVSMs (Kernel Support Vector Machines) classify instances of ASDs from the input datasets. The experimental results of suggested method are evaluated for classifying ASDs from datasets instances of Toddlers, Children, Adolescents, and Adults in terms of recalls, precisions, F-measures, and classification errors.

Keywords:Autism Spectrum Disorder, Adaptive Grey Wolf Optimization (AGWO), Classification Techniques, Kernel Support Vector Machine (KSVM), ASD Detection.

1. INTRODUCTION

ASDs are neurological developmental disorders that have a significant impact on a person's ability to interact or communicate socially and can appear even in very young kids [1]. They are chronic conditions that do not have treatment options. One research found that 33% of kids with difficulties other than ASDs exhibit some symptoms of ASDs but fall short of complete criteria for diagnosis. ASDs have significant financial effects due to their increased case counts throughout the world as well as the time and expenses associated with diagnostics. Early identification of ASDs can save money in the long run by allowing for the prescription of appropriate therapy and/or medication, which benefits both patients and healthcare professionals [2]. Traditional clinical approaches for detecting ASDs like ADI-R (Autism Diagnostic Interview Revised) and ADOS-R (Autism Diagnostic Observation Schedule Revised) are cumbersome and inconvenient traditional clinical methods as verbal portions of ADI-R cannot be appropriately answered by: patients; young children with delayed speeches scoredon only 25% of the items. Additionally, professional examiners take 90 to 150 minutes to speak with caregivers [3], thus it takes a long time. Moreover, results are summarized with missing values making it less accurate.

On the other hand, identification of ASDs by ADOS-R are based on scoring measurements of provided replies. The propensity to over classify kids who have other clinical diseases is another major issue [4]. As a result, healthcare professionals are in dire need of screening tools for detecting ASDs that can accurately determine if patients have ASDs based on measurable attributes and inform them of the necessity for formal clinical diagnosis. Patients start showing signs of ASDs throughout the first three years of life. Children, however, can develop normally until the age of 18 to 36 months, at which point they may start to show signs. Despite many studies, diagnosing ASDs has proven to be challenging tasks.Other than studying the patient's behaviour and progress, there are no other identifiable symptoms for ASDs [5]. Diagnostic Observationsthrough schedules and Interviews, which comprisedASD patient diagnostics.. However, because there are no identifiable behaviours that can be defined as ASDs, these identification methods take time and can be inaccurate at times [6]. As a result, it is vital to develop methods for diagnosing ASDs that are more reliable and efficient than relying on behavioural patterns.

Numerous brain illnesses have become better understood via the study of neuroimages. Neuroimages are used with MLTs to establish diagnosis procedures and find biomarkers for ASDs. Using MRIs (Magnetic Resonance Images), which can extract data on anatomical and functional activities of the brain are superior methods for investigating neuroimages [7]. rs-fMRIs (Resting state functional MRIs), one of the MRI imaging methods, provides data on brain's neuronal activities. Instead of analysing raw MRIs, graph-theoretic or network-based methods are more efficient for analyses.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

The human brain is divided into many ROIs using network based approaches, which then generate brain networks from MRI data. Although clinicians use standardised diagnostic tools to identify ASDs,the primary drawback lies in diagnostic time required by instruments to complete examinations and analyze findings [8]. Intelligent MLTs for ASD detections have been proposed and primarily aim to speed up diagnosis while improving accuracies of findings and thus helping patients with ASDs by early interventions. Lowering the dimensionalities of relevant inputs, MLTs discover top most ASD traits. Using research datasets, MLTs seek to build superior prediction models and encompass mathematical modelling, artificial intelligences, search algorithms, and other areas of predictions [9]. Standard MLTs include NNs (Neural Networks), DTs (Decision Trees), rule-based classifiers, and SVMs (Support Vector Machines). These are automated devices that process data with little or no human involvement. The main step in the approach for diagnosing ASDs is selecting the appropriate class—ASDs or non-ASDs—based on input features. The method might be viewed as a prediction task that makes use of intelligent MLTs. Hence, automated algorithms or classifiers based on MLTs assessed children with ASDs where input datasets and test cases determined predictions.

Classifications of ASDs have also been solved using supervised MLTs. In supervised learning, goal variables (dependent variables) are predicted using algorithms from inputs where target variables can be continuous or categorical. Many supervised MLTs have been proposed for predicting ASDs where tasks included identifying patterns [10], balancing data, examining multi-dimensional information and choosing required features for competent executions. The assessment metrics used included computing values of accuracies, sensitivities, specificities, ROC curves, and UARs. MLTs can process vast datasets with successful analysis of even complicated datasets containing information on ASDs. The development of knowledge bases, prediction rules and therapeutic responses via indicators and features can also help in learning about ASDspatterns through appropriate MLTs [11] which significantly shorten time taken to diagnose issues, according to earlier studies. The complete feature space might be used by machine learning with adaptive feature selection to explore different subsets with impacts. The Feature selection, on the other hand, is a non-deterministic polynomial time and complex task. AGWOs were used in this study to determine most significant features and efficient classification strategies in datasets of ASDs.

The remainder of the research study is as follows: Section 2 reviews some of the most modern strategies for detecting ASDs. Section 3 describes the proposed detection methodology's process. Section 4 contains the findings and discussion. Section 5 is dedicated to the conclusion and future work.

2. Literature Review

Several existing methods for using classic data mining and soft computing techniques have been presented. The role and effectiveness of various supervised learning and nature-inspired methodologies used for diagnosis of the specified human psychological diseases have been accessed and presented in this part. Furthermore, the publication trends of relevant publications have been examined from several angles. Finally, future avenues for applying these approaches to diagnose psychological problems have been identified.

DLTs (deep learning methods) were proposed by Heinsfeld et al. [12] to detect patients suffering from ASDs based on brain activities recorded in voluminous brain image datasets namely ABIDE global multi-site database (Autism Brain Imaging Data Exchange). ASDs are neurological disorders that result in repeated behaviours and social impairments where one in every 68 kids in the US have ASDs (Disease Control Centre). In an effort to understand the neural patterns that resulted from the categorization, the scientists investigated functional connectivity patterns that can be used to conclusively identify individuals with ASDs using functional brain imaging data. When compared to control of patients, the study's results enhanced recognitions of ASDs in dataset with 70% accuracy. The categorization patternsindicated anti-correlations of brain activities between brain ‘s anterior and posterior parts caused in by ASDs. The study summarized their findings and identified brain areas based on their proposed DLTs which developed controls for diagnosing ASDs. Raj et al. [13] addressed usages of NBs (Nave Bayes), SVMs, LRs (Logistic Regressions), KNNs (K Nearest neighbours), NNs, and CNNs (Convolution Neural Networks) for forecasting and analysing the challenges associated with identifying ASDs in children and adolescents. The study tested their suggested approach on three non-clinical datasets of ASDs all having 21 features. The initial screening for ASDs in children had 292 cases, the second dataset had 704 cases while the third dataset had 104 occurrences of adolescents. Their results strongly suggested that by handling missing values of datasets MLTs performed better as their predictions based on CNNs scored higher accuracy percentages in all three datasets (99.53, 98.3, 96.88). Intelligent diagnosis based on MLTs were suggested by Hossain et al. [14] to enhance routine clinical tests, which can be time-consuming and costly. The study used existing classification techniques to discover most important characteristics and automate diagnostic procedures of ASDs from datasets with information on toddlers, children, adolescents, and adults. Their assessments of recalls, precisions, F-measures, and classification errors evaluated classifiers on the above described binary datasets. SVMs based on SMOs (sequential minimum optimizations) surpassed all other MLTs with better accuracy in benchmarks for diagnosis of ASDs while minimizing classification errors. The study also discovered that Relief Quality methods effectively located important qualities of ASDs in datasets.

For predicting ASDs in people of any age, Omar et al. [15] developed a mobile application and offered a successful prediction model based on MLTs. In this work, RFs (Random Forests), CART (Classification and Regression Trees), and Random Forest-Id3 (Iterative Dichotomiser 3) were merged to construct prediction models for ASDs. Mobile applicationswere also developed based on the study's proposed model and tested using AQ-10 dataset along with 250 real-world datasets of ASDs. The technique outperformed other corresponding methods in terms of accuracies, specificities, sensitivities, precisions, and FPRs (false positive rates) for both types of datasets.MLTs were utilised for ASD categorizations by Thabtah et al [16] who highlighted the advantages and disadvantages of each. The dependability of these measures based DSM-IV manuals in place of DSM-5 manuals was another significant problem with existing screening techniques for ASDs. As a result, present screening methods must be altered, particularly diagnostic algorithms built into them, to match updated classification criteria for ASDs in the DSM-5. The most effective SNP subset may be identified by identifying the most informative SNPs, according to Alzubi et al[17] .'s recommendation. The recommended approach, known as SVM-RFE, combines two approaches: CMIMs (Conditional Mutual Information Maximizations) and RFEs (Recursive Feature Eliminations) based on SVMs. Performance of the suggested approach was evaluated against 3 other methods namelymRMRs (Minimum Redundancy Maximum Relevancies), CMIMs, and ReliefF, using the classifiers, SVMs, NBs, LDAs (Linear Discriminant Analyses), and k-NNsfor ASDs on SNP datasets obtained from the National Center for Biotechnology Information (NCBI) Gene. Their findings showed 89% accuracy in classifying the testdataset and that the adopted feature selection technique was better than other feature selection algorithms. Abdolzadegan et.al., [18] used density-based clustering, artefacts are

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

eliminated and robustness is increased. In addition, metrics like Mutual Information (MI), Information Gain (IG), mRmRs and GAs (Genetic Algorithms) were used to choose features and final conclusions were obtained using KNNs and SVMs. According to their findings, SVMs had a classification accuracy of 90.57% compared to KNNs accuracy of 72.77%. Furthermore, their recommended technique showed sensitivity of 99.91% for SVMs and 91.96% of KNNs. Additionally, their experiments showed that SL, DFA, LE, and Entropy characteristics had significant influences on classification accuracies.

Shi et al. [19] proposed MSTs (minimum spanning trees) for feature selectionsof ASD's neuromarkers. Initially, undirected networks of nodes representing possible characteristics were constructed followed by weight computations which considered feature's redundancy and discriminative ability. MSTs were built from fundamental graph structures using the Prim techniques. Nodes in MSTs and total edge weights of linked nodes were rated. In classifications, N qualities corresponding to nodes with first N least sums were chosen. Finally, SVMs evaluated discriminative performances of feature selections. The findings of the study's comparative experiments revealed that their proposed technique improvedclassifications of ASDs and scored with 86.7% accuracy, 87.5% sensitivity, and 85.7% specificity, respectively. For the purpose of identifying ASDs, Mostafa et al [20] developed an autoencoder-based method. Additionally the study employed as features were the topological centralities of the brain network and the spectrum of the Laplacian matrix. Multiple machine learning algorithms were trained on the recovered features to identify ASDs after the autoencoder extracted discriminant features from the given feature set. The autoencoder was then utilised to pre-train a neural network classifier, yielding a classification accuracy of 79.2 percent. In addition, the autoencoder was pretrained with a neural network to provide a more discriminating representation of the characteristics, yielding a classification accuracy of 74.6 percent. On the whole Autism Brain Imaging Data Exchange 1 (ABIDE 1) dataset, the study's proposed methodologies beat other methods in terms of classification accuracy. Delowar et al[21] used pre-existing classification techniques for improved diagnosis and automate diagnostic processes of ASDs where datasets of toddlers, kids, teens, and adults were input. The study evaluated classification and feature selection methodologies using these datasets. Their experimental findings demonstrated that MLPs (multilayer Perceptrons) performed better than with cent percent accuracy, even with limited samples of toddlers, children, adolescent, and adult datasets.Additionally, they found that all four datasets for ASDs performed well when the "relief F" feature selection technique was used to rank the most important characteristics. A strategy for identifying ASDs was put out by Mohanty et al. The proposed study is based on an analysis of a toddler dataset with imbalanced ASDs from the UCI data repository. Three steps are taken to complete this task. In the initial step of pre-processing, categorical attributes are converted to numeric values via frequency encoding, then numeric attribute standardisation. In the second stage, PCAs (Principal Component Analyses) minimized input dimensions. Last but not least, the data from toddlers with ASDs were categorised using MLTs in two stepsnamely training parameters and k-fold cross validations (k=10). In comparison to other methods, their experimentation results demonstrated good classification performances.

3. Proposed Methodology

This paper presented an AGWOs technique for identifying the most important features and efficient classification techniques in datasets of ASDs. The objective is to identify most relevant characteristics necessary for classification techniques to automate processes and select best performing algorithms. This work uses SMOTEs in its pre-processing while deleting features with missing values as well as those that provide no benefit throughout the analyses. The AGWOs then repeats this procedure to find the smallest feature with the maximum classification recall and accuracy. Finally, KVSMs determined whether an instance in the dataset hadASDs or not.

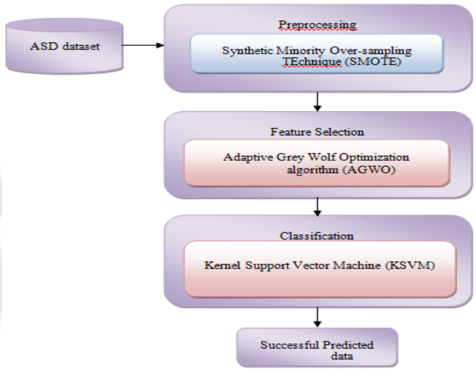

Figure 1. The overall process of the proposed methodology

3.1. Preprocessing

Preparing data for training and analyses with MLTs are frequent initial steps to get the best possible outcomes and extract relevant characteristics [23]. Pre-processing data for computer vision MLTs entails multiple processes, including normalizations of given data. The purpose of these is to remove some of the insignificant identifying traits across the various data sets. To make the borders of each class as clear-cut as possible, the majority of classification algorithms seek to collect pure examples to learn from. The synthetic examples that are farther from the border are simpler to classify than the ones that are closer to it, which present a considerable learning challenge for most classifiers. In light of these results, we provide a novel, sophisticated method for preprocessing unbalanced training sets using SMOTEs, which aims to precisely identify boundaries and produce pure synthetic samples by generalising.

The suggested strategy is split into two sections, which are described below: initial phase, In the first stage, synthetic instances are produced using SMOTEs [24] using the following formula:

$$N = 2 * (r - z) + z \qquad (1)$$

where N *stands for* initial synthetic instance numbers, $r$stands for majority class sample counts, and $z$represents minority class sample counts

In the second step, artificial instances which match majority classes and instances closer to SMOTEsborders are eliminated. Next the supplied data's cleansing, the feature selection process is carried out as explained in the following section.

3.2. Feature Selections based on GWOs (Grey Wolf Optimizations)

The feature selection procedure entails the efficient selection of a subset of variables while avoiding the effect of noise and irrelevant factors on predictive findings [25]. The procedures using filters, wrappers, and embedded techniques scan entire datasets to achieve efficient feature subsets. The correct feature set increases the diagnostic system's performance.

- **Grey Wolf Optimization (GWO)**

This work proposes unique optimizations using AGWOsbased on GWOs inspired by behaviours of grey wolves. Mathematically in GWOs, theirsocial structures are representedas fittest solutions being alpha (α) wolves followed by (β) and delta (δ) wolves as next levels of solutions [26]. It is assumed that the remaining possibilities are omega () wolves. In GWOs, three wolves lead the hunts (optimizations), while others hound the best prey by following them. The following equations were provided to represent grey wolf encircling behaviour during a hunt in addition to social leadership.

Another special characteristic of wolves is their hunting of preys collaboratively where three general rules are followed:

- Tracking, circling, and disrupting targets until they stop moving; and
- Following, chasing, and approaching preys
- Assault on preys

Figure 2 depicts these steps.

Fig. 2 – Hunting styles of Grey wolves: (A) track (B) approach (C) chase (D) pursuit, harassing, encircle (E) stationary situationsbefore attack

3.2.1. Mathematical model of GWOs

This section details on mathematical approaches of wolves social hierarchies and tracing, encircling, attacks on preys.

a) Social hierarchies:

GWOs mimic social hierarchies of wolves by finding fittest solutions as α followed by β, and δ. Additionally Omega ω, and α are also considered as potential solutions.

b) Encircling preys:

Grey wolves, as was previously said, circle their victim when hunting. To represent encircling behaviour analytically, the following equations are presented:

$$\vec{D} = \left|\vec{C}.\vec{X_p}(t) - \vec{X}(t)\right| \qquad (2)$$

$$\vec{X}(t+1) = \vec{X_p}(t) - \vec{A}.\vec{D} \qquad (3)$$

When *t stands for* current iterationcoefficient vectors, $\vec{X_p}$implies positional vectors of preys. Vectors are calculated using:

$$\vec{A} = 2\vec{a}.\vec{r_1} - \vec{a} \qquad (4)$$

$$\vec{C} = 2.\vec{r_2} \qquad (5)$$

Where r_1 and r_2 implyrandom vectors in the interval [0, 1] and get reduced from 2 to 0 linearly across iterations. Figure 3 depicts positional vectors in 2D and probable neighbours as the consequences of equations (2) and (3). Grey wolves locations (X,Y) are updatedcorresponding to prey locations (X*,Y*). Best agents are found at a number of places relative to the present location by altering the values of the vectors and. For instance, setting and will result in (X*-X,Y*). Notably, random vectors allow wolves to go to the areas shown in Fig. 3. Grey wolves use equations (2) and (3) to update their position within the search spaces around prey and random locations (3).

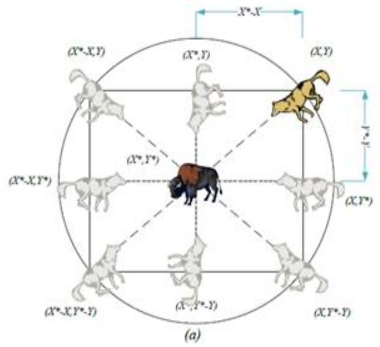

Fig. 3. 2D position vectors and their possible next locations

Search spaces with n dimensions can also exploit the same concepts as grey wolves circle best solutions in hypercubes/spheres).

c) Hunting:

Grey wolves are capable of tracking down and encircling victims with α wolves leading the packs. Alternatively, β, and δ also join the hunting of preys.However, in abstract search spaces, positions of perfect spots (preys) are not clear. Mathematically recreations of GWOs assume alphas as best candidate solutions while betas and deltas are knowledgeable of prospective prey spots. Hence, top three results are only considered and the positions of others including omega wolves are changed to correspond to the positions of α wolves using Equation (6).

$$\vec{D_\alpha} = \left|\vec{C_1}.\vec{X_\alpha} - \vec{X}\right|, \vec{D_\beta} = \left|\vec{C_2}.\vec{X_\beta} - \vec{X}\right|, \vec{D_\delta} = \left|\vec{C_3}.\vec{X_\delta} - \vec{X}\right| \qquad (6)$$

$$\vec{X}(t+1) = \frac{\vec{X_1}+\vec{X_2}+\vec{X_3}}{3} \qquad (7)$$

d) Attacking preys (exploitations):

As previously stated, grey wolves conclude their hunts by attacking targets when they stop moving. The numerical values are reduced to depict the approaching of preys. It should be emphasised that fluctuation ranges are also confined to random numbers between [-a,a] decreasing from two to zero in iterative procedures. The randomly generated values between -1 and 1 imply search agent's future positions can be anywhere between current locations and prey's locations. Figure 4 (a) shows how |A|1 coerces the wolves into attacking the prey.

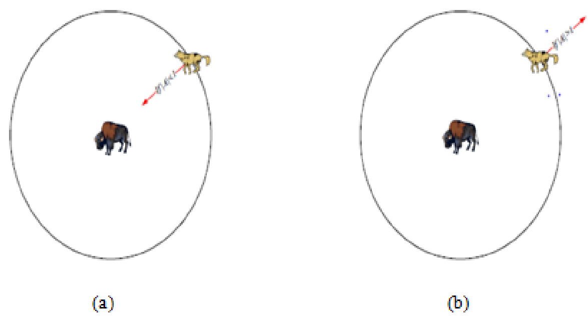

Fig. 4. Attacking prey versus searching for prey

ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

With the suggested operators, GWOs enables its search agents to alter the positions of α, β, δ and ω wolves to launch assaultson preys. With these operators, GWOs are likewise prone to stalling in local solutions. While the suggested encircling strategy does show some exploration, GWO needs additional operators to prioritise exploration.

e) **Prey discovery (explorations):**
Grey wolves generally use locations of α, β, δ in searches (divergence) for preys while attacks are launched in a united manner (convergences). Divergences are theoretically emulated by pushing search agents to diverge from preys using random values in the interval [-1, 1], as this fosters exploration and allows GWOs to explore more broadly. Fig. 4(b) depicts |A|>1 resulting in GWO's divergences in search for fitter preys. The resultant vectors contain randomized values in the interval [0, 2], as shown in Equation (5). Random weights are assigned to preys (C>1) or reduce highlightson preys with (*C<1*) i.e. determining distances in Equation (2). This enables optimizations of GWOs in explorations and avoiding local optimums where the values of A fall linearly unlike that of C. For emphasize explorations not only during first rounds but during final iterations (C produces random values) where they are specifically useful in last levels of iterations and stagnations of local optimums.
C vectors by nature may also be interpreted as effects of obstacles placed in paths to preys as wolf hunting routes have natural impediments which stops them from accessing preys easily. Vector C successfully does this. The prey may be randomly given weight and strengthened making it difficult for wolves to attai9n them.

To summarise, GWOs establish random populations of grey wolves to begin their searches (potential solutions) where α, β, δ wolves predict preysin iterations and option changes as they near preys. To emphasize on explorations and exploitations, values A gradually reduce zero from two. When | A |>1 and | A |1 candidate solutions diverge from prey and tend to converge near the prey. Finally, when end conditions are reached, GWOs get terminated.

3.2.2. Adaptive GWOs

GWOs start with initial randomized populations which are optimized. The three top solutions α, β, δ are found during optimizations. When random values of A are in the range [-1,1] and When $|A|<1$ is reached, α, β, δ converge towards estimated prey positions. omega wolves trigger location changes of search agents. As the iteration progresses, the parameters a and A linearly decrease resulting in divergence of search agents when $|\vec{A}|>1$ and convergence towards preys when $|\vec{A}|<1$. On reaching termination conditions, scores/locations of alpha wolves are considered top results from optimizations. Two additional elements are added to GWOs to enable multi-objective optimizations where the first are repositories for discovered non-dominated Pareto optimal solutions. α, β, δ solutions are selected from the repositories to act as leaders in hunting processes. A straightforward storage unit, the archive allows for the saving or retrieval of previously discovered non-dominated Pareto optimal solutions where primary modules are archive controllers controlling enters or exits in the archive. It should be mentioned that the archive only allows a certain number of members. The archive inhabitants are contrasted with the non-dominated solutions discovered so far during the iteration.

The likelihood of removing solutions grow proportional to counts of solutions in hypercubes (segments). Most packed segments are selected first, and solutions are discarded at random for accommodating new solutions i.e. solutions are put outside hypercubes. In this scenario, all areas have been expanded to accommodate new solutions and as a result, alternate solution segments differ.

The second element is the process for choosing leaders. In GWOs, alpha, beta, and delta wolves are greatest solutions developed which assist other search agents locate solutions that are closer to global optimums and by pointing them towards promising areas of search spaces. However, solutions in multi-objective search spaces cannot be directly compared due to Pareto optimality rules mentioned in the prior paragraph. Throughout leader selections, the best non-dominated solutions ever created are found in the repository. Leader selections are also non-dominated solutions where least crowded search space segments are chosen as alpha, beta, or delta wolves. Through the use of roulette wheels, the required hypercubes are chosen and with the following probability:

$$P_i = \frac{c}{N_i} \qquad (8)$$

where c is a constant greater than one and N is the number of Pareto optimum solutions obtained in the i-th segment

Hypercubes with lesser crowds are more likely to be new leaders based on Eq. (8). The possibilities of choosing hypercubes where leaders are chosen decrease as fewer solutions are located in hypercubes. Certain exclusions occur in the need for choosing three new leaders. If the least-crowded region contains three solutions, the alpha, beta, and delta solutions are each given three of them at random. The second least crowded hypercube is chosen to choose other leaders from if the least crowded hypercube has less than three solutions. The delta leader should be selected from the third least crowded hypercube if the second least crowded hypercube has just one solution.

This technique prevents AGWOs from selecting similar leaders for alpha, beta, or delta. The searches are focussed on on unexplored/unexposed areas of search spaces since leader selection mechanisms prioritize leastly packed hypercubes and distributes leader from other segments when there aren't enough leaders (less than 3) in least congested segments.

Algorithm 1. Pseudocode of proposed adaptive GWO

```
Create a grey wolf population Xi (i = 1, 2,..., n)
Set up a, A, and C.
Determine the objectives for each search agent.
Find the non-dominant solutions and use them to populate the archive.
Xα=SelectLeader(archive)
Remove alpha from the archive
Xβ= SelectLeader(archive)
Remove beta from the archive
Xδ= SelectLeader(archive)
t=1;
while (t < Max number of iterations)
for each search agent
Using Equations (3)-(8) finish updating current search agent's positions.
end for
a, A, and C should be updated.
Determine the objective values of each search agent.
Find non-dominant solutions.
Refresh the archive with the acquired non-dominated solutions.
If the archive is full, use the grid technique to remove one of the present members.
Add the new solution to the archive
end if
If so, add the updated solution to the archive's end.
If any of the new solutions added to the archive are outside the hypercubes, update the
grids to cover the new solution(s) end if
end if
Xα=SelectLeader(archive)
Remove alpha from the archive
Xβ= SelectLeader(archive)
Remove beta from the archive
Xδ= SelectLeader(archive)
Xδ= SelectLeader(archive)
t=t+1
end while
returnarchive
```

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

3.3. Classification using kernel based SVMs

SVMs are Global classification models that frequently incorporate all characteristics and construct non-overlapping divisions. Training and test sets of data examples are frequently used in categorization taskseach search engine. Each training set instance has a single objective value (a class label) and a number of attributes (features). Building a model that can predict target values for data instances in the testing set for which just the attributes are known is the goal of a classifier [27]. It is possible to think of the classification problem as a two-class problem in which the objective is to distinguish between the two classes using a function deduced from the samples given. A classifier that generalises well or performs well in situations that have never been seen before is what is desired. The separating hyperplane that increases the distance (margin) between itself is the best one and the nearest example of each class. It should be anticipated that this classifier will generalize more effectively than the alternatives. SVMs main goal is to choose the hyperplane with the largest margin.

SVMs minimise both the geometric margin and the empirical classification error. The term "maximum margin classifier" also applies to SVMs. When adopting kernel techniques, which implicitly map inputs into high-dimensional feature spaces, SVMs perform well in non-linear classifications. The kernel approaches allow for creation of classifiers without prior knowledge of feature spaces. SVM models represent cases as points in space that are projected into greater space between examples of categories. SVMs, for example, can find a hyperplane with the largest percentage of points from the same class on the same plane when given a collection of points from either of the two classes. These OSHs (Optimal Separating Hyperplanes) are separating hyperplanes that improve the separation between the two parallel hyperplanes and reduce the likelihood of misclassifying test dataset samples. given a set of data points that contain labelled training data

$$M = \{(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)\} \tag{9}$$

where $y_n = 1/-1$, a constant that denotes the class to which that point x_n belongs. Where, n=number of the data sample.

Each x_n is a real vector with p dimensions. Before categorising input vectors, SVMs turn them into decision values using suitable threshold values. The dividing hyperplane for viewing training data is stated as:

$$Mapping: \quad w^T . x + b = 0 \tag{10}$$

where w is a weight vector in p dimensions and b is a scalar . The separating hyperplane and the vector w are perpendicular. We may raise the margin by employing offset option b. We choose hyperplanes without any points in them when the training data can be linearly separated, and we aim to maximise the distance between them. Distances between hyperplanes are calculated as 2/ |w|. To reduce |W|, all must be ensured.

$$w. x_i - b \geq 1 \; or \; w. x_i - b \leq -1 \tag{11}$$

- **Radial Basis Kernel Functions**

Because kernel functions of RBFs (Radial Basis Functions) can analyze higher-dimensional data, RBFs are used as the core of SVMs. The Euclidean distance from the source determines the kernel's output (one of these will be the support vector and the other will be the testing data point). The support vector, which establishes the region of influence this support vector has over the data space, will be at the centre of the RBF. Following is a definition of the RBF's Kernel functions:

$$k(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \gamma > 0 \tag{12}$$

where k is the training vector and is a kernel parameter. A greater value produces a more uniform decision border and a smoother decision surface. This is because larger RBFs enable support vectors to have a significant impacts across larger regions. The optimal parameter set is used to the training dataset to produce the classifier. The suggested classifier approach is utilised to accurately categorize the data from ASDs.

1. Results and Discussion

To evaluate the effectiveness of the proposed method in differentiating various classes of ASDs dataset (DATASET:https://www.kaggle.com/fabdelja/autism-screening-for toddlers?select=Toddler+Autism+dataset+July+2018.csv). In this work, four dataset are used namely, ADULT, CHILD, ADOLESCENT and TODDLER for experimentation.The methods are Conditional Mutual Information Maximization (CMIM), Sequential minimal optimization - Support Vector Machine (SMO-SVM) and Adaptive Grey Wolf Optimization with Support Vector Machine (AGWO-SVM) based model is compared for evaluating the performance of the classifier.

In all trials, the training set was utilised to optimise model parameters, and the validation set was used to tune model and training procedure hyperparameters. Several criteria that are often used in binary classification were used in the trials to evaluate the utility of the different techniques in the prediction of ASDs data. Following the determination of the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) rates, additional performance indicators were created. Precision, which was defined as the percentage of pertinent retrieved instances, was the initial performance metric. The second performance parameter was recall, which is measured as the percentage of retrieved relevant occurrences. Precision and recall matter for assessing a prediction strategy's effectiveness, despite the fact that they frequently contradict one another. As a result, these two metrics may be combined and given equal weights to create the F-measure, a single metric. The percentage of correctly predicted cases to all anticipated occurrences was the final definition of accuracy. The first table contrasts the effectiveness of the suggested and current approaches.

Precision is defined as the ratio of positively discovered observations to all predicted positive observations.

$$Precision = TP/(TP + FP) \tag{13}$$

The ratio of accurately detected positive observations to total observations is described as sensitivity or recall.

$$Recall = TP/(TP + FN) \tag{14}$$

The weighted average of Precision and Recall is defined as the F - measure. As a result, it suffers from false positives and false negatives.

$$F - measure = 2 * (Recall * Precision)/(Recall + Precision) \tag{15}$$

The following is how accuracy is calculated in terms of positives and negatives:

$$Accuracy = (TP + FP)/(TP + TN + FP + FN) \tag{16}$$

Table 1. Performance comparison results between the proposed and existing methods

Dataset	Metrics	CMIM	SMO-SVM	AGWO-SVM
ADULT	Accuracy	91.19	93.24	93.98
	Precision	85.35	89.97	90.87
	Sensitivity	93.24	94.00	95.04

	F-measure	89.02	91.94	92.91
CHILD	Accuracy	92.65	96.55	97.25
	Precision	93.18	96.53	97.23
	Sensitivity	93.14	96.60	97.27
	F-measure	93.01	96.56	97.25
ADOLESCENT	Accuracy	96.41	98	99
	Precision	95.65	97.67	98.80
	Sensitivity	95.67	98.30	99.15
	F-measure	93.34	97.98	98.98
TODDLER	Accuracy	93.01	94.34	95.03
	Precision	90.67	93.19	96.06
	Sensitivity	91.23	93.88	94.56
	F-measure	91.87	93.54	95.30

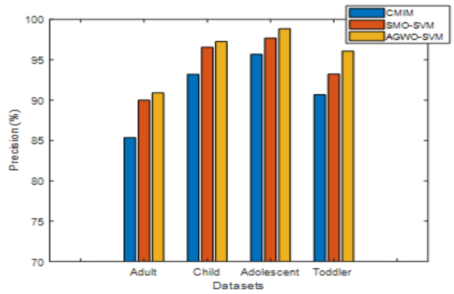

Fig.5. Comparative precisions of suggested and existing methods for classifying ASDs

The accuracy comparison between the proposed and current methods for identifying the ASDs data is shown in Fig. 5. Overall, the findings demonstrated that the proposed classification model outperformed the other machine-learning algorithms on the datasets. These results are consistent with prior discovered error rates and can be attributed to non-redundant rule sets of the proposed classification model. According to the findings, the suggested AGWO-SVM approach provides higher precision results than the ones currently used for classification.

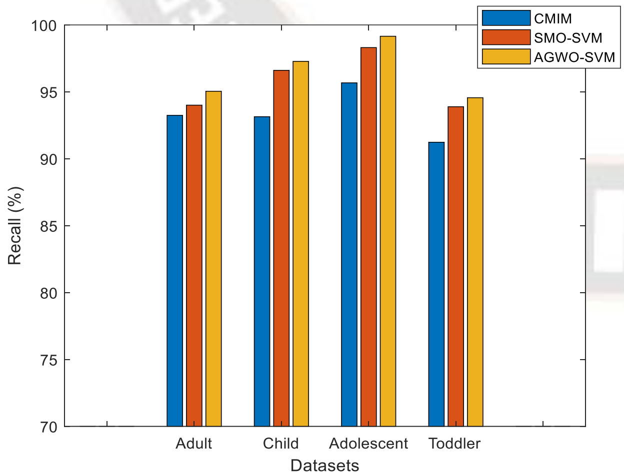

Fig.6. Comparative recalls of suggested and existing methods for classifying ASDs

The memory comparison between the proposed and current methods for identifying the ASDs data is shown in Fig. 6. The information utilised in this article was largely concerned with identifying people who had ASD symptoms, including different traits that frequently affect the diagnosis. The prediction model is therefore viewed as a classification difficulty that results from having ASDs or not. The results of applying the suggested supervised models to the given problem were then analysed and appraised. The feature selection approach must be utilised to support the assessment findings and the accuracy of the models prior to deploying them by deleting the weak variables from the databases. The offered supervised classification models and feature selection techniques based on AGWO were deemed appropriate for diagnosis of ASDs.

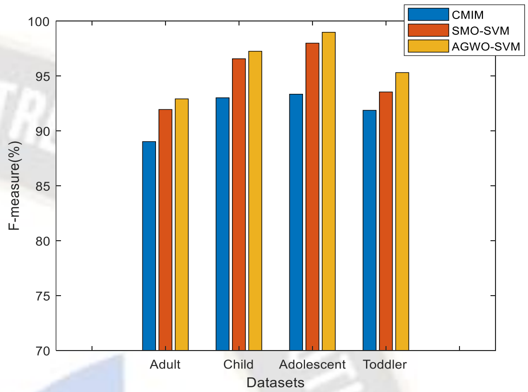

Fig.7. Comparative f-measures of suggested and existing methods for classifying ASDs

The results of the F-measure comparison between the suggested and actual approaches for diagnosing ASDs are shown in Figure 7. The toddler database variable, as indicated by the ASDs test, has the strongest association with the target class, as shown by both the feature-selection and classification algorithms. When compared to other MLTs, the suggested model has higher accuracy values in all datasets and in comparisons with teenage databases, the toddler database produces the best f-measure values.. In terms of f-measures, the graph reveals that the proposed AGWO-SVM model beats the present techniques.

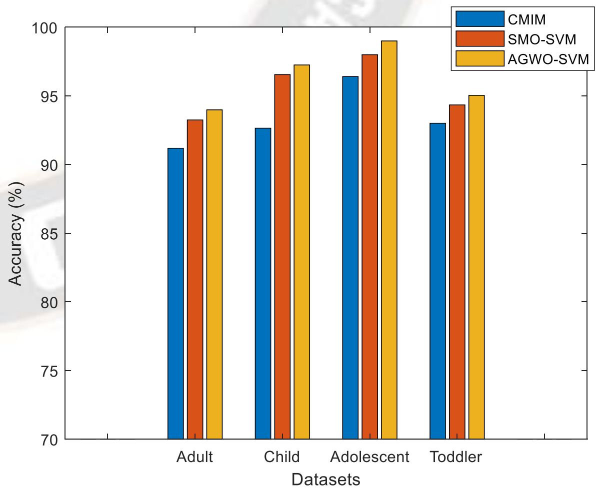

Fig.8. Comparative accuracies of suggested and existing methods for classifying ASDs

Figure 8 compares the classification accuracy of the data for ASDs using the proposed and existing approaches. A supervised machine learning model that can accurately anticipate the objective and generalise new instance predictions is considered successful. The accuracy of a model is often measured, and accuracy has two subtypes: sensitivity and specificity. According to the findings, the suggested AGWO-SVM strategy outperforms other current classification methods in terms of accuracy.

5. Conclusion

The main objective was to offer good model based on MLTs to identifysigns of ASDs in patients. Multiple steps have to be executed for selecting the best executing MLTs for these identifications. Also, the best questionnairesfor diagnosing ASDs for building database based on age groups from the executed steps. AGWOs were used in this study to determine the most significant features and efficient classification strategies in datasets of ASDs. Initially, the SMOTEs-based preprocessing strategy is used to remove extraneous data from the ASDs dataset. The AGWOs then repeats this procedure to find the smallest feature with the maximum classification recall and accuracy.

In order to assess whether a dataset instance has ASDs or not, KVSMs categorization is utilised. To find the best performing classifier for these binary datasets while taking recall, precision, F-measures, and classification errors into consideration, the experimental findings looked at datasets of ASDs from toddlers, children, adolescents, and adults. According to the results, the suggested AGWO-SVM technique has high accuracy results of 99 %for the teenage dataset, 93.98 %for the adult dataset, 97.25 %for the child dataset, and 95.03 %for the toddler dataset. In the future, the robust clustering algorithm will be used to efficiently classify ASDs data with enhanced accuracy outcomes.

REFERENCES

1. Abdullah, A. A., Rijal, S., & Dash, S. R. (2019, November). Evaluation on machine learning algorithms for classification of autism spectrum disorder (ASD). In *Journal of Physics: Conference Series* (Vol. 1372, No. 1, p. 012052). IOP Publishing.
2. Hyde, K. K., Novack, M. N., LaHaye, N., Parlett-Pelleriti, C., Anden, R., Dixon, D. R., &Linstead, E. (2019). Applications of supervised machine learning in autism spectrum disorder research: a review. *Review Journal of Autism and Developmental Disorders*, *6*(2), 128-146.
3. Rakesh, U. Ramya, V Senthil Murugan, V.(2023). Develop a 7 Layers Convolution Neural Network and IoT-Based Garbage Classification System. International Journal of Intelligent Systems and Applications in Engineering, 2023, 11(3), pp. 268–275.
4. Rahman, M. M., Usman, O. L., Muniyandi, R. C., Sahran, S., Mohamed, S., &Razak, R. A. (2020). A Review of machine learning methods of feature selection and classification for autism spectrum disorder. *Brain sciences*, *10*(12), 949.
5. Maenner, M. J., Yeargin-Allsopp, M., Van Naarden Braun, K., Christensen, D. L., &Schieve, L. A. (2016). Development of a machine learning algorithm for the surveillance of autism spectrum disorder. *PloS one*, *11*(12), e0168224.
6. Veeresh Babu, M. Ramya, V. Senthil Murugan, V.(2022) Design and development of embedded/SQL home based upper limb rehabilitation. International Journal of Healthcare ManagementThis link is disabled., 2023, 16(4), pp. 513–523.
7. VERMA, M., KUMAR, D., & ARORA, T. (2020). A review of machine learning techniques for feature based classification of autism spectrum disorder. *Adv. Appl. Math. Sci*, *19*(5), 387-396.
8. Praveena, T. L., & Lakshmi, N. M. (2020). Detection of Autism Spectrum Disorder Effectively Using Modified Regression Algorithm. *Emerging Research in Data Engineering Systems and Computer Communications*, 163-175.
9. Rakesh, U. Ramya, V Senthil Murugan, V.(2023). Classification, Collection, and Notification of Medical Waste Using IoT Based Smart Dust Bins. Ingenierie des Systemes d'InformationThis link is disabled., 2023, 28(1), pp. 149–154.
10. Sharma, M. (2019). Improved autistic spectrum disorder estimation using Cfs subset with greedy stepwise feature selection technique. *International Journal of Information Technology*, 1-11.
11. Abirami, S. P., Kousalya, G., &Balakrishnan, P. (2020). A Meta-Heuristic Model Based Computational Intelligence in Exploration and Classification. *Intelligent Systems, Technologies and Applications: Proceedings of Fifth ISTA 2019, India*, *1148*, 61.
12. Selvaraj, S., Palanisamy, P., &Parveen, S. (2019, September). Autism spectrum disorder prediction using machine learning algorithms. In *International Conference On Computational Vision and Bio Inspired Computing* (pp. 496-503). Springer, Cham.
13. Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., ... &Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature reviews Disease primers*, *6*(1), 1-23.
14. Nation, K., Clarke, P., Wright, B., & Williams, C. (2006). Patterns of reading ability in children with autism spectrum disorder. *Journal of autism and developmental disorders*, *36*(7), 911-919.
12. Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., &Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, *17*, 16-23.
13. Raj, S., & Masood, S. (2020). Analysis and detection of autism spectrum disorder using machine learning techniques. *Procedia Computer Science*, *167*, 994-1004.
14. Hossain, M. D., Kabir, M. A., Anwar, A., & Islam, M. Z. (2021). Detecting autism spectrum disorder using machine learning techniques. *Health Information Science and Systems*, *9*(1), 1-13.
15. Omar, K. S., Mondal, P., Khan, N. S., Rizvi, M. R. K., & Islam, M. N. (2019, February). A machine learning approach to predict autism spectrum disorder. In *2019 International conference on electrical, computer and communication engineering (ECCE)* (pp. 1-6). IEEE.
16. Thabtah, F. (2017, May). Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment. In *Proceedings of the 1st International Conference on Medical and health Informatics 2017* (pp. 1-6).
17. Alzubi, R., Ramzan, N., &Alzoubi, H. (2017, August). Hybrid feature selection method for autism spectrum disorder SNPs. In *2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (pp. 1-7). IEEE.
18. Abdolzadegan, D., Moattar, M. H., &Ghoshuni, M. (2020). A robust method for early diagnosis of autism spectrum disorder from EEG signals based on feature selection and

DBSCAN method. *Biocybernetics and Biomedical Engineering*, *40*(1), 482-493.
19. Shi, C., Zhang, J., & Wu, X. (2020). An fMRI feature selection method based on a minimum spanning tree for identifying patients with autism. *Symmetry*, *12*(12), 1995.
20. Mostafa, S., Yin, W., & Wu, F. X. (2019, November). Autoencoder based methods for diagnosis of autism spectrum disorder. In *International Conference on Computational Advances in Bio and Medical Sciences* (pp. 39-51). Springer, Cham.