An Efficient Autism Spectrum Disorder Classification using Modified Artificial Immune System

Abinaya S[1], Dr. W. Rose Varuna[2]
[1]*Ph.D Research Scholar, Department of Information Technology, Bharathiar University, Coimbatore – 641046, contactdrabinaya@gmail.com*
[2]*Assistant Professor, Department of Information Technology, Bharathiar University, Coimbatore – 641046, rosevaruna@buc.edu.in*

KEYWORDS
Accuracy, machine learning, AIS, classification, mathematical model, and Markov chain.

ABSTRACT:
Autism Spectrum Disorder (ASD) is a childhood disability that interferes with social interaction and communication, as well as patterns of behavior. This paper introduces a method for the classification of Autism Spectrum Disorder (ASD) employing a Modified Artificial Immune System (M-AIS). The proposed system improves the AIS by adding dynamic feature extraction and optimization that improves the classification of sensory, motor and genetic condition related data. The model fits to existing ASD diagnostic models that have shortcomings like static features classification and rigidity in data features. The modified AIS uses clonal selection, mutation, and affinity maturation to refine decision boundaries to increase diagnostic accuracy. The proposed system was evaluated on features from ASD data; the system was accurate and fast in its classification. The modified AIS offered improved real-time adaptability and highest 95.12% accuracy of the predictions thatis more robust than existing machine learning models. This method is a good solution for early diagnosis of ASDs because it offers clinicians a better, flexible instrument for testing existence of ASD features in people.

1. Introduction

Autism Spectrum Disorder (ASD) is a permanent neurological disorder characterised by problems in communication, interaction, and repetitive or persistent behaviours. It is manifested by impaired speaking, writing, gestures, and comprehension of ideas and feelings, as well as by repetitive actions [1]. The term 'spectrum' has been used because there are a wide variety of symptoms, skills, and intensity with which individuals with ASD. ASD is best diagnosed in early childhood, with the onset of symptoms in the second and third year of life but may be diagnosed later in life. Machine learning and Artificial Intelligence have offered new possibilities for exploring big data connected to ASD. These computational methods have the potential to enhance the ascertainment of the initial symptoms, customize management strategies, and enhance the knowledge of the pathophysiology of ASD [2]. Scientists are now increasingly trying to use advanced statistical analysis techniques for analyzing behavioural, genetic and neuroimaging data for better diagnosis of ASD [3].

ASD data analysis includes dealing with different types of data, behavioural data, neuroimaging data, and genetics data, to gain a better understanding of ASD and increase diagnostic accuracy. This information is gathered from clinical evaluation tools, structured and self-administered questionnaires or observational research and then processed using statistical tools or artificial neural networks (ANN) respectively [4]. Data pre-processing, cleaning, normalization, and dealing with missing values, which is important before applying any analytical models. Subset selection methods are employed to determine which are the most significant input variables, for example, social communication impairments, stereotyped behaviours, or sensory concerns. Studies that have applied decision trees (DT), support vector machines (SVM), and neural networks (NN) in machine learning have recorded success in recognizing features that set ASD away from other developmental complications [5].

The nature of ASD and its wide range of possible manifestations within the spectrum makes high-dimensional data analysis an intricate process. More complex approaches, including deep learning and ensemble learning, are gradually becoming more popular to increase the correspondence of ASD classification and the potential for individual assistance [6]. Metaheuristic learning algorithms are optimization methodologies which are specifically engineered to reduce candidate solution problems. They are being used more and more for the classification of ASD because of the nature and variability of data on ASD. Most of the conventional machine learning approaches fail to work well for high-dimensional and noisy data that are inherent to ASD [7].

Several metaheuristic optimization algorithms such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) techniques in which feature selection, model parameters or classifiers' performances can be optimized [8]. It would therefore appear reasonable to be able to understand the metaheuristic learning algorithms as working by searching the solution space more efficiently thus not getting trapped in local optima which is disadvantageous to the conventional machine learning algorithms. They are especially useful when sorting and categorizing ASDs where there are large and multiple data sets that need to be exploited fully regarding their features [9]. For instance, the hyperparameters of neural networks, or the performance of the ensemble models that will improve the classification can be optimized by the metaheuristic algorithms.

The use of metaheuristic learning algorithms in the classification of ASD is necessitated by the characteristics of ASD data [10]. ASD is frequently structured with numerous variables concerning behaviour, cognition, and genetics, and it is challenging to train a standard machine learning technique that extracts meaningful patterns without overfitting or losing generality [11]. Metaheuristic algorithms become a viable solution when global search algorithms are included to perform feature selection and model training. A problem with this classification is that the symptoms in ASD are quite diverse and hence it cannot be easily categorized into a particular model.

Metaheuristic algorithms can be used to fine-tune the selection of the features to the specific individuals with ASD in different subgroups. This flexibility leads to more accurate and detailed classifying models. ASD datasets are noisy and missing some values, which may affect the performance of conventional classification methods. Metaheuristic algorithms are less vulnerable to such problems as they can perform global optimization, and can easily look for an optimum solution within large, noisy search space. This leads to models of greater accuracy that are constructed to address characteristics of the data associated with ASD [12].

The metaheuristic learning algorithms can enhance the classification models of the ASDwhich increases the possibilities of better generalization in other data sets that eventually turn out to be very helpful for the early screening of the ASD and formulation of individual educational plans. This requiresthe identification of models that reflect the relationships between the ASD manifestations, behaviour, and neural elements of the disorder.

Artificial Immune System (AIS) a biologically inspired computational model is incorporated in the classification process of ASD to enhance the performance by minimizing feature selection and maximizing detection as well as prediction [13]. The algorithms, like clonal selection, and negative selection models of AISs, can categorize ASDs based on identifying the irregular pattern in the data. AIS can be characterized by a high level of flexibility and stability when working with many features and great data complexity of ASD, which allows for maximising the classification performance and increases the accuracy of the diagnostics [14].

The research article for ASD classification using AIS is organised as follows: the overview of ASD, need for ASD classification and AIS for ASD is detailed in Section 1, the recent comprehensive analysis with research gap and how it is addressed by proposed Modified AIS is discussed in Section 2, the proposed Modified AIS (M-AIS) for ASD classification is given in Section 3, experimental analysis with discussion is given in Section 4, and the article is concluded with further research direction in Section 5.

2. Related Works

Harrison et al., 2021 explored the motor and sensory qualities in ASD classification. Sensory features were integrated with RDoC features by support vector classification, and the diagnostic accuracy was significantly enhanced. Sensory features did a good job of differentiating ASD from DCD when used by themselves. The study is crucial for decoding the ASD phenotype and improving the RDoC framework's diagnostic classification of sensory and motor atypicality's [15].

To diagnose Autism Spectrum Disorder (ASD), Andrade et al. (2021) suggested a machine learning and verbal decision analysis protocol. The hybrid methodology is expected to improve efficiency since criterion in the existing protocols like the ICD-10 can be refined. Using the machine learning models, the study enhanced diagnostic accuracy and shortened the evaluation period. The proposed

approach was evaluated on a dataset of diagnosed individuals in Brazil, and the results indicate the possibility of using the approach for early ASD diagnosis [16].
Sterrett et al. (2024) examined the application of functional classifications of autism in adulthood. In the clinical interviews, they developed descriptions of the characteristics of autistic adults in form of vignettes. Three clinicians from expert settings used a feature list to sort the participants in terms of support requirements and found that the autism spectrum has discernible subgroups. These subgroups were compared with other types, for example, Profound Autism, also pointing to the need for dichotomous classifications in research and practice [17].
Ismail et al. proposed a new hybrid ensemble-based classification model, known as HEC-ASD, for the prediction of ASD genes using gene ontology data. The identified genes were predicted using HGS methods and gradient boosting machines to improve ASD gene prediction. When compared with the previous gene network-based model, the HEC-ASD model achieved the high classification accuracy of 88% evaluated by the SFARI gene database of ASD, which underlines the potential and efficiency of the ensemble learning in the prediction of ASD genes [18].
In their recent study, Kim et al. (2022) employed multimodal MRI data to differentiate children with low-functioning ASD at a preschool age. Based on T1-weighted MRI and DTI data, the machine learning model was highly accurate with ASD classification, with 88.8% correct classification rate compared to typically developing controls. It was also found that cortical thickness and diffusion parameters are the features that are significant to the model. Finally, the study focused on the use of multimodal MRI analysis for early diagnosis of ASD and stressed on the neuroimaging techniques[19].
Farooq et al.(2023) used federated learning (FL) in the diagnosis of Autism Spectrum Disorder (ASD) in children and adults through the use of machine learning models. FL was used for training local classifiers (logistic regression and support vector machine) before the results were combined at the central server. FL model pointed to a 98% accuracy in children, and 81% in adults, proving that FL is capable of diagnosing ASD without compromising data privacy[20].
Liao et al. (2022) built a hybrid deep learning model using both EEG and eye fixation and facial expression signals for ASD classification in children. A weighted Naive Bayes classifier was trained with 87.50% classification accuracy on the given data. The study thus opens up the question of the benefits of combined physiological and behavioral data in enhancing the diagnosis of early stage ASD, with EEG being the most discriminative[21].
The research gaps in this literature point to a need for better combination of motor and sensory aspects, accurate early identification procedures, and enhancement of genetic and neuroimaging methods for Autism Spectrum Disorder (ASD) categorization. real-time and dynamic features of sensory-motor classification were not addressed by Harrison et al. (2021) [15] and more effective optimization methods were required for early-stage identification by Andrade et al. (2021) [16]. These shortcomings can be mitigated by mathematically modifying AIS by incorporating dynamic feature extraction and classification techniques which respond to variations in sensory-motor signals.
These self-learning features make this system capable of enhancing more rigid approaches to machine learning, for example, support vector classification, where the decision boundary can be tweaked according to dynamic patterns in sensory data. AIS also incorporates gene-based models for correct identification of ASD, where immune-inspired techniques are included to enhance the ensemble learning models. This approach improves classification accuracy and time and can be used for early diagnosis of ASD.

3. Proposed Methodology- Classification using Modified Artificial Immune System

This section explains the classification of Autism Spectrum Disorder (ASD) using Mathematically Modified Artificial Immune System (M-AIS). Initially, the data is processed using feature scaling and dimensionality is handled using Principal Component Analysis (PCA). Further, the processed data is utilized in classification based on diverse class labels and the overall methodology is given in Figure 1.

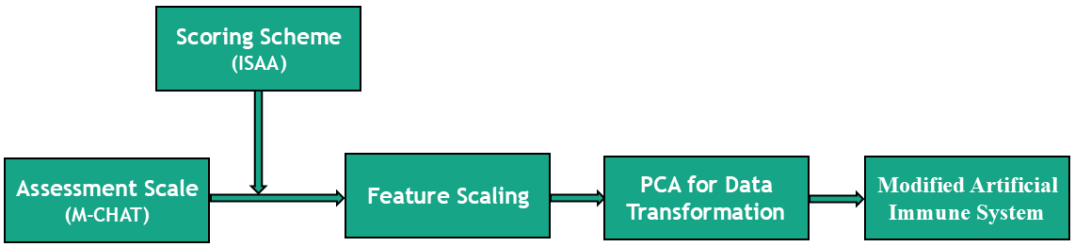

Figure 1. Overall Research Methodology

3.1. Pre-Processing and Feature Selection

Feature scaling is an essential component of preprocessing, which is needed when the data is given to machine learning models. When using this dataset feature scaling is applied to prevent features with large scales or units from influencing the model more than the other features. For instance, the answers to the questions in M-CHAT are on different scales and when the scales are not well-adjusted, the model is skewed. Two popular scaling methods are used, which includes normalization and standardization. Normalization makes the features between 0-1 while Standardization makes the data have a mean of 0 and standard deviation of 1. It also helps to balance between the features because none of them are allowed to dominate the learning process.

Principal Component Analysis (PCA) is used to reduce the dimensionality of the dataset once again. PCA seeks to find the direction along which most of the data variation exists thus allows one to concentrate on the most prominent data features. The first principal component (PC1) has a reasonable amount of data spread having a score of – 3.2791547 to 3.2677400. The second component (PC2) which contributes to the second largest variance varies between -4.6311653 and 1.5933235. The next components PC3 to PC5 account for minor variation. This works for reducing the dimensionality of the model and in effect offers accuracy in the important features which are essential for correct classification while reducing general information otherwise that can be overwhelming to the model.

3.2. Classification

The idea behind the AIS mimics the workings of the biological immune system to resolve computational problems. The AIS consists of several key processes: The model usually includes algorithms such as Negative Selection and Clonal Selection to represent recognition as well as response capabilities. Fusing the Clonal Selection Algorithm (CSA) with the Immune Network Model (INM) and Mathematically Modified Markov Chains (MMC) can boost diversity where the memory-based learning in tasks of optimization and classification. Joining the interaction and memory characteristics of the immune system with Markov chains' stochastic state changes will craft a more efficient algorithm for autism severity categorization (Mild Autism to Severe Autism). The CSA is the essential element by developing replicas of highly effective antibodies and enhancing their efficacy through mutations. The Immune Network Model (INM) creates interactions among antibodies that stimulate and suppress one another and promotes diversity while preserving long-term memory. MMC controls the movement between distinct states or solutions by reflecting past data probabilities which facilitates better prediction accuracy and aids in tweaking mutation processes. In a multi-step approach this model allows each antibody (classifier) to diversify and mutate driven by both the interactions within the network (from INM) and state alterations (from MMC).

Antibodies and its State Initialization

Let A indicate the antibodies population $\{a_1, a_2, a_3, \ldots\ldots, a_n\}$, where the severity of autism is categorized antibody a_i is given as four states as $S = \{S_0, S_1, S_2, S_3\}$. where S_0 indicates no autism, S_1

indicates mild autism, S_2indicates moderate autism, and S_3 indicates severe autism.Every a_i antibody is related with a categorisation state S_i that is determined by a probabilistic transitionis given in Equation 1.

$P(S_i|a_i) = \{p_{i0}, p_{i1}, p_{i2}, p_{i3}\}$---------(1)

where p_{ij} indicates the antibody probability a_i that categorize the input information into S_j.

The affinity function measures how well an antibody will fit the input data we are working with. However, the immune network provides the stimulation and suppression interactions which are not considered in the machine learning model. The antibody function for antibody is given in Equation 2.

$f(a_i, x) = \frac{1}{1+d(a_i,x)}$---------(2)

where $d(a_i, x)$ is the distance between the antibodya_i and input$x$, that is the measure of the differences between them.

The network interaction matrix is given in Equation 3.

$A = [a_{ij}]$ ---------(3)

The simulation function in the network is given in Equation 4.

$S(a_i, a_j) = \alpha \cdot e^{-\beta d(a_i,a_j)}$---------(4)

where α and β are constant values and the distance among antibodies a_i and a_j is $d(a_i, a_j)$.

The suppression function is given in Equation 5.

$P(a_i, a_j) = \gamma \cdot \frac{1}{1+d(a_i,a_j)}$---------(5)

where the suppression intensity is indicated by γ controls.

The concentration update for every antibody is given in Equation 6 and it update the concentration based on interactions among the network.

$c(a_i) = c(a_i) + \sum_j S(a_i, a_j) - \sum_j P(a_i, a_j)$---------(6)

A stochastic matrix P represents the transition from these classification states based on the historical records. Originally each antibody goes through state transitions according to past data. Let P=[p_{ij}] be the Markov transition matrix, where p_{ij} is the probability of moving from state S_i to state S_j. This Markov chain transition is expressed in Equation 7.

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{bmatrix}$$---------(7)

Every element p_{ij} satisfies the condition

$$\sum_j p_{ij} = 1$$

The state transition probability for antibody a_i, the probability of transitioning from state S_i to state S_j is given by Equation 8.

$P(S_j|S_i, a_i) = P_{ij}$---------(8)

These transition probabilities affect the mutation process so that antibodies are guided to states observed in the state transition process. For example, if, from one time to the other, a classifier frequently changes from Mild to Moderate, the Markov chain will help to mutate to the Moderate state when classifying the similar patients.

Cloning and mutation

Cloning and mutation are two of the activities of the clonal selection, which has been extended in this paper to include state transitions and network effects. High-performing classifiers (antibodies with high concentration $c(a_i)$) are cloned. The number of clones C_i is proportional to their affinity and concentration is given in Equation 9.

$C_i = \alpha \cdot f(a_i, x) \cdot c(a_i)$ ---------(9)

where the scaling factor is indicated by α.

The mutation rate $\mu_i$is regulated by both the states namely Markov and affinity state of transition probabilities is given in Equation 10.

$\mu_i = \frac{1}{f(a_i,x)} \cdot P(S_i, S_j)$ ---------(10)
where $P(S_i, S_j)$ is the Markov transition probability from state S_i to state S_j.
The mutated clones $A'=\{a_1',a_2',...,a_m'\}$ are generated as in Equation 11.
$A' = \{a_i' | \mu_i \cdot f(a_i, x) \cdot c(a_i) \geq \tau\}$---------(11)
Mutations are performed on the cloned antibodies. Here, the dynamics of mutation rates are regulated by Markov chains considering transition probabilities. Any antibodies that are expected to change frequently between specific states will change its structure to favour such states. This makes it possible for the mutation process to search for new classifications at the same time as it keeps the likely state transitions.

Selection and Network Adjustment
During the selective cloning and mutation, the affinity of each antibody is retested again. High-affinity antibodies are taken forward to the next generation. The new generation of antibodies A_{next} selected based on affinity and network interactions is given in Equation 12.
$A_{next} = select(A \cup A')$ ---------(12)
Such that

$$f(a_i, x) + c(a_i) \geq \theta$$

Where θ is considered as a selection threshold.
The final affinity function integrates both the immune network model and the Markov chain state transitions that is given in Equation 13.
$f'(a_i, x) = f(a_i, x) + \lambda_1 \sum_j S(a_i, a_j) - \lambda_2 \sum_j P(a_i, a_j) + \lambda_3 P(a_i, a_j)$---------(13)
where the weight factors namely stimulation, suppression, and transition state are indicated by λ_1, λ_2, and λ_3, respectively.
The transition probability from no autism to mild autism is given in Equation 14.
$P(S_1 | S_o, a_i) = p_{01}$---------(14)
The probability for severe autism from mild autism is given in Equation 15.
$P(S_3 | S_1, a_i) = p_{13}$---------(15)
The final concentration of after update of network and mutation in the antibody is given in Equation 16.
$c(a_i) = \sum_j S(a_i, a_j) - \sum_j P(a_i, a_j) + \mu_i$---------(16)
The immune network is modified, considering new values of stimulation and suppression for the new population. Markov chain is again modified with the latest classification behaviour to enhance the transition probabilities. This is accomplished by this hybrid affinity function which guarantees that only the best antibodies are chosen while the interactions and state transitions of the population controls the population dynamics.
No Autism to Mild Autism Transition: An antibody a_i may be highly selective of classifying as No Autism. If Markov chain analysis makes transitions between No Autism and Mild Autism as common, then mutations make a_i classify similar cases as Mild Autism. Classifiers for similar cases are antibodies so they can help each other that become more forceful while keeping the variety of classifiers down by suppressing and mutating them. Such integration improves CSA since it combines the memory of the immune network and the maintenance of diversity, with the state transition model of Markov chains. This hybrid system is especially suitable for classification problems such as Autism severity, as the interactions between classification states can have a prominent impact on the performance of the algorithm over time.

4. Result and Discussion

The data collected for this study was obtained through 550 parents' and caregivers completed Modified Checklist for Autism in Toddlers (M-CHAT). These responses are useful in establishing the biomarkers of Autism Spectrum Disorder (ASD) during infancy and toddlerhood. The M-CHATis an assessment of early manifestations of ASD, which makes it possible to investigate how particular

behavior in a child can be associated with the disorder. In this case, the data must be split into several sets so that it can have an accurate training, validation and test set to avoid over-fitting and improve model generalization.

This paper follows a training, validation, and test-data ratio of 60:20:20 whereas such division of data helps the model to get enough variation to learn from and generalize without memorizing details from the training set. This task and testing can be instrumental in assessing performance on new-data and tuning the parameters of the model. To increase the size and quality of the dataset, synthesized data is produced using statistical characteristics obtained from 550 responses. The synthetic data and the original dataset are then analyzed using the newest version of R, which is the programming language mostly used for statistical computing and machine learning.

The emphasis of the study is to assess the effectiveness of the proposed M-AIS algorithm for solving a given problem. This algorithm is evaluated on the M-CHAT dataset using several performance indicators including accuracy, precision, recall, F1 score, sensitivity, and specificity. These metrics give a complete depiction of the algorithm classification characteristics, and crucial for evaluating the algorithm's ability to diagnose ASD. The effectiveness of the M-AIS algorithm is evaluated against the existing methods, namely DT, HEC-ASD, and FL.

4.1. Performance Metrics

Accuracy is the ability of the model in the differentiation of precise incidence or cases from the imprecise or rather unrelated ones, providing overall efficiency. Recall tests the validity of the model in correctly classifying all the children with autism. Recall is the interest of the total population of individuals with ASD that the model canacquire, or the ability of the model to correctly identify ASD people. The F1-score integrates precision and recall for a balanced measure, and is valuable when the data classes are imbalanced. Sensitivity, another performance measurement index, represents the likelihood with which the model discerns the total correct persons with autism. This is well understood in clinical diagnosis because a false negative outcome can be counterproductive. Specificity shows the model's ability to capture the children without autism correctly while minimizing misidentification. This result can assist in constructing more accessible and accurate ASD diagnostic tools, which is a severe deficiency in early childhood.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \text{---------(17)}$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \text{---------(18)}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \text{---------(19)}$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \text{---------(20)}$$

$$Specificity = \frac{True\ Negative}{True\ Positive + False\ Positive} \text{---------(21)}$$

Table 1. Comparison of Accuracy

Sample Count	DT	HEC-ASD	FL	M-AIS
100	83.02	79.56	81.23	89.67
200	83.67	81.43	82.7	90.89
300	84.45	82.3	84.3	91.53
400	85.82	82.9	85.66	93.46
500	86.88	84.09	86.2	95.12

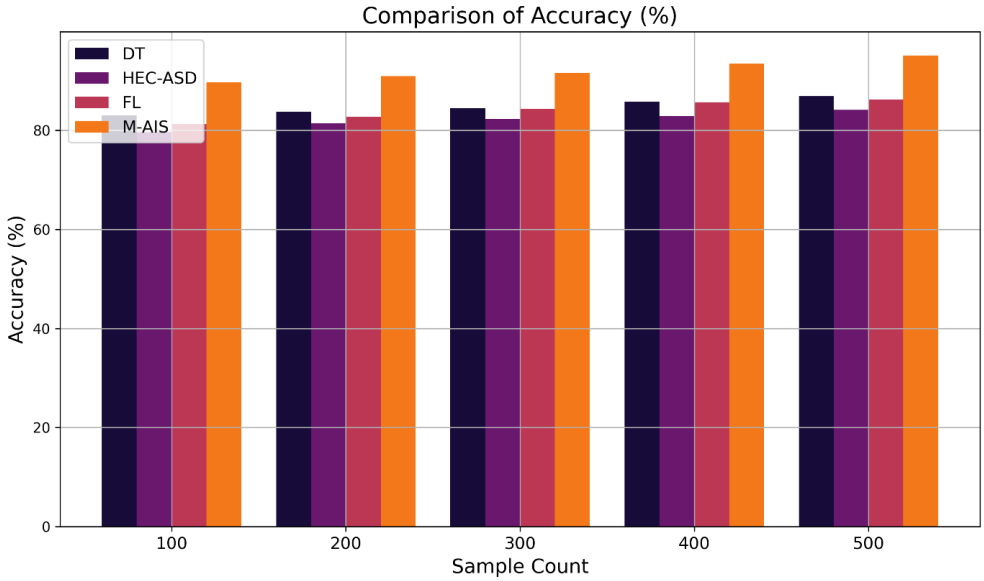

Figure 2. Comparison of Accuracy

In Table 1, the results of experimental comparison of M-AIS are accomplished with three existing methods include Decision Tree (DT), Hybrid Ensemble based Classification for Autism Spectrum Disorder (HEC-ASD), and Federated Learning (FL). This shows that the M-AIS performs higher in terms of accuracy in all cases of sample sizes. 87% for 100 samples, 95.12% for 500 samples, which is much higher than DT's value of 86.88% and HEC-ASD's value of 84.09%. From this it can be deduced that M-AIS can provide a better classification of ASD data. The further incremental improvements in the accuracy depicts good generalization properties of the M-AIS approaches particularly when the sample size is large.

Table 2. Comparison of Precision

Sample Count	DT	HEC-ASD	FL	M-AIS
100	86.71	82.23	83.14	90.56
200	86.96	82.34	83.27	90.9
300	87.61	83.46	83.92	91.45
400	88.85	84.54	84.33	92.2
500	89.19	85.09	85.62	93.67

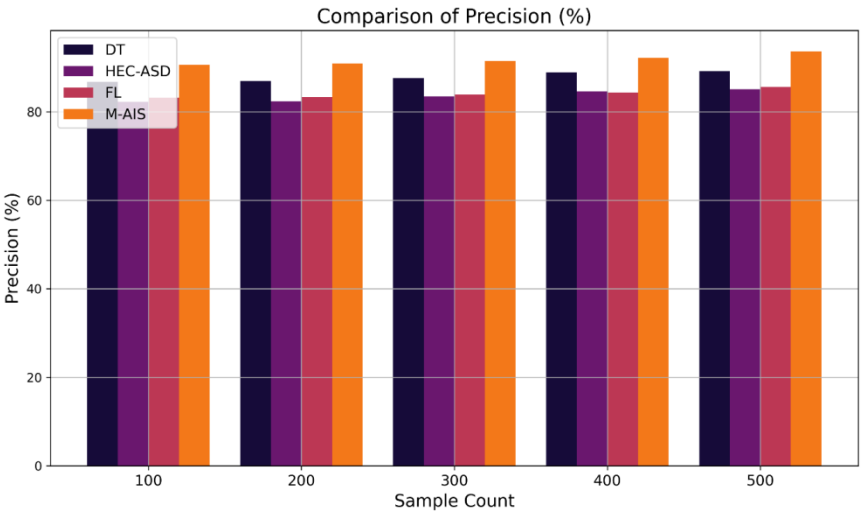

Figure 3. Comparison of Precision

Table 2 depicts the precision metrics and like in the previous results, M-AIS outcompeted the other methods. For example, with 100 samples, M-AIS reaches the highest precision of 90.56% while DT, HEC-ASD, and FL have lower, scores of 86.71%, 82.23% and 83.14% respectively. This difference remains consistent in relation to the sample sizes, and with M-AIS achieving 93.67% precision at 500 samples. High precision means that M-AIS has fewer false positives than other classifiers, and thus M-AIS is a better classifier for situations where the false positive rate is expensive since it is vital in diagnosing the diseases.

Table 3. Comparison of F1-Score

Sample Count	DT	HEC-ASD	FL	M-AIS
100	88.91	84.7	85.63	92.88
200	89.03	84.93	86.88	93.16
300	89.9	85.46	87.09	93.55
400	91.9	86.12	87.83	94.58
500	92.56	86.99	89.01	94.91

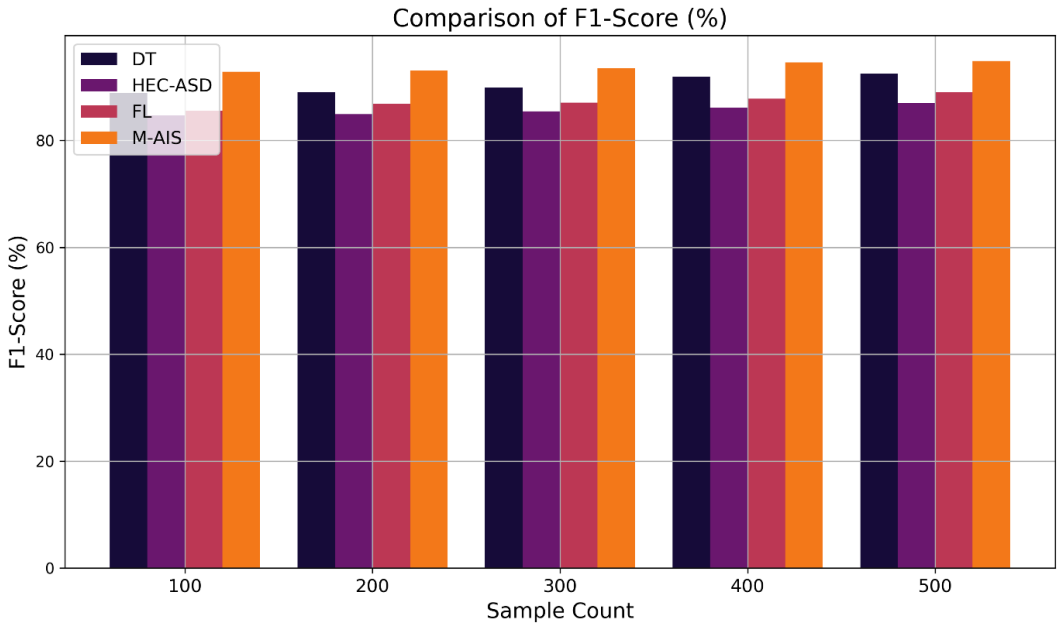

Figure 4. Comparison of F1-Score

In Table 3, the F1-Score, indicates precision as well as the recall in the same model, and reasserts the efficiency of M-AIS. The performance of the proposed method on 100 samples is as follows: M-AIS: 92.88%, DT: 88.91%, HEC-ASD: 84.7%, and FL: 85.63%. The F1-Score for the M-AIS approach remains consistently higher in all cases and the score reaches a 94.91% at 500 samples. This shows that the model does not have a problem in handling imbalances in class by affording the right balance in precision and recall.

Table 4. Comparison of Sensitivity

Sample Count	DT	HEC-ASD	FL	M-AIS
100	88.71	85.03	83.18	90.9
200	88.82	85.23	84.49	91.7
300	89.46	85.59	85.47	92.4
400	89.69	86.22	86.04	92.66
500	90.12	87.3	86.33	93.1

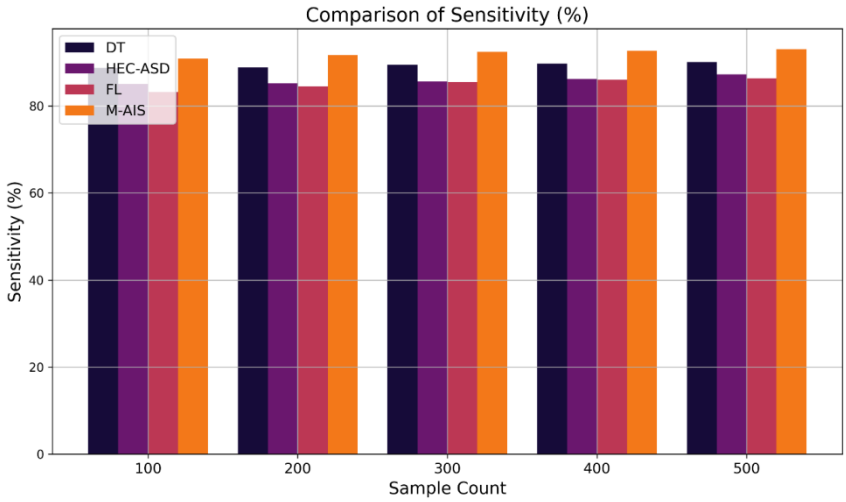

Figure 5. Comparison of Sensitivity

Sensitivity is shown in Table 4 and reflects the model's performance in correctly predicting true positive, that is, people with ASD. M-AIS has the highest sensitivity, 90.9% at 100 samples and 93.1% at 500 samples, better than DT 90.12%, HEC-ASD 87.3%, and FL 86.33%. M-AIS has a high sensitivity thus rare cases of ASD are likely to be overlooked, making it a very efficient tool whenever detection of ASD cases must be accurate.

Table 5. Comparison of Specificity

Sample Count	DT	HEC-ASD	FL	M-AIS
100	4.56	3.12	3.85	10.23
200	5.33	3.54	4.16	11.25
300	7.45	5.76	5.01	15.45
400	8.63	8.97	6.9	10.56
500	9.15	9.32	7.46	12.17

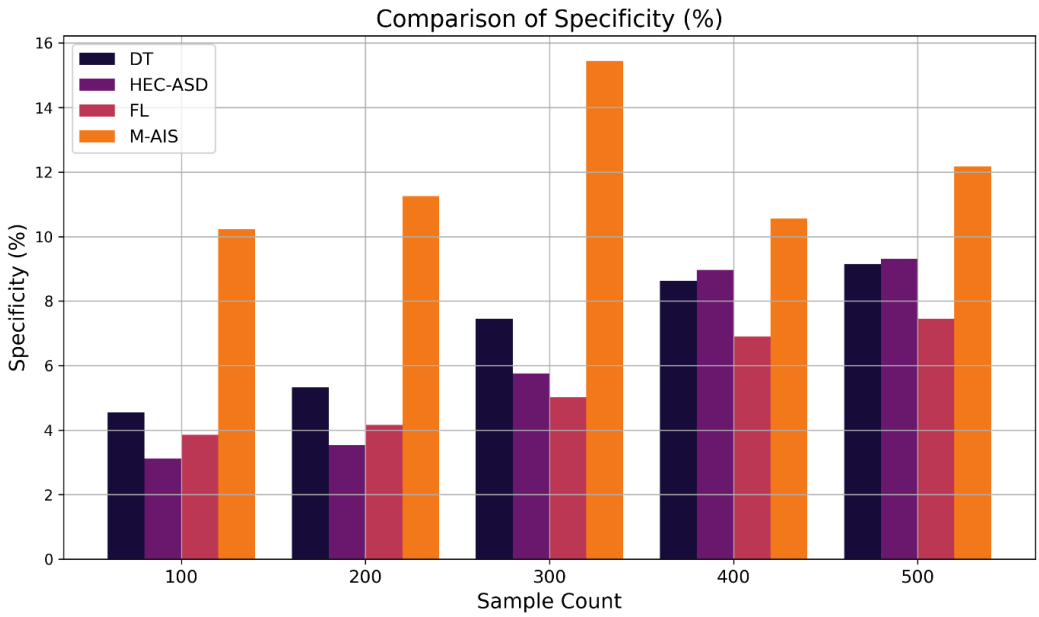

Figure 6. Comparison of Specificity

Table 5 represents specificity which measures the ability of the model to correctly categorize negative class (people without ASD). More specifically, using M-AIS is both more accurate and efficient than the other methods explored here, especially when the sample size is small. For 100 samples, M-AIS

yields 10.23% specificity while DT is 4.56%, HEC-ASD is 3.12%, FL is 3.85%. This trend goes on and on as the number of samples increases, M-AIS is 0.1217 at 500 samples for instance. Greater specificity means that the M-AIS avoids high number of missed rates, making it a well-rounded and highly reliable classifier that does not generalize too many people as non-ASD. In each of the accuracy, precision, F1-Score, sensitivity, and specificity measures M-AIS provides a stronger classification of ASD than DT, HEC-ASD or FL. Because it performs well even when there is a variation of sample size, it is more accurate and efficient for diagnosing ASD.

5. Conclusion

Autism Spectrum Disorder (ASD) affects children's social interactions, communication, and their behaviors. This paper proposes the classification of ASD through a newly developed method known as the Modified Artificial Immune System (M-AIS), as it also improves the extraction and optimization of features. The dynamic aspect of M-AIS overcomes the shortcomings of the diagnosis model by implementing the clonal selection, mutation, and the affinity maturation for fine tune of the decision boundary. The proposed system has a high accuracy rate of 95.12%, and is more flexible and efficient than most other machine learning models. These improvements in accuracy and flexibility make M-AIS a useful tool in the early identification of ASD and provide clinicians with a more accurate and real-time adaptable means of identifying ASD traits. The M-AIS framework that avoids the drawbacks of traditional static models can become the basis for a radical rethinking of the diagnostics of ASD within the framework of pediatric healthcare practice on a large scale, with high accuracy and efficiency.

Future upgrades can explore a bigger sample size by including various resources, though it is also possible to include multimodal data such as EEG, MRI, or genetic data to classify individuals with ASD accurately. Additionally, switching some of the activities to deep learning models or applying hyperparameter optimization steps will increase M–AIS performance. Other potential areas of adaptation might also encompass the incorporation of federated learning frameworks aimed at maintaining privacy whilst sharing global learning models as an approach to construct safer and more distributed diagnostic systems.

Reference

1. Sharma, S. R., Gonda, X., & Tarazi, F. I. (2018). Autism spectrum disorder: classification, diagnosis and therapy. *Pharmacology & therapeutics*, *190*, 91-104.
2. 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.
3. Yang, X., Zhang, N., & Schrader, P. (2022). A study of brain networks for autism spectrum disorder classification using resting-state functional connectivity. *Machine Learning with Applications*, *8*, 100290.
4. Jiang, W., Liu, S., Zhang, H., Sun, X., Wang, S. H., Zhao, J., & Yan, J. (2022). CNNG: a convolutional neural networks with gated recurrent units for autism spectrum disorder classification. *Frontiers in Aging Neuroscience*, *14*, 948704.
5. Gao, J., Chen, M., Li, Y., Gao, Y., Li, Y., Cai, S., & Wang, J. (2021). Multisite autism spectrum disorder classification using convolutional neural network classifier and individual morphological brain networks. *Frontiers in Neuroscience*, *14*, 629630.
6. Ali, N. A., Syafeeza, A. R., Jaafar, A. S., Alif, M. K. M. F., & Ali, N. A. (2020). Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm. *IAES International Journal of Artificial Intelligence*, *9*(1), 91-99.
7. Tripathy, H. K., Mallick, P. K., & Mishra, S. (2021). Application and evaluation of classification model to detect autistic spectrum disorders in children. *International Journal of Computer Applications in Technology*, *65*(4), 368-377.
8. Hu, Z., Wang, J., Zhang, C., Luo, Z., Luo, X., Xiao, L., & Shi, J. (2021). Uncertainty modeling for multicenter autism spectrum disorder classification using Takagi–Sugeno–Kang fuzzy systems. *IEEE Transactions on Cognitive and Developmental Systems*, *14*(2), 730-739.

9. Sujatha, R., Aarthy, S. L., Chatterjee, J., Alaboudi, A., &Jhanjhi, N. Z. (2021). A machine learning way to classify autism spectrum disorder. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(6), 182-200.
10. Zhang, M., Zhao, X., Zhang, W., Chaddad, A., Evans, A., &Poline, J. B. (2020). Deep discriminative learning for autism spectrum disorder classification. In *Database and Expert Systems Applications: 31st International Conference, DEXA 2020, Bratislava, Slovakia, September 14–17, 2020, Proceedings, Part I 31* (pp. 435-443). Springer International Publishing.
11. Mashudi, N. A., Ahmad, N., & Noor, N. M. (2021). Classification of adult autistic spectrum disorder using machine learning approach. *IAES International Journal of Artificial Intelligence*, *10*(3), 743.
12. Elshoky, B. R. G., Younis, E. M., Ali, A. A., & Ibrahim, O. A. S. (2022). Comparing automated and non-automated machine learning for autism spectrum disorders classification using facial images. *Etri Journal*, *44*(4), 613-623.
13. Devika Varshini, G., &Chinnaiyan, R. (2020). Optimized machine learning classification approaches for prediction of autism spectrum disorder. *Ann Autism Dev Disord. 2020; 1 (1), 1001*.
14. Gaspar, A., Oliva, D., Hinojosa, S., Aranguren, I., & Zaldivar, D. (2022). An optimized Kernel Extreme Learning Machine for the classification of the autism spectrum disorder by using gaze tracking images. *Applied Soft Computing*, *120*, 108654.
15. Harrison, L. A., Kats, A., Kilroy, E., Butera, C., Jayashankar, A., Keles, U., & Aziz-Zadeh, L. (2021). Motor and sensory features successfully decode autism spectrum disorder and combine with the original RDoC framework to boost diagnostic classification. *Scientific Reports*, *11*(1), 7839.
16. Andrade, E., Portela, S., Pinheiro, P. R., Nunes, L. C., Filho, M. S., Costa, W. S., & Pinheiro, M. C. D. (2021). A protocol for the diagnosis of autism spectrum disorder structured in machine learning and verbal decision analysis. *Computational and Mathematical Methods in Medicine*, *2021*(1), 1628959.
17. Sterrett, K., Clarke, E., Nofer, J., Piven, J., & Lord, C. (2024). Toward a functional classification for autism in adulthood. *Autism Research*.
18. Ismail, E., Gad, W., & Hashem, M. (2022). HEC-ASD: a hybrid ensemble-based classification model for predicting autism spectrum disorder disease genes. *BMC bioinformatics*, *23*(1), 554.
19. Kim, J. I., Bang, S., Yang, J. J., Kwon, H., Jang, S., Roh, S., ... & Kim, B. N. (2022). Classification of preschoolers with low-functioning autism spectrum disorder using multimodal MRI data. *Journal of Autism and Developmental Disorders*, 1-13.
20. Farooq, M. S., Tehseen, R., Sabir, M., & Atal, Z. (2023). Detection of autism spectrum disorder (ASD) in children and adults using machine learning. *scientific reports*, *13*(1), 9605.
21. Liao, M., Duan, H., & Wang, G. (2022). Application of machine learning techniques to detect the children with autism spectrum disorder. *Journal of Healthcare Engineering*, *2022*(1), 9340027.
22. R. Renukadevi et al. "An Improved Collaborative User Product Recommendation System Using Computational Intelligence with Association Rules",Communications on Applied Nonlinear Analysis**,** Volume 31, Issue 6s (2024),pp. 554 - 564**.** https://doi.org/10.52783/cana.v31.1243
23. K T, S., P, S. S., Kumar E, B., L R, S., J, V., & R., N. (2024). Experimental Assessment between Dissimilar Techniques and Methodologies to Sports Knee Injury using Magnetic Resonance Imaging. *South Eastern European Journal of Public Health*, 1635–1644. https://doi.org/10.70135/seejph.vi.2167
24. Ashfauk Ahamed, A. K., et al. "Prediction of The Growing Stock In Stock Market On Analysis Of The Opinions Using Sentiment Lexicon Extraction And Deep Learning Architectures." *Frontiers in Health Informatics* 13.3 (2024): 1382-1392.
25. R. Renukadevi et al. "Optimized Computational Intelligence with Association Rules foraCollaborative Consumer Product Recommendation System", Advances in Nonlinear Variational Inequalities, Volume 31, Issue 3s (2025), pp. pp. 375 – 385. https://doi.org/10.52783/anvi.v28.3065