









An Attention-Based Hybrid Optimized Residual Memory Network (AHRML) Method for Autism Spectrum Disorder (ASD) Detection

Muhanna K. A. Al-Muhanna^{1,*}, Amani Ahmed Alghamdi², Bahauddeen Alrfaei^{3,4}, Mohammad Afzal⁵, Reema Al-Subaiee^{3,4}
and Rania Haddadi⁶

¹Materials Science Research Institute, King Abdulaziz City for Science and Technology (KACST), Riyadh 11442, Saudi Arabia[✉]

²Department of Biochemistry, College of Science, King Saud University, Riyadh 11451, Saudi Arabia[✉]

³Stem Cells and Regenerative Medicine Unit, King Abdullah International Medical Research Center (KAIMRC), Riyadh 11481, Saudi Arabia[✉]

⁴King Saud bin Abdulaziz University for Health Sciences (KSAU-HS), Riyadh 11481, Saudi Arabia[✉]

⁵Department of Chemistry, College of Science, King Saud University, Riyadh 11451, Saudi Arabia[✉]

⁶Department of Zoology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia[✉]

Correspondence to:

Muhanna K. A. Al-Muhanna*, e-mail: mmuhanna@kacst.edu.sa, Mob: +966-560022888

Received: February 4 2024; Revised: March 13 2024; Accepted: March 13 2024; Published Online: April 17 2024

ABSTRACT

A neurological condition known as autism spectrum disorder (ASD) is marked by issues with speech, socializing, and monotonous activities. Even if genetics is the primary cause, prompt identification is vital, and using machine learning presents an intriguing opportunity to diagnose the condition more quickly and affordably. Even so, the specific problems of increased computational costs, extended execution times, and decreased efficacy concern the conventional approaches. In order to provide the highest level of disease prediction accuracy, the objective of the proposed study is to develop an automated tool for ASD detection that integrates a number of cutting-edge mining approaches. This study proposes a computer-aided and ultra-light framework called attention-based hybrid optimized residual memory network (AHRML) for accurate and efficient ASD detection. Here, a new hybridized Arithmetic Harris Hawks Optimizer is employed to minimize the dimensionality of features in order to streamline the disability identification process. Moreover, a sophisticated deep learning technique called attention-based residual term memory is developed to reliably and less frequently identify ASD from the provided data. The authors employed the ASD dataset to train and test the proposed model. The dataset includes demographic data (age and gender), behavioral characteristics (social skills and communication abilities), and ASD diagnosis data. In addition, a range of parameters were used to validate and test the proposed AHRML model's performance using the popular ASD dataset.

KEYWORDS

autism spectrum disorder (ASD), deep learning, artificial intelligence (AI), neurodisabilities, mental healthcare, optimization, attention-based hybrid optimized residual memory network (AHRML), rehabilitation, disability support

INTRODUCTION

A neurodegenerative disease, known as autism spectrum disorder (ASD), results in long-lasting deficiencies in children's behavior and social communication abilities (Bahathiq et al., 2024; Nogay and Adeli, 2024). The World Health Organization reports that 0.63% of children have an ASD diagnosis. It first manifests in childhood and then spreads to people of all ages. Most of the time, symptoms start to show up in the first 5 years of life. ASD is a severe neurodevelopmental disease that carries a substantial cost for medical care. Stereotypical behavior and a persistent lack of social contact are hallmarks of ASD (Kumar et al., 2024), and they are frequently coupled with a generalized decline in communication

abilities. There are neurological and genetic components to ASD. Social interaction, an aptitude for imagination and thinking, obsessive behaviors, and issues with communicating with others are all examples of behaviors associated with ASD (Gao et al., 2024; Saponaro et al., 2024). The precise processes that explain the connection between ASD and neurological problems are not well understood and are believed to include intricate interactions among genetic, environmental, and neurobiological variables. ASD is mostly defined by impairments in social communication and interaction, along with limited and repetitive behavior patterns. However, persons with ASD often have other neurological problems that

occur alongside. Many persons with ASD sometimes have difficulties in processing sensory information, which may manifest as either heightened sensitivity or reduced sensitivity to sensory stimuli, such as sound, touch, or light. The presence of sensory processing impairments may have a substantial effect on everyday activities and contribute to the behavioral symptoms associated with ASD. The three categories for social communication skills are deficiencies in interpersonal–emotional cooperation; problems in non-verbal communication actions; and deficiencies in growing, sustaining, and comprehending connections. Additionally, prohibited and monotonous patterns of behavior as well as obsessed interest are considered behavioral deficiencies in children diagnosed with autism. These behaviors are characterized by the following (Dhamale et al., 2023): (i) stereotypical or recurrent motor actions, (ii) reliance on uniformity, (iii) obstacles adapting to abrupt shifts in everyday or rigid sequences of either nonverbal or verbal conduct, or (iv) distinctive attitude toward visual facets. Hyper- or hypo-sensitivity to a variety of modalities, including hearing, smell, taste, and body awareness, can occur in children with learning problems. For instance, noises from air conditioners may be upsetting to kids who are extremely sensitive to noise. In addition, self-harm, standing on their tiptoe, thrashing of the fingers or firearms, and swaying from side to side are other activities that frequently correspond with perceptual problems (Wang et al., 2023). Early detection and treatment are crucial. Parents, educators, and other people without specialization or credentials can screen children for autism.

Automated methods for detecting ASD may aid clinicians by simplifying the process of screening and diagnosis, hence decreasing the amount of time and effort needed for evaluation. This enables physicians to concentrate on more intricate situations and provide individualized treatment to persons with ASD. Timely identification of ASD is essential for prompt intervention and treatment, resulting in improved outcomes for persons with ASD. Automated detection systems may expedite the process of identifying potential issues in children, enabling timely access to suitable therapies and support services. ASD detection technologies that are automated may be implemented on platforms that can be scaled, such as mobile apps or web-based platforms, which allows them to be available to a larger audience. This facilitates screening in many locations, such as schools, healthcare institutions, and distant places, hence enhancing the availability of early diagnosis and intervention services.

According to the American Academy of Pediatrics, screening should be performed on all children on a frequent basis to make sure they are receiving enough support to achieve their greatest potential. A healthcare provider can use the data acquired from the evaluation and observation procedure to gain a good knowledge of an individual's conduct or social skills prior to the diagnosing phase (Alvarez-Jimenez et al., 2020; Das Biswas et al., 2020). Only medical professionals are qualified to make diagnoses. Using established diagnostic tools, medical professionals will diagnose youngsters with autism. In a conversation with parents or other carers, they will also go over the child's past developmental trajectory and present behavior. Early assessment and therapy are possible for ASD. Clinical diagnosis

of ASD depends critically on early detection. Customized therapies can then be developed based on this diagnosis to enhance the quality of life of children with ASD and their carers. However, using costly testing procedures, the diagnosis of ASD can be a drawn-out process. The rise in ASD cases observed globally in recent times has prompted medical and scientific researchers to develop more efficacious screening techniques (Mohan and Paramasivam, 2021). Owing to the development of modern technology, we are now able to maintain enormous quantities of data. To make judgments according to the data gathered, data mining is an essential responsibility. It has to do with machine learning. The field of machine learning has made significant strides in the last several years, and applied fields such as medicine and biology are beginning to understand its value. Machine learning techniques are employed or indicated in medical decision-making and diagnosis to help with the interpretation of data (Lamani and Julian Benadit, 2023). As a result, machine learning-assisted disease evaluation methods are extensively researched. There are currently a number of studies on ASD, although they are not without limits. Some research just looks at medical professionals, and certain studies concentrate on gathering ASD data. It is also crucial to provide a rapid, simple, and easily accessible way to assist with the early identification of ASD. Seeking treatment from specialists is beneficial for families of patients with ASD. Many machine learning and deep learning techniques have been developed for ASD diagnosis and classification in the literature materials (Epalle et al., 2021; Yang et al., 2022). However, the traditional methods struggle with the particular issues of higher computing expenses, longer execution times, and reduced effectiveness. Therefore, the goal of the proposed study is to build an automated artificial intelligence (AI) tool for ASD detection that combines a number of cutting-edge mining techniques to guarantee the highest level of disease prediction accuracy (Chen et al., 2020; Fu et al., 2021). Additionally, the public University of California, Irvine (UCI) databases are used in this work to detect ASD. The extensive and widely used ASD datasets have a connection with the clinical diagnosis of ASD in people of all ages. The specific contributions of this work are listed below:

- For accurate and efficient ASD detection, this study presents an automated and lightweight framework named attention-based hybrid optimized residual memory network (AHRML).
- In order to streamline the identification process for the disabled, a new hybridized Arithmetic Harris Hawks Optimizer (AHHO) is used to reduce the dimensionality of features.
- In addition, attention-based residual term memory (ARTM), a clever deep learning algorithm, is created to accurately and less frequently identify ASD from the given data.
- Using the well-known ASD dataset that was taken from the UCI repository, the suggested AHRML model's findings were validated and tested using a variety of parameters.

The remainder of the components of this work are split into the following sections: a thorough review of the literature is provided in the Related Works section to look at

several approaches currently in use for ASD classification and identification. It also talks about the issues and difficulties that traditional approaches encounter. The Proposed Methodology section presents a concise synopsis of the suggested methodology, complete with a block diagram, mathematical explanations, and algorithms. In the Results and Discussion section, specific public ASD datasets and assessment metrics are used to validate the performance outcomes and results of the suggested model. In the Conclusion section, the entire study is finally summarized together with the results, conclusions, and next steps.

RELATED WORKS

This section explores a number of current approaches to the identification and categorization of ASDs. It also evaluates the benefits, issues, and difficulties that traditional methods present in light of their performance results.

Khodatars et al. (2021) did a thorough analysis of the literature to look at several deep learning techniques for ASD identification. The study's authors accurately determined the condition by using neuroimaging technologies. This study presents a comprehensive review of various deep learning strategies, encompassing the following techniques: autoencoders (AEs), convolutional neural networks (CNNs), recurrent neural networks, generative adversarial networks, deep belief networks, and deep Boltzmann machines. The results of this study show that deep learning approaches are effective in detecting ASD because of their strong capacity to handle large datasets with guaranteed accuracy. Sherkatghanad et al. (2020) developed an efficient and automated ASD detection and recognition system using the CNN technique. In this case, the layered architecture is created to carefully examine the inputs in order to identify the ASD with accuracy. Nevertheless, the recommended method takes longer for the validation and training phases, which has a big impact on the system's overall efficacy. Ke et al. (2020) conducted a thorough evaluation of the literature to evaluate the efficacy and performance of 14 different learning approaches used to diagnose ASD. Deep learning approaches have been adopted in a number of studies recently for the identification of mental disorders in the literature. This is because deep learning techniques have proven to be more effective in complicated pattern analysis and disease diagnosis. Preethi et al. (2022) used a two-dimensional CNN technique to identify ASDs efficiently by examining the input pattern analysis. This work uses the AdaDelta optimization methodology to enhance the detection performance of disorder detection. Qiang et al. (2023) developed a hierarchical brain network method to efficiently identify and recognize ASD. The primary focus of their study was on examining the input feature analysis in order to minimize the mistake rate and forecast disorder. Liu et al. (2021) applied a multi-regional data-driven attention learning strategy for the identification of ASD. Here, the AE and long short-term memory (LSTM) deep learning algorithms are combined to provide the recommended learning strategy. This study has made use of the Autism Brain Image Data Exchange (ABIDE) database to examine the results of

the proposed model. Yang et al. (2020) utilized a deep neural network (DNN) classification approach for an effective identification and classification of ASD. The classification of data using both voxels and connectivity functions has been effectively accomplished through the use of DNNs. In order to get enough examples for the DNN study, the multi-layer perceptron is applied with four distinct configurations leveraging the ABIDE repository.

Sewani and Kashef (2020) utilized an AE-based deep learning technique for an effective prediction of autism. AEs are essential for extracting low-dimensional features, which may then be fed into a neural network or learning models for classification problems. This architecture uses nonlinear optimization to reduce dimensionality. With fewer neurons in the hidden layer than in the input, the suggested approach concentrated on exploiting AE to extract pertinent data from the input layer. Zhang et al. (2022) implemented a variational AE model combined with a special feature selection method to efficiently identify and categorize ASD. In order to swap out the primary tanh function with a modified one, this work constructed a pipeline and used a threshold-shifting strategy, which might further improve the efficiency of classification. Furthermore, two restrictions that aid in training models with increased specificity or sensitivity have been applied. This strategy may be a useful supplementary technique for early diagnosis of ASD, given its exceptional classification performance on the heterogeneous dataset. This implies that this technique improves the state-of-the-art approaches even further. Subah et al. (2021) used a DNN to identify ASD from the ABIDE dataset as accurately as possible. Here, the most important characteristics from the provided photos have been extracted using the time series feature extraction model. Several machine learning methods have been used in earlier research to predict ASD, in which some of the most well-liked approaches are discussed below (Sujatha et al., 2021; Hasan et al., 2022).

Decision tree

One member of the supervised learning algorithm group is this algorithm. Regression and classification issues can be addressed through involvement. By examining fundamental decision rules (training data) drawn from previous data, a decision tree (DT) is used to construct a model for training that may be applied in predicting the type or amount of the variable being targeted. This algorithm has the benefit of requiring less work to set up the data during preliminary processing; however, it also has the downside of becoming unstable if a small change in any component of the data triggers a significant change in the process of decision-making stump.

Random forest

In order to create the class and facilitate learning, random forests (RFs) generate a huge number of DTs ahead of time. Overfitting the training set is optimal for DTs. The usual bootstrap approach of pooling is applied to tree learners by

the RF modeling approach. Bagging often selects a random sample in place of a training set and compares trees with samples of this kind.

Support vector machine

It is a linear algorithm that handles classification and regression issues. Support vector machine (SVM) has a straightforward concept: the data are divided into groups by the algorithm, which produces a line or hyper-plane. This algorithm's effective memory management is one of the advantages it offers, although large datasets are not a good fit for it. Moreover, the SVM is thought to be a classifier that maximizes the difference between specific important variables. As the foundation of this approach is optimizing the distance among the support vectors that produce the optimum hyper-plane, it is more likely to produce more effective data separation. But, it is striving to cross a threshold that raises the margin, necessitating the application of an optimizing process and requires that every point be on the appropriate segment of the hyper-plane.

Naive Bayes

This statistical classification technique is based on the Bayes theorem. This is one of the simplest supervised learning algorithms and is fast, precise, and trustworthy. On massive datasets, naive Bayes classifiers offer great speed as well as precision. One of the most potent and adaptable learning algorithms for data analysis and algorithmic learning, it has a high prediction efficiency. In a controlled setting, it is successfully trainable. This can be used with different attributes for small and large training data collection, and it often contains very little training data. This sorting algorithm has the knack of handling missing values.

The majority of earlier works (Bala et al., 2022; Farooq et al., 2023) target to establish an efficient intelligence-based learning methodology for detecting ASD from available data, according to the literature evaluation and analysis. However, the main issues with conventional techniques are their lower detection rate, longer prediction times, and mathematical modeling complexity. There is a lack of research on deep learning models for longitudinal monitoring and progress tracking of individuals with ASD over time. Developing models that can track changes in behavior and symptom severity longitudinally could provide valuable insights for personalized intervention planning and treatment evaluation. Therefore, the goal of the proposed research project is to put into practice a novel, lightweight AI-based detection framework for the precise identification and prediction of ASD based on the provided medical data.

PROPOSED METHODOLOGY

A detailed explanation of the proposed ASD detection methodology, including an overview, a block diagram, and

algorithmic details, is given in this section. Typically, the ASD is a chronic neurological disease. In order to provide proper medical care, ASD requires early, fast, precise, and successful detection. As an alternative to the traditional approaches, which focus on a quicker and cheaper patient diagnosis, data mining techniques have been used to identify the injured person. The traditional processes that have been developed and carried out in the past have been costly and laborious. Although much research in this area has been conducted with effectiveness in assessing ASD, these studies have not yet achieved outstanding performance in a time that is less expensive. The design and implementation of a unique and distinctive framework, known as AHRML for the identification and categorization of ASD from the provided medical data, constitutes the work's original contribution. Data preprocessing, feature optimization, ASD prediction, and performance estimation are the most significant aspects of the proposed AHRML system, which is presented in Figure 1. For system development and validation, the public open-source dataset is taken into consideration in this work initially. Following the acquisition of the data, typical preprocessing steps—which include attribute balancing, scaling, outlier removal, and missing value imputation—are used to produce normalized data. As a consequence, the previously processed information is used to determine the crucial and required features, which reduces the training and validation time and the computational load of classification. This is accomplished by implementing the novel hybrid optimization technique, AHHO. Subsequently, the acquired features are passed to an intelligence-based classifier, ARTM, which uses processing speed and detection rate to reliably identify ASD. Lastly, a thorough simulation and performance analysis are done to look at the results of the suggested AHRML system.

The integration of residual connections in a hybrid optimized residual memory network enhances interpretability by establishing unambiguous routes for the transmission of information inside the network. Furthermore, the use of memory processes may improve the comprehensibility of the model by allowing it to store and recall pertinent information from previous observations. This facilitates doctors' understanding and confidence in the model's predictions. Prior studies have shown that the hybrid optimized residual memory network is very successful in a range of machine learning tasks, such as medical diagnosis and prediction. The hybrid optimized residual memory network utilizes its hybrid architecture and optimization methodologies to produce exceptional performance in automated ASD diagnosis, beyond the capabilities of traditional deep learning models. This enhances the accuracy and reliability of diagnostic outcomes.

Dataset characteristics

The dataset comprises several aspects including demographic information (such as age and gender), behavioral factors (including social skills and communication abilities), and diagnostic information pertaining to ASD. The dataset has a certain quantity of instances or samples, with each

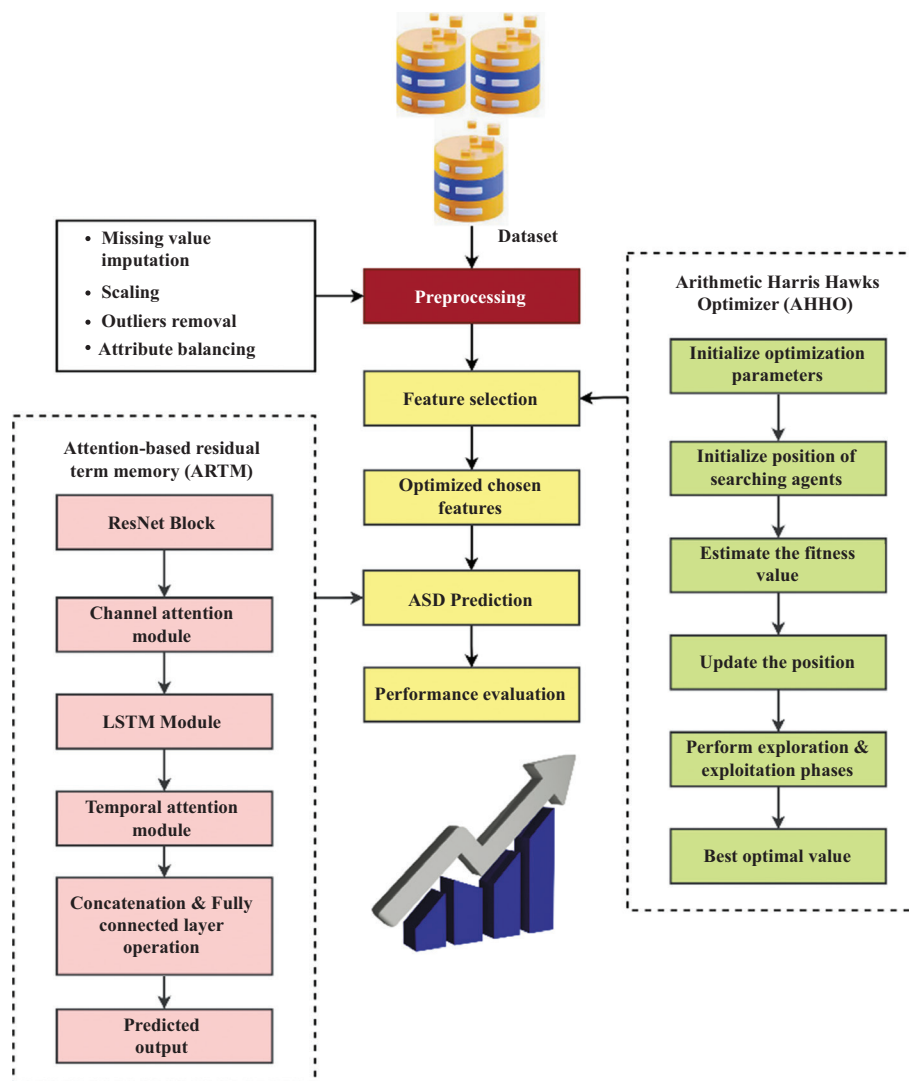


Figure 1: Overview of the proposed AHRML model. Abbreviations: AHRML, attention-based hybrid optimized residual memory network; ASD, autism spectrum disorder; LSTM, long short-term memory; ResNet, residual network.

instance representing a child. The target variable frequently denotes the existence or nonexistence of ASD, functioning as the label for classification tasks. The dataset is often used for the purpose of researching and developing machine learning models pertaining to the diagnosis, prediction, and categorization of autism. Preprocessing stages include several tasks such as managing missing data, encoding categorical variables, and scaling numerical characteristics, among other tasks.

AHHO for feature selection

At this point, the most essential and useful information qualities have been selected from the provided data using the cutting-edge hybrid optimization technique, AHHO. The ASD detection system's overall processing efficiency and speed are greatly increased by using this optimizer. The development of this hybridized model involves the integration of two distinct and widely used optimization techniques, namely the Harris hawks optimization (HHO) algorithm and the arithmetic algorithm (AA). As a result,

many optimization strategies for feature selection are being explored in the literature.

HHO is known for its ability to perform global optimization by leveraging the collaborative behavior of Harris hawks. In the context of dimensionality reduction, global optimization can help identify a low-dimensional embedding that captures the most informative features of the dataset while minimizing information loss or distortion. Many dimensionality reduction algorithms involve nonconvex optimization problems, where traditional optimization techniques may struggle to find globally optimal solutions. HHO's capacity to escape local optima and explore diverse regions of the search space can be advantageous for optimizing nonconvex objective functions commonly encountered in dimensionality reduction. The key advantages of the AA and HHO algorithms, in comparison to others, are their higher searching efficiency, low number of iterations needed to locate the optimal value, and quick solution determination time. Therefore, in order to develop a novel and distinctive optimizer for feature selection, the proposed work combines these two models. The architecture is composed of a primary layer at the top of the hierarchy that contains a number of HHO agents for

searching and a supplementary layer at the bottom that comprises subgroups with a number of arithmetic optimization algorithm (AOA) populations. The procedure of upgrading the search agents' locations is started by the AOA operation. It is essential to make sure that the locations of every searching agent in the highest layer are synchronized with the most effective solution found by the group that corresponds in the bottom layer in order to locate the best possible solution. A solution that is even more efficient can be achieved by developing new equations for each of the exploration and exploitation stages. This method greatly enhances the outcome and enables an in-depth examination of the problem at hand. The system takes its information from how Harris hawks hunt, making use of their phenomenal ability to locate and identify the prey with accuracy. Sometimes, though, a potential victim might not seem easily identifiable, so the hawks have to stay observant and keep an eye on the target area for a while before they may spot their next supper. In order to evaluate the most suitable option with the expected outcome, the system expertly utilizes Harris hawks to symbolize multiple options. By positioning the hawks in strategic positions and employing a variety of techniques, the algorithm accurately mimics the hawks' movements until the intended outcome is obtained. Each tactic is applied with equal care and attention to guarantee the greatest efficiency. At first, the exploration strategy of AHHO is executed by updating the location of searching agents in the topmost layer, which is mathematically represented as shown below:

$$\delta_{j+1}^i = \begin{cases} h_r - \alpha | h_r - 2\beta(h_j \div (mop + \varepsilon) \times ((u_j - l_j) \times m + l_j) |, \\ \quad \mathfrak{R} \geq 0.5 \ \& \ x \geq 0.5 \\ h_r - \alpha | h_r - 2\beta(h_j \div mop \times ((u_j - l_j) \times m + l_j) |, \\ \quad \mathfrak{R} \geq 0.5 \ \& \ x < 0.5 \\ h_b - h_m - \gamma[l_j + \vartheta], \ \mathfrak{R} < 0.5 \end{cases} \quad (1)$$

where δ_{j+1}^i indicates the location of searching agents in the topmost layer, j indicates the present iteration, α , β , γ , ϑ , x , \mathfrak{R} are the random numbers in the range of 0 to 1, h_r denotes the random hawk, h_b defines the best location of the hawk, h_m indicates the average mean of the hawk, l_j , u_j are the lower and upper bounds respectively, and mop represents the math optimizer probability coefficient. It is estimated as shown in the following equations:

$$mop(j) = 1 - \left(\frac{k}{K}\right)^{1/h} \quad (2)$$

$$h_m(j) = \frac{1}{\aleph} \sum_{j=1}^{\aleph} h_i(j) \quad (3)$$

where h is the parameter that determines the precision of exploitation, k and K are the current and maximum number of iterations, \aleph defines the total number of hawks in the population, and $h_i(j)$ denotes the position of the hawk. After that, the transition is performed that balances the modes of exploration and exploitation according to the energy level of the prey, which is computed based on the following equation:

$$\mathcal{E} = 2\mathcal{E}_0 \left(1 - \frac{k}{K}\right), \quad k = \{1, 2, 3 \dots K\} \quad (4)$$

where \mathcal{E} is the energy level, and \mathcal{E}_0 denotes the initial energy. Harris hawks use an unanticipated swoop technique in the attacking phase to seize their prey, which is detected in the earlier stage. The rabbit has enough energy to escape by making erratic and cunning leaps during $|\mathcal{E}| \geq 0.5$. Moreover, the rabbit grows more and more exhausted as the Harris hawks close in on it, and then they ambush attack, which is illustrated by the following equations:

$$\delta_{j+1}^i = \begin{cases} h_b - [\mathfrak{f}_G] - \mathcal{E}|2(1 - \beta)h_b - [\mathfrak{f}_G]|, \ x < 0.5 \\ h_b - [\mathfrak{f}_P] - \mathcal{E}|2(1 - \beta)h_b - [\mathfrak{f}_P]|, \ x \geq 0.5 \end{cases} \quad (5)$$

$$\mathfrak{f}_G = h_j - mop \times ((u_j - l_j) \times m + l_j) \quad (6)$$

$$\mathfrak{f}_P = h_j + mop \times ((u_j - l_j) \times m + l_j) \quad (7)$$

The hybrid algorithm makes use of the concept of incorporating the Levy flight pattern in order to statistically represent the leapfrog motions of the predator and the escape patterns of the prey. It mimics the unpredictable, abrupt, and swift dives that hawks make when they approach their intended prey, as well as the circular movements of the prey—especially rabbits—during their flight. Then, the position update is performed based on the following equations:

$$\delta_{j+1}^i = \begin{cases} Q \text{ if } \mathfrak{f}(Q) < \mathfrak{f}(h_j) \text{ and} \\ \quad h_j = \begin{cases} \mathfrak{f}_G, \ x < 0.5 \\ \mathfrak{f}_P, \ x \geq 0.5 \end{cases} \\ S \text{ if } \mathfrak{f}(S) < \mathfrak{f}(h_j) \text{ and} \\ \quad h_j = \begin{cases} \mathfrak{f}_G, \ x < 0.5 \\ \mathfrak{f}_P, \ x \geq 0.5 \end{cases} \end{cases} \quad (8)$$

$$Q = S + w \times LF(d) \quad (9)$$

$$S = h_b - \mathcal{E}|V_{h_b - h_m}|; \ V = 2 - 2\beta \quad (10)$$

The Harris hawks will encircle the rabbit fiercely and make surprise hits in an effort to capture it. This phase of the resilient attack is described in the following equations:

$$\delta_{j+1}^i = \begin{cases} h_b - \mathcal{E} | h_b - [h_j - mop \times (u_j - l_j) \times \\ \quad m + l_j] |, \ x < 0.5 \\ h_b - \mathcal{E} | h_b - [h_j - mop \times (u_j - l_j) \times \\ \quad m + l_j] |, \ x \geq 0.5 \end{cases} \quad (11)$$

$$\delta_{j+1}^i = \begin{cases} Q \text{ if } \mathfrak{f}(Q) < \mathfrak{f}(h_j) \text{ and} \\ \quad h_j = \begin{cases} \mathfrak{f}_G, \ x < 0.5 \\ \mathfrak{f}_P, \ x \geq 0.5 \end{cases} \\ S \text{ if } \mathfrak{f}(S) < \mathfrak{f}(h_j) \text{ and} \\ \quad h_j = \begin{cases} \mathfrak{f}_G, \ x < 0.5 \\ \mathfrak{f}_P, \ x \geq 0.5 \end{cases} \end{cases} \quad (12)$$

$$S = h_b - \mathcal{E}|V_{h_b - h_m}|; \ V = 2 - 2\beta \quad (13)$$

This process culminates in the output of the best optimal value, which is utilized to select the most important and necessary characteristics from the normalized dataset. The suggested model's overall ASD detection performance and

results are significantly enhanced by using this hybridized optimizer.

ARTM classification

In this stage, a novel ARTM classification strategy is used to predict ASD using the features that were acquired in the previous stage. For the diagnosis of ASD, a range of learning algorithms are typically created in earlier literature publications. Deep learning techniques work well with a high disease diagnosis rate, making them superior to machine learning algorithms. However, the typical issues of increased error rate, high time complexity, and complexity in data training and validation plague standard deep learning methods. Therefore, the goal of the presented study is to design a deep learning system for ASD diagnosis that is both efficient and lightweight. The attention block, residual network (ResNet) block, and LSTM block are combined to develop the proposed ARTM approach. The human attention process is modeled after the biological mechanisms that focus promptly on the most unique and significant aspects of the problem at hand. After training, typical neural networks frequently acquire stable weight matrices, which do not alter even if the network is fed entirely new inputs. As a result, classic neural network models find it challenging to adjust to rapidly changing settings.

ARTM allows the model to selectively attend to relevant features and ignore irrelevant ones during different stages of processing. By dynamically adjusting the attention weights, the model can focus on important features while suppressing noise or irrelevant information, leading to more effective feature representation and improved discrimination between classes. The memory component in ARTM enables the model to maintain information over multiple time steps or layers. This is particularly beneficial for tasks involving sequential data or long-range dependencies, such as natural language processing or time series analysis. By preserving the context and temporal relationships, ARTM can capture complex patterns and dependencies that may be crucial for accurate prediction or classification.

A neural network could create self-modifying weights and focus on the most important parts of the input through the incorporation of the attention mechanism. The idea put forward makes use of the temporal attention structure and the channel attention mechanism, which are two of the attention mechanism approaches that have been put forth at this point. By combining spatial and channel-wise data through convolution processes, a ResNet can obtain useful features from layered data. It has been shown that by clearly modeling the interrelated relationships between the channels of attribute maps, the efficiency of a network can be significantly enhanced. Consequently, by assigning each channel of feature maps a weight parameter, we incorporate the channel attention mechanism with ResNet in order to adapt and adjust channel-wise information. The backpropagation method can still be used to train such weight factors. These procedures are made possible by the inclusion of the squeeze-and-excitation blocks. The architecture model of the proposed ARTM model is shown in Figure 2.

Channel-wise information gets generated by means of global average pooling in the squeezing phase. Two fully connected (FC) layers are used in the following excitation process to identify channel interactions and restrict cost. Specifically, one FC layer with a reduction ratio lowers the inputs' scale, and the additional FC layer eventually recovers the channel dimension to its initial value. Implementing a channel-wise convolution among stream weights and feature maps produced by the ResNet yields the block's final outcome. Putting the last hidden state as the output continues to be the default in a standard LSTM prediction process. Nevertheless, as the input data sequence grows longer, the prediction accuracy will drop. In order to address this issue, we incorporate a temporal attention layer into the final LSTM layer, enabling the consideration of all concealed states. The last time step's concealed state is designated as a typical state since it has historical data in it. After that, it is compared with all hidden states in order to determine their respective scores. Higher weights will be allocated to the hidden state with the greatest score. As a result, the system has the potential to effectively highlight the informative elements of various inputs and can adjust to input sequences that are both short- and long-term. The temporal attention operations are mathematically illustrated below:

$$\mathbb{W}(\mathbf{h}_s, \mathbf{h}_S) = x^S \tanh([\mathbf{h}_s; \mathbf{h}_S]) \quad (14)$$

$$\varpi_s = \frac{\exp(\mathbb{W}(\mathbf{h}_s, \mathbf{h}_S))}{\sum_{m=1}^S \exp(\mathbb{W}(\mathbf{h}_s, \mathbf{h}_S))} \quad (15)$$

$$\mathbf{h}_o = \sum_{m=1}^S \varpi_s \mathbf{h}_s \quad (16)$$

where \mathbf{h}_s indicates the hidden state at time step s , \mathbf{h}_S denotes the final hidden state at time step S , ϖ_s represents the attention weight value, x^S defines the learnable matrix, and \mathbf{h}_o is the final output of the temporal attention layer. Then, the final prediction result is obtained as shown in the following equation:

$$\mathcal{O} = \sum_{i=1}^I \varpi_i k_i + \varpi_0 \quad (17)$$

where σ denotes the predicted result, ϖ_i defines the coefficient variable of i th transformed variable, and ϖ_0 represents the optional bias term. This hybrid learning model improves performance outcomes by accurately predicting ASD in the suggested framework.

RESULTS AND DISCUSSION

This section uses the widely used ASD dataset and assessment metrics to validate the effectiveness and results of the proposed AHRML system. A Windows 10 PC with specifications including a CPU speed of 2.9 GHz core i7, GPU Intel 620, RAM of 12 GB, and a spare hard drive space of a minimum of 5 GB has been utilized for the experiments in order to successfully run and validate the proposed configuration. Although the prevalence of ASD is rising globally, there are very few publicly accessible statistics dedicated to

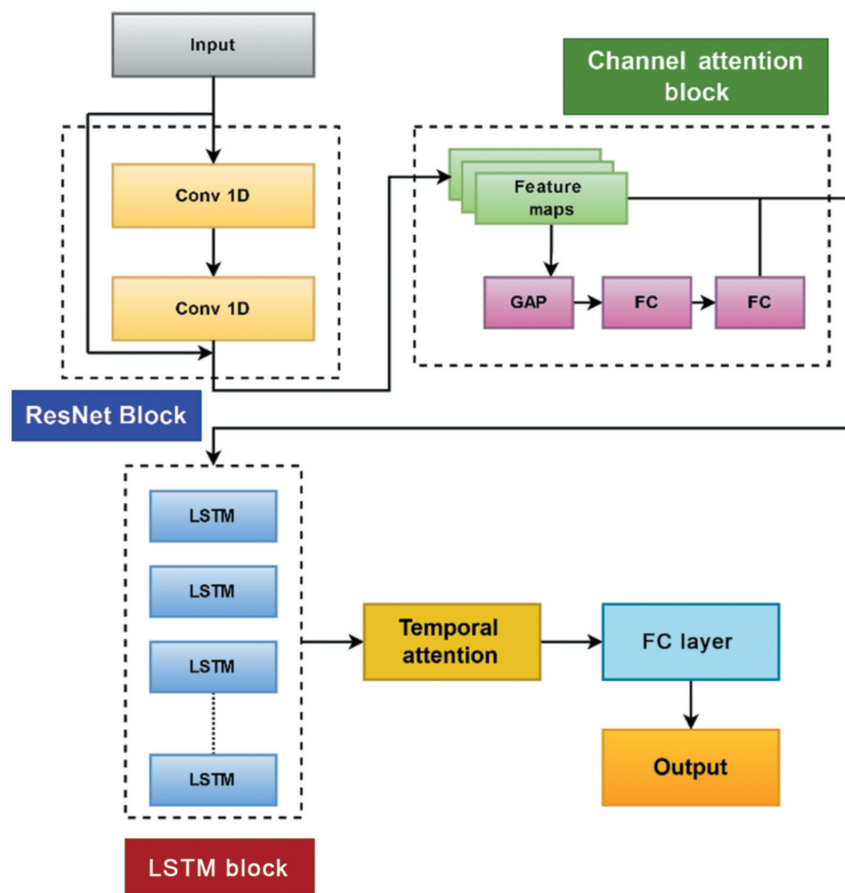


Figure 2: Architecture of the ARTM model. Abbreviations: ARTM, attention-based residual term memory; FC, fully connected; GAP, global average pooling; LSTM, long short-term memory; ResNet, residual network.

the disorder's research. There are currently too few clinical screening datasets for autism, and the majority of resources are genetics-focused. The ASD in the adult and children's datasets, which comprised 704 instances and 21 attributes in the adult dataset and 292 instances and 21 attributes in the children's dataset, was taken from a public UCI repository (Hossain et al., 2021; Sujatha et al., 2021) and employed in the present study. This children's dataset contains 292 observations, with ages ranging from 4 to 11. Working with the ASD dataset presents many challenges due to the multitude of respondents, and the dataset details are given in Table 1.

Accuracy in predictive modeling refers to the percentage of total data points that are successfully estimated. It shows the proportion of accurate forecasts in a simulated set of data. It is crucial since AI learns on its own, making it challenging to assess the accuracy of supplied data:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (18)$$

The ratio of positive events to all genuine positive examples is known as precision. To put it plainly, precision tells us how true a model is when it says it is correct. In order to compute the same value, apply the following equation:

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (19)$$

Recall, which is calculated as the ratio of true positives to the total number of real positive occurrences, quantifies the percentage of actual positive events that a system or model properly detected. The outcome is computed using the formula below.

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (20)$$

The F1-score is the combined precision and recall metric, with 1 denoting the best possible outcome and 0 the worst. An accurate disease detection system that strikes a balance

Table 1: ASD dataset description.

S. No	Data information	Type of attributes	No. of samples	Yes (Y) or no (N)
1	ASD screening for children	Binary, categorical, and continuous	292	Y—141; N—151
2	ASD screening for adult	Binary, categorical, and continuous	704	Y—189; N—515;
3	ASD screening for combined information	Binary, categorical, and continuous	996	Y—330; N—666

Abbreviation: ASD, autism spectrum disorder.

Table 2: Performance comparative study for children’s ASD dataset.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
NB	94.25	90.91	97.56	94.12
KNN	86.21	87.18	82.93	85
SVM	98	98.5	99	99
RF	96.55	93.18	99	96.47
DT	87.36	87.50	85.37	86.42
XGB	97.70	97.56	97.56	97.5
LR	98	98.5	98	98
ANN	98.85	97.62	98.8	98.8
Proposed	99	98.9	99	99.2

Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine; LR, linear regression; XGB, XGBoost.

Table 3: Comparison based on AUC, kappa, and log loss value for children’s ASD dataset.

Methods	AUC	Kappa	Log loss value
NB	0.94	0.88	2.071
KNN	0.86	0.72	4.972
SVM	0.85	0.84	2.698
RF	0.96	0.93	1.243
DT	0.87	0.74	4.557
XGB	0.97	0.91	0.829
LR	0.88	0.91	0.878
ANN	0.98	0.97	0.414
Proposed	0.99	0.99	0.152

Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; AUC, area under the curve; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

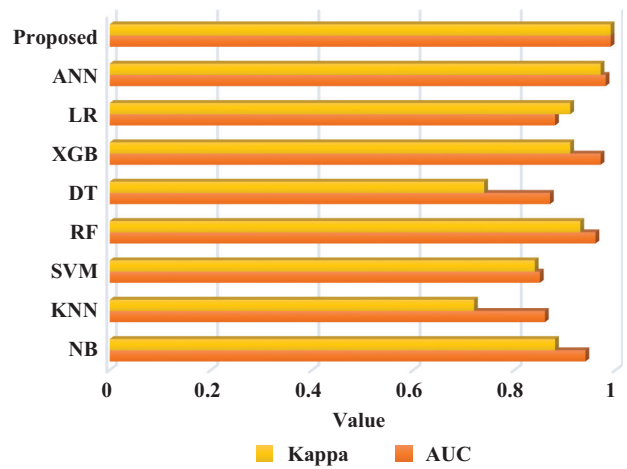


Figure 4: AUC and kappa analysis for children’s ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; AUC, area under the curve; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

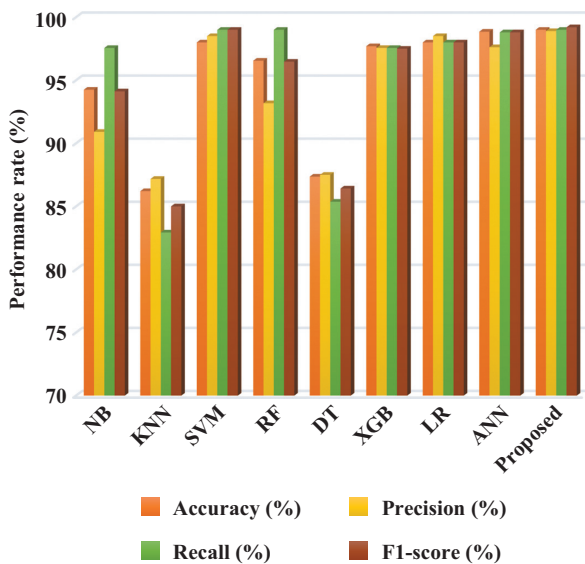


Figure 3: Overall performance analysis using children’s ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

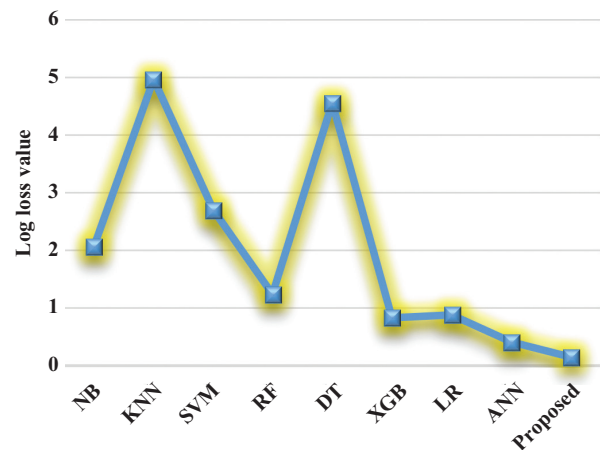


Figure 5: Log loss value for children’s ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

Table 4: Performance comparative study for the adult ASD dataset.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
NB	96.19	95	91.94	93.44
KNN	89.05	84.21	77.42	80.67
SVM	96.19	93.55	93.55	93.55
RF	95.71	94.92	90.32	92.56
DT	87.14	84.31	69.35	76.11
XGB	96.19	95	91.94	93.44
LR	97.14	96.67	93.55	95.05
ANN	96.67	91.04	98.39	94.57
Proposed	99	98.8	98.9	99.1

Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

between false positives and false negatives is indicated by an increased F1-score.

$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% \quad (21)$$

Table 5: Comparison based on AUC, kappa, and log loss value for the adult ASD dataset.

Methods	AUC	Kappa	Log loss value
NB	0.94	0.90	1.373
KNN	0.85	0.73	3.948
SVM	0.95	0.90	1.373
RF	0.94	0.89	1.545
DT	0.81	0.80	4.634
XGB	0.94	0.90	1.373
LR	0.96	0.93	1.029
ANN	0.97	0.92	1.201
Proposed	0.99	0.989	0.211

Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; AUC, area under the curve; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

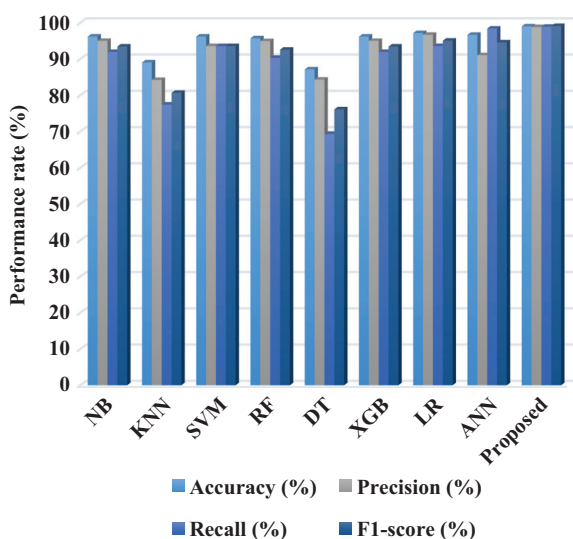


Figure 6: Overall performance analysis using the adult ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

where \mathcal{TP} indicates the true positives, \mathcal{TN} indicates true negatives, \mathcal{FP} represents the false positives, and \mathcal{FN} represents the false negatives. Using a dataset of children with

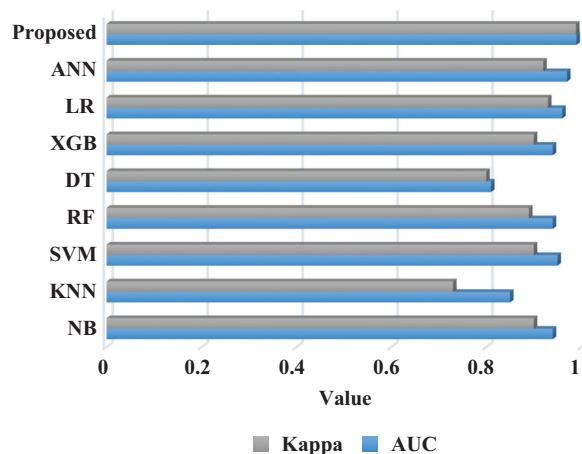


Figure 7: AUC and kappa analysis for the adult ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; AUC, area under the curve; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

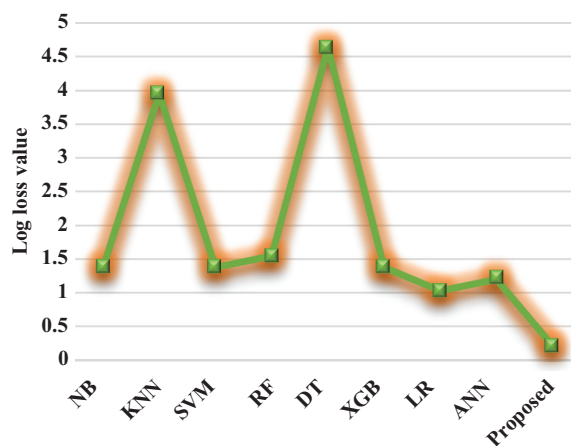


Figure 8: Log loss value for the adult ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

Table 6: Performance comparative study for the combined ASD dataset.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
NB	93.27	86.41	93.68	89.89
KNN	83.16	72.73	75.79	74.23
SVM	91.58	88.04	85.26	86.63
RF	93.27	92.13	86.32	89.13
DT	91.92	87.37	87.37	87.37
XGB	92.59	91.01	85.26	88.04
LR	91.58	91.67	81.05	86.03
ANN	94.28	89.80	92.63	91.19
Proposed	99.1	99	98.8	99

Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

Table 7: Comparison based on AUC, kappa, and log loss value for the combined ASD dataset.

Methods	AUC	Kappa	Log loss value
NB	0.93	0.84	2.42
KNN	0.81	0.61	6.068
SVM	0.89	0.80	3.033
RF	0.91	0.84	2.427
DT	0.90	0.81	2.913
XGB	0.90	0.82	2.669
LR	0.88	0.80	3.034
ANN	0.93	0.86	2.063
Proposed	0.99	0.99	0.154

Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; AUC, area under the curve; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

ASD, Tables 2 and 3 assess and contrast the overall effectiveness of the suggested learning techniques with the standard approaches. Next, corresponding graphical illustrations for each of them are shown in Figures 3-5. The following factors were taken into account for this analysis: F1-score, kappa coefficient, accuracy, precision, recall, log loss value, and area under the curve. This comparative analysis shows that, in comparison to the other current classification approaches, the suggested AHRML model performs well and offers better performance yields.

In the same way, the adult ASD dataset’s overall performance comparison study is carried out, as indicated in Tables 4 and 5. Next, their corresponding graphical representations are shown in Figures 6-8. These results also show that, in comparison to the other classification methods, the AHRML model yields better results for ASD detection.

Furthermore, as indicated in Tables 6 and 7, the performance outcomes are estimated and contrasted for the combined children + adult dataset. Figures 9-11 provide the relevant graphical depictions of them. When compared to all other current techniques, the comparison data show that the suggested AHRML produces an ASD prediction result with a lower log loss value because adopting hybrid AHHO is the primary means of achieving better categorization outcomes.

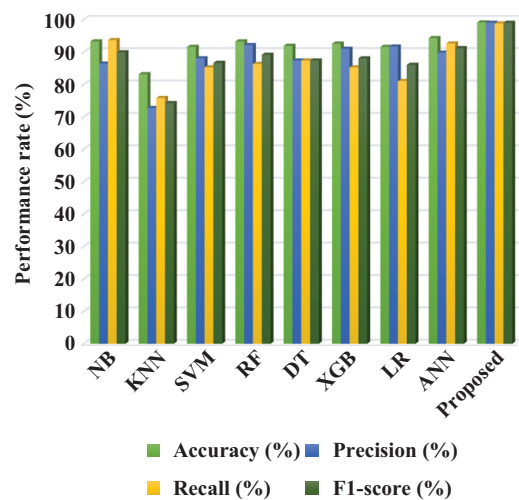


Figure 9: Overall performance analysis using the combined ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

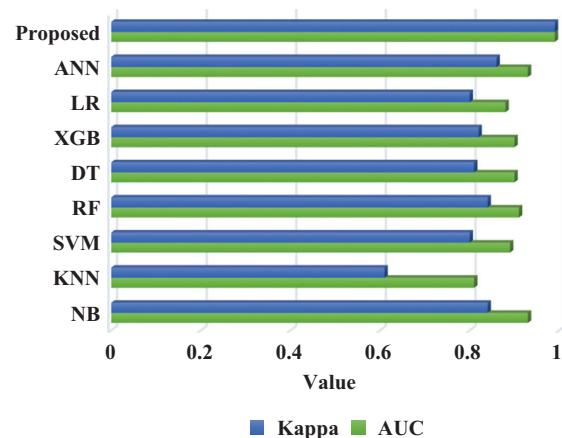


Figure 10: AUC and kappa analysis for the combined ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; AUC, area under the curve; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

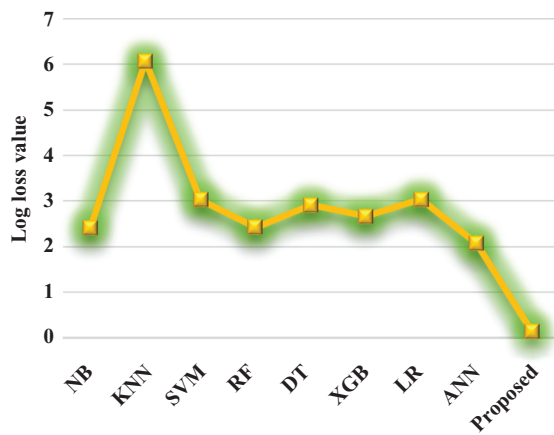


Figure 11: Log loss value for the combined ASD dataset. Abbreviations: ANN, artificial neural network; ASD, autism spectrum disorder; DT, decision tree; KNN, k-nearest neighbor; NB, naive Bayes; RF, random forest; SVM, support vector machine.

CONCLUSION

The study has proposed a model to identify and diagnose ASD using advanced deep learning methods. It addressed the existing limitations by presenting a model for detecting ASD symptoms over time. The ASD is typically a long-term neurological condition. ASD needs to be detected early, quickly, precisely, and successfully in order to receive the appropriate medical therapy. Rather than relying on more expensive and time-consuming traditional methods, data mining techniques have been employed to identify the damaged party. The conventional procedures that were created and used in the past have been expensive and time-consuming. While some studies in this field have been undertaken with success in evaluating ASD, these investigations have not yet reached exceptional results at a lower cost. The work's original contribution is the formulation and application of the distinctive and one-of-a-kind AHRML framework for the diagnosis and classification of ASD from supplied medical data. The most important components of the suggested AHRML system are feature optimization, performance estimates, ASD

prediction, and data preprocessing. In this work, the public open-source dataset is first considered for system development and validation. The normalized data are obtained via standard preprocessing procedures such as attribute balancing, scaling, outlier removal, and missing value imputation, which are applied after the data are acquired. Consequently, this lowers the training and validation time as well as the classification computing burden by using the preprocessed data to identify the important and necessary features. The AHHO, a novel hybrid optimization technique, is used to achieve this. The obtained attributes are then fed into an intelligence-based classifier called ARTM, which employs detection rate and processing speed to accurately identify ASD. Finally, a comprehensive performance analysis and simulation are conducted to examine the outcomes of the proposed AHRML system. When compared to all current approaches, the total analysis concludes that the proposed AHRML offers better detection performance and outcomes because of appropriate feature reduction, training, and testing operations. The results improve our comprehension of the intricate connections between behavioral characteristics and ASD, opening up possibilities for future study in uncovering the fundamental causes of the illness. Deep learning models for ASD diagnosis are useful tools for doctors, providing help in diagnosing and planning therapy. These models enhance diagnostic accuracy and consistency by using objective and standardized tests that work with clinical experience. The effectiveness of deep learning models in detecting ASD is highly dependent on the quality and amount of annotated data that are accessible for training purposes. The restricted availability of extensive, top-notch datasets could limit the ability of models to generalize and perform well.

ACKNOWLEDGMENTS

The authors extend their appreciation to the King Salman Center for Disability Research (funder ID: <http://dx.doi.org/10.13039/501100019345>) for funding this work through Research Group No. KSRG-2022-155.

REFERENCES

- Alvarez-Jimenez C., Múnica-Garzón N., Zuluaga M.A., Velasco N.F. and Romero E. (2020). Autism spectrum disorder characterization in children by capturing local-regional brain changes in MRI. *Med. Phys.*, 47, 119-131.
- Bahathiq R.A., Banjar H., Jarraya S.K., Bamaga A.K. and Almoallim R. (2024). Efficient diagnosis of autism spectrum disorder using optimized machine learning models based on structural MRI. *Appl. Sci.*, 14, 473.
- Bala M., Ali M.H., Satu M.S., Hasan K.F. and Moni M.A. (2022). Efficient machine learning models for early stage detection of autism spectrum disorder. *Algorithms*, 15, 166.
- Chen T., Chen Y., Yuan M., Gerstein M., Li T., Liang H., et al. (2020). The development of a practical artificial intelligence tool for diagnosing and evaluating autism spectrum disorder: multicenter study. *JMIR Med. Inform.*, 8, e15767.
- Das Biswas S., Chakraborty R. and Pramanik A. (2020). On prediction models for the detection of autism spectrum disorder. In: *Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2020*, Springer, Poland, pp. 359-371.
- Dhamale T.D., Bhandari S.U. and Harpale V.K. (2023). Fusion of features: a technique to improve autism spectrum disorder detection using brain MRI images. *Biomed. Pharmacol. J.*, 16, 2443-2455.
- Epalle T.M., Song Y., Liu Z. and Lu H. (2021). Multi-atlas classification of autism spectrum disorder with hinge loss trained deep architectures: ABIDE I results. *Appl. Soft Comput.*, 107, 107375.
- Farooq M.S., Tehseen R., Sabir M. and Atal Z. (2023). Detection of autism spectrum disorder (ASD) in children and adults using machine learning. *Sci. Rep.*, 13, 9605.
- Fu Y., Zhang J., Li Y., Shi J., Zou Y., Guo H., et al. (2021). A novel pipeline leveraging surface-based features of small subcortical structures

- to classify individuals with autism spectrum disorder. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry*, 104, 109989.
- Gao J., Xu Y., Li Y., Lu F. and Wang Z. (2024). Comprehensive exploration of multi-modal and multi-branch imaging markers for autism diagnosis and interpretation: insights from an advanced deep learning model. *Cereb. Cortex*, 34, bhad521.
- Hasan S.M., Uddin M.P., Al Mamun M., Sharif M.I., Ulhaq A. and Krishnamoorthy G. (2022). A machine learning framework for early-stage detection of autism spectrum disorders. *IEEE Access*, 11, 15038-15057.
- Hossain M.D., Kabir M.A., Anwar A. and Islam M.Z. (2021). Detecting autism spectrum disorder using machine learning techniques: an experimental analysis on toddler, child, adolescent and adult datasets. *Health Inf. Sci. Syst.*, 9, 1-13.
- Ke F., Choi S., Kang Y.H., Cheon K.-A. and Lee S.W. (2020). Exploring the structural and strategic bases of autism spectrum disorders with deep learning. *IEEE Access*, 8, 153341-153352.
- Khodatars M., Shoeibi A., Sadeghi D., Ghaasemi N., Jafari M., Moridian P., et al. (2021). Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: a review. *Comput. Biol. Med.*, 139, 104949.
- Kumar S.S., Selvakumar K. and Murugan V.S. (2024). Identification of autism spectrum disorder using modified convolutional neural network (MCNN) and feature selection techniques. *Int. J. Intell. Syst. Appl. Eng.*, 12, 678-691.
- Lamani M.R. and Julian Benadit P. (2023). An early detection of autism spectrum disorder using PDNN and ABIDE I&II dataset. In: *International Conference on Artificial Intelligence on Textile and Apparel*, Springer, pp. 295-310.
- Liu Y., Xu L., Yu J., Li J. and Yu X. (2021). Identification of autism spectrum disorder using multi-regional resting-state data through an attention learning approach. *Biomed. Signal Process. Control*, 69, 102833.
- Mohan P. and Paramasivam I. (2021). Feature reduction using SVM-RFE technique to detect autism spectrum disorder. *Evol. Intell.*, 14, 989-997.
- Nogay H.S. and Adeli H. (2024). Multiple classification of brain MRI autism spectrum disorder by age and gender using deep learning. *J. Med. Syst.*, 48, 15.
- Preethi S., Arun Prakash A., Ramyea R., Ramya S. and Ishwarya D. (2022). Classification of autism spectrum disorder using deep learning. In: *Intelligent Systems: Proceedings of ICMIB 2021*, Springer, pp. 247-255.
- Qiang N., Gao J., Dong Q., Li J., Zhang S., Liang H., et al. (2023). A deep learning method for autism spectrum disorder identification based on interactions of hierarchical brain networks. *Behav. Brain Res.*, 452, 114603.
- Saponaro S., Lizzi F., Serra G., Mainas F., Oliva P., Giuliano A., et al. (2024). Deep learning based joint fusion approach to exploit anatomical and functional brain information in autism spectrum disorders. *Brain Inform.*, 11, 1-13.
- Sewani H. and Kashef R. (2020). An autoencoder-based deep learning classifier for efficient diagnosis of autism. *Children*, 7, 182.
- Sherkatghanad Z., Akhondzadeh M., Salari S., Zomorodi-Moghadam M., Abdar M., Acharya U.R., et al. (2020). Automated detection of autism spectrum disorder using a convolutional neural network. *Front. Neurosci.*, 13, 1325.
- Subah F.Z., Deb K., Dhar P.K. and Koshiha T. (2021). A deep learning approach to predict autism spectrum disorder using multisite resting-state fMRI. *Appl. Sci.*, 11, 3636.
- Sujatha R., Aarthy S., Chatterjee J., Alaboudi A. and Jhanjhi N. (2021). A machine learning way to classify autism spectrum disorder. *Int. J. Emerg. Technol. Learn. (iJET)*, 16, 182-200.
- Wang D., Yang X. and Ding W. (2023). Autism spectrum disorder (ASD) classification with three types of correlations based on ABIDE I data. *Math. Found. Comput.* 10.1038/s41598-020-61213-w.
- Yang X., Schrader P.T. and Zhang N. (2020). A deep neural network study of the ABIDE repository on autism spectrum classification. *Int. J. Adv. Comput. Sci. Appl.*, 11. <http://dx.doi.org/10.14569/IJACSA.2020.0110401>.
- Yang X., Zhang N. and Schrader P. (2022). A study of brain networks for autism spectrum disorder classification using resting-state functional connectivity. *Mach. Learn. Appl.*, 8, 100290.
- Zhang F., Wei Y., Liu J., Wang Y., Xi W. and Pan Y. (2022). Identification of autism spectrum disorder based on a novel feature selection method and variational autoencoder. *Comput. Biol. Med.*, 148, 105854.