An Auto-encoded Warm Equilibrium Automated Learning (AE$_2$L) Model for Automatic Recognition and Classification of Autism Spectrum Disorder

Muhanna K. Al-Muhanna$^{1,*}$, Amani Ahmed Alghamdi$^2$, Bahauddeen Alfaesi$^{3,4}$$^*$, Mohammad Afzaal$^5$, Reema Al-Subaiee$^{3,4}$$^*$ and Rania Haddadi$^6$

$^1$Materials Science Research Institute, King Abdulaziz City for Science and Technology (KACST), Riyadh-11442, Saudi Arabia
$^2$Department of Biochemistry, College of Science, King Saud University, Riyadh 11451, Saudi Arabia
$^3$Stem Cells and Regenerative Medicine Unit, King Abdullah International Medical Research Center (KAIMRC), Riyadh 11481, Saudi Arabia
$^4$King Saud Bin Abdulaziz University for Health Sciences (KSAU-HS), Riyadh 11481, Saudi Arabia
$^5$Department of Chemistry, College of Science, King Saud University, Riyadh 11451, Saudi Arabia
$^6$Department of Zoology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

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

Received: February 29 2024; Revised: March 24 2024; Accepted: March 24 2024; Published Online: May 8 2024

ABSTRACT

Autism spectrum disorder (ASD) is a neurological condition characterized by difficulties with communication and socializing, and repetitive activities. If the underlying reason is hereditary, early detection is still important, and machine learning offers a fascinating way to identify the condition more rapidly and economically. However, the unique issues of higher computational costs, longer execution times, and lower effectiveness affect the traditional methods. The proposed project aims to create an automated artificial intelligence tool for ASD identification that combines several state-of-the-art mining techniques to deliver the best possible level of disease prediction accuracy. For accurate and effective ASD identification, this research suggests an automated and lightweight method dubbed the auto-encoded warm equilibrium automated learner. To speed up the handicap detection process, a unique warm optimized feature selection methodology is applied to minimize the dimensionality of attributes. In addition, auto-encoded term memory equilibrium learning, a powerful deep learning technique, is designed to accurately and less frequently detect ASD from the given data. Moreover, the classifier performs better when hyperparameters are tuned using the equilibrium optimization model. The results of the proposed AE$_2$L model have been tested and validated using a variety of parameters utilizing the well-known ASD dataset that was taken from the UCI repository.

KEYWORDS

autism spectrum disorder (ASD), deep learning, artificial intelligence (AI), neurodisabilities, mental healthcare, optimization, auto-encoded warm equilibrium automated learner (AE$_2$L)

INTRODUCTION

The prevalence of autism spectrum disorder (ASD) has increased more recently than it has ever been. Identifying autistic traits through screening tests is exceptionally costly and time-consuming (Choi et al., 2024; Koehler et al., 2024). A person with ASD will experience lifelong difficulties in communicating and interacting with others. As symptoms of autism typically manifest in the early stages of life, it is referred to as a “behavioral disease”. It can be diagnosed at any point in a person’s life. The ASD theory states that the issue arises during infancy and persists through teenage and maturity. It is difficult to screen identical items, and the classification needs to be done carefully (Ahmad et al., 2024). The problem of ASD has to do with how the human brain develops. In general, social involvement and interaction with others are difficult for someone with ASD. This typically has an impact on the life of an individual for the entire duration of their lifespan. Considering that this condition could have hereditary and environmental causes is interesting. The problem’s signs could appear as early as age three and last for the rest of a person’s life. Although the patient’s condition cannot be fully treated, its effects may be temporarily mitigated if the symptoms are identified early on (Hajjej et al., 2024; Pan and Foroughi, 2024). Machine learning (ML) is quite effective in diagnosing this problem. The majority of
investigators (Mohamed et al., 2023) have used ML techniques—among which deep learning (DL) methods are widely used—to define individuals as well as typical controls. The diagnosing statistics guide indicates that autism and intellectual impairment develop together. The normal and ASD-affected brain images are shown in Figure 1.

On the other hand, the International Classification of Diseases offers a comprehensive framework for differentiating between ASDs and those that are not; it also considers past data regarding the loss of prior skills during the diagnostic process (Trudel et al., 2023; Menaka et al., 2024). The most challenging part of recognizing ASD is that there is no identifiable pathognomonic characteristic; instead, all symptoms center on altering the behavior profile of the affected person, which varies depending on the intensity and age. The method of determining ASD (Shen et al., 2023) is complex and differs from person to person. As people get older, their symptoms also vary. Therefore, early identification and treatment could significantly improve the prognosis, which could stop the brain from adapting to a negative environment. Early intervention can effectively improve children’s linguistic and cognitive abilities before behavioral issues arise, as previous investigations have demonstrated that children’s brain resilience turns with development. Consequently, it is crucial to identify ASD early on. The majority of ASD detection techniques available today, regrettably, generate human observation-based diagnoses that are laborious and challenging to use. For instance, experts deliver the Modified Checklist for Autism in Toddlers, a common questionnaire for parents, in closely monitored clinical settings. It typically takes hours to complete. An intelligent automatic detection tool is therefore required to increase operability and detection efficiency. The early identification of ASD can be challenging, and parents often become concerned about their children only after they have been watched closely, delaying the process. It is a difficult task that necessitates rigorous analysis to identify the most notable traits from a large dataset (Archana et al., 2023; Halim et al., 2023). Handling missing values in attributes can also be a challenge when handling data processing operations. The quantity and accuracy of the information taken into account will largely determine how ML is applied in the steps that follow. It is necessary to customize automation depending on the analytical viewpoint.

ASD manifests differently across individuals, with diverse phenotypic expressions and genetic predispositions. DL models are capable of capturing subtle patterns and variations in ASD-related data, enabling the detection of nuanced differences that may not be apparent to human observers. The proliferation of digital healthcare data, including electronic health records, medical imaging, and wearable sensor data, has provided rich sources of information for ASD detection. DL methods, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), can efficiently process large-scale datasets and scale to accommodate diverse data modalities, making them well-suited for ASD detection tasks. DL algorithms can automatically learn hierarchical representations of data, allowing them to adaptively extract relevant features from raw input. This is particularly advantageous for ASD detection, where the underlying patterns may be complex and multi-faceted, spanning spatial, temporal, and semantic domains. DL-driven methods offer the potential for automated and efficient ASD detection, reducing the burden on clinicians and healthcare professionals. By streamlining the diagnostic process and providing objective decision support, these methods can facilitate timely intervention and improve patient outcomes.

The current diagnostic methods often rely heavily on subjective assessments by clinicians, leading to variability and delays in diagnosis. The development of automated systems for the recognition and classification of ASD has the potential to revolutionize the diagnostic process, enabling early identification and intervention. Traditional machine-learning approaches have shown promise in this domain, but there remains a need for more sophisticated models capable of capturing complex patterns and subtle nuances in ASD-related data. ASD-related data encompass a wide range of modalities, including behavioral assessments, neuroimaging data, genetic information, and clinical notes. Integrating heterogeneous data sources and extracting relevant features pose significant challenges. As automated systems become increasingly complex, ensuring the interpretability of model predictions is essential for gaining insights into the underlying mechanisms of ASD and fostering trust among clinicians and stakeholders.

ML is being used these days to identify a number of diseases, such as ASD and depression. The principal goals of utilizing ML techniques are to expedite access to healthcare services by increasing precision in diagnosis and reducing the diagnostic time. The diagnosis method of a particular instance can be considered a task related to classification in ML as it entails determining the appropriate class (ASD, No-ASD) according to the properties of the input case (Mumenin et al., 2023). In previous studies, researchers employed a variety of classification strategies to achieve higher accuracy in identifying ASD cases using different datasets. However, they face significant challenges related to an increased processing load, a longer processing time, low efficiency, and a high rate of incorrect predictions (Lawan et al., 2023; Ranjana, 2023; Swedha and Devendran, 2023).

Therefore, the proposed study aims to develop a novel hybrid DL-tied optimization model for precise diagnosis and prediction of ASD. The following is a summary of this paper’s primary contributions:

- This research proposes an automated and lightweight system called auto-encoded warm equilibrium
• A novel warm optimized feature selection (WOFS) methodology is used to minimize the dimensionality of characteristics in order to expedite the handicap identification procedure.
• Moreover, a sophisticated DL method called auto-encoded term memory equilibrium learning (AET-MEL) is developed to reliably and less frequently identify ASD from the provided data.
• Furthermore, hyperparameter tuning with the equilibrium optimization (EO) model enhances the classifier’s performance.
• A range of parameters have been employed to validate and test the suggested AE\textsubscript{L} model’s results using the popular ASD dataset that was extracted from the UCI repository.

The following sections constitute the remaining elements of this work: In the Related Studies section, a comprehensive assessment of the literature is offered that looks at several methods currently employed for the identification and classification of ASD. It also discusses the problems and challenges that conventional methods face. A brief summary of the proposed methodology is provided in the next section, along with a schematic representation, theoretical justifications, and algorithms. In the Results and Discussion section, the performance outcomes and results associated with the proposed model are validated using specific available ASD datasets and assessment criteria. The Conclusion section provides a final summary of the research along with the findings, recommendations, and future actions.

RELATED STUDIES
The term ASD pertains to a number of sophisticated neurodevelopmental cognitive illnesses, including Asperger’s disorder, autism, and developmental disintegrative conditions (Griff et al., 2023). The word “spectrum” indicates that these medical conditions have a wide range of signs and severity levels. These conditions have been included under intellectual and cognitive disorders (Chan et al., 2023) and persistent learning syndromes in the World Statistical Taxonomy of Disorders and Relevant Health Problems. Early warning signs of ASD can include nonresponse to insulting others, difficulty in making eye contact, and dearth of interest in providers. Such signs often appear in the initial year of life. A tiny percentage of toddlers appear to grow up normally during the early stages of life, but around the ages of 18-24 months, they begin showing symptoms of autism, such as restricted and monotonous behavioral patterns, a limited selection of interests and activities, and limited proficiency in language (Saleh and Rabie, 2023). In the first 5 years of life, children who struggle to interact and communicate with society may suddenly become violent or introverted due to these diseases, which also impact how someone perceives and communicates with other people. Even though ASD first manifests in the early years (Rashid and Shaker, 2023), it typically lasts throughout childhood as well as adulthood.

Alsaeed and Alzahrani (2022) applied a DL algorithm for effective identification and classification of ASD. This study assessed the effectiveness of MobileNet, Xception, and VGG19 DL models in identifying ASD using face features. But the main flaw in this effort might be that the models that were suggested were not accurate enough. Oh et al. (2021) developed a unique automated detection framework for ASD detection. Here, the linear support vector machine (SVM) method-based computerized method has been implemented for the diagnosis of ASD. Altemeijji et al. (2020) looked into a number of ML methods to find an appropriate method for ASD prediction. Additionally, the goal of this work is to use algorithms trained with ML to diagnose autism more quickly so that therapy can be given to children at earlier and more significant developmental phases. It primarily offers a broad overview of ASD. It also signifies the intended audience and the advantages for the region. A hierarchical brain network technique has been developed by Qiang et al. (2023) to effectively diagnose and recognize ASD. The authors of the study primarily looked at input feature analysis to reduce forecast disorder and mistake rate. A deep neural network classification technique was used by Yang et al. (2020) in order to successfully identify and classify ASD. Deep neural networks have proven to be a useful tool for classifying data using both voxels and connection functions. Using the ABIDE repository, the MLP has been used in four different configurations to obtain enough samples for the DNN study. Ahmed et al. (2022) evaluated how well DL and ML methods performed in predicting ASD using eye-tracking data. The ability to pay abnormal attention to images is one of the most crucial components of effective learning.

For an early and precise diagnosis, the eye-tracking technology offers valuable information concerning children’s cognitive behavior. In order to evaluate the behavior of children with autism, it operates by scanning the pathways taken by the eyes in order to obtain a series of ocular extension points on the image. Three artificial intelligence (AI) methods for early autism detection are being investigated in this study: ML, DL, and a hybrid method combining the two. According to the study, it is determined that the hybrid classification methodology performs well and provides an improved disease diagnosis accuracy. Liao et al. (2022) investigated the use of ML techniques for diagnosing children with ASD. This study presented a multimodal paradigm that may identify children with ASD automatically. There were four phases taken into consideration. When the youngsters were supplying the stimulus during the data gathering stage, three sensors were utilized for gathering multimodal data. In the feature extraction stage, electroencephalogram, eye fixation, and facial expression features were extracted. The physiological and behavioral information were fed into a classification model to generate subdecisions after the behavioral elements had been incorporated with the eye fixation and facial expression features. The decision fusion utilized the subdecisions as components, and the weighted
Naïve Bayes (NB) algorithm produced the ideal classification outcome.

Nogay and Adeli (2020) aim to investigate and present an extensive evaluation of the most recent state-of-the-art ML research for determining the presence of ASD using structural, functional and hybrid imaging approaches. More extensive research needs to be done as the accuracy of studies using a larger number of respondents is generally lower than that of studies involving fewer individuals. The accuracy of the computerized diagnosis of ASD is higher in a younger age range, according to an analysis of the participants’ ages. It is anticipated that in the upcoming years, ML technology will become accessible to physicians and play a major role in the early and quick identification of ASDs. Hossain et al. (2021) aimed to identify the most important characteristics with the use of existing classification techniques and streamline the procedure of diagnosis for more accurate outcomes. For this purpose, they have examined youngster, child, adolescent, and adulthood ASD datasets. To find the best-performing classifier and feature set for each of these four ASD datasets, they assessed cutting-edge classification and feature selection strategies. Vakadkar et al. (2021) performed a comprehensive comparative study to examine the effectiveness of different ML techniques used for ASD prognosis. It includes the models of logistic regression (LR), decision tree, NB, K-nearest neighbor, and SVM. They may therefore conclude that the optimal model for the current dataset is LR, which has the maximum accuracy, based on the values acquired. When the training data are binary and small in size, LR works effectively. Even in cases where there are minimal correlations between the variables, the linearly split feature space remains effective. Nonetheless, NB makes the assumption that every feature is independent of the condition. Thus, the projected outcome may be off if there is interdependence among any of the variables. Parlett-Pelleriti et al. (2023) carried out a thorough analysis to look into the various unsupervised learning approaches used to diagnose ASD. Additionally, this paper discusses the many clustering strategies that have been employed in previous research projects to diagnose ASD. Sherkatghanad et al. (2020) created an effective and automated system for detecting and identifying ASDs using the CNN technique. In this instance, the layered architecture is designed to thoroughly review the inputs in order to accurately detect ASD. However, the suggested approach takes longer for the training and validation stages, which significantly affects the overall effectiveness of the system.

According to a review and analysis of the literature, the majority of earlier initiatives have focused on creating an efficient intelligence-based learning system that can recognize ASD from the data that are already available. However, the main issues with conventional approaches are their reduced detection rate, longer prediction times, and more difficult mathematical modelling. Thus, by developing a novel, inexpensive AI-based detection system, the intended study seeks to accurately identify and anticipate ASD based on the medical data provided.

**PROPOSED METHODOLOGY**

This section provides a thorough explanation of the proposed ASD detection process, together with a background, block diagrams, and mathematical information. ASD is typically a long-term neurological condition. ASD must be detected early, quickly, precisely, and successfully to receive an 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 achieved exceptional results at a lower cost. The work’s unique contribution is creating and implementing the AE\_L framework, which is used to identify and classify ASD using the medical data provided. Figure 2 presents the concept for the AE\_L system, and its main components are data preprocessing, feature engineering, ASD prediction, and performance evaluation.

AE\_L can automatically extract discriminative features from raw input data, such as neuroimaging scans, behavioral assessments, or genetic markers, without the need for manual feature engineering. By encoding the input data into a lower-dimensional latent space through autoencoding, AE\_L learns compact and informative representations that capture relevant patterns and structures associated with ASD. The warm equilibrium mechanism ensures that the model reaches a stable state during training, promoting the convergence of feature representations and preventing overfitting. AE\_L dynamically adjusts its parameters to maintain equilibrium between exploration and exploitation during learning, facilitating robust and stable model performance. The model’s adaptive nature enables it to adapt to variations and complexities in ASD-related data, including subtle patterns, heterogeneity across individuals, and variability in data modalities. By incorporating warm equilibrium learning principles, AE\_L enhances its capacity to learn from limited or noisy data, which are often encountered in clinical datasets for ASD detection. AE\_L can integrate information from multiple data modalities, such as imaging, genetic, and clinical data, into a unified representation for ASD detection. AE\_L can leverage complementary information from diverse sources through multi-modal fusion techniques to improve diagnostic accuracy and robustness. AE\_L can provide insights into the underlying features and patterns driving ASD detection through interpretable representations and decision-making processes. By analyzing the learned latent representations, clinicians can understand and validate the features contributing to ASD diagnosis, enhancing confidence in the model’s predictions. AE\_L’s automated learning framework allows for scalable and efficient training on large-scale datasets, enabling its deployment in real-world clinical settings. The model’s computational efficiency and scalability make it suitable for processing high-dimensional and heterogeneous data commonly encountered in ASD detection tasks.
The employment of an AE₂L approach for ASD recognition and classification is justified by its ability to handle the complexity of ASD diagnosis, extract informative features from heterogeneous data, adapt to data variability, maintain equilibrium dynamics, potentially enhance interpretability, and offer scalability and efficiency for clinical deployment. These advantages position AE₂L as a promising methodology for advancing the state of the art in ASD diagnosis and improving outcomes for individuals with ASD.

The warm equilibrium principle embedded within the AE₂L framework facilitates self-organization and stability in learning dynamics, promoting convergence to optimal solutions while avoiding overfitting and catastrophic forgetting. By maintaining a balance between exploration and exploitation during training, AE₂L mitigates the risk of model instability. It improves the capacity to learn from limited or noisy data, which is particularly relevant in ASD diagnosis, where data availability may be constrained.

Usually, an evaluation procedure may not be completed with only the collected data. As a result, the raw data must be filtered and structured into a format that can be used. Data preparation, also known as maintaining the data, entails separating some continuous variables, like an individual’s age, and adding or deleting missing values. The variables that are deemed superfluous or redundant must be eliminated from the repository before applying the given data. However, as part of data transformation, the data dimension is tailored by picking features based on the requirements of the model. In order to further enhance the ML prediction models’ reliability and efficiency, these two actions must be taken. The subsequent phase of the data identification procedure is called data exploration, and it involves using algorithms for the analysis of information and identification to generate a specific list of models over the data.

To figure out the appropriate ML method to anticipate ASD results and find more stringent restrictions, it is necessary to determine the data type for the user data. A classification

---

**Figure 2**: Block diagram of the proposed AE₂L framework. Abbreviations: AE₂L, auto-encoded warm equilibrium automated learner; ASD, autism spectrum disorder.
system that predicts the value of the unknown examples will identify the ruling stage. The confusion metrics will be used to assess each potential automated learning algorithm in order to determine its accuracy, sensitivity, and precision. The final step is to impart the newly acquired knowledge to an expert in healthcare so that it can be applied professionally to benefit society and assist parents in learning about their children’s status at a younger age. Following preprocessing, the innovative warm optimization (WO) technique is used to extract the most important and relevant features from the preprocessed medical data. Because of this, a special AET-MEL methodology is also used in order to correctly identify ASD based on the selected features from the medical data. An equilibrium optimizer is used to adjust hyperparameters during this classification process in order to make efficient decisions. The ASD diagnosis is successfully completed using the suggested AE₃L model, which is lightweight, easy to implement, and requires little processing time.

Warm optimized feature selection

The proposed study primarily uses the WO technique to select the most important features from the preprocessed medical data to effectively diagnose the disease. Earlier research studies have established numerous optimization techniques for feature selection and optimization. However, the bulk of techniques suffer from three key issues: high computational complexity, a lengthy time to obtain the best solution, and a significant failure rate at the local optimum. Thus, the proposed research project aims to apply a special optimization technique called WO for feature engineering. On the basis of the two varieties of worm breeding found in nature, WO has been established. The final pupal worm is generated by the weighted summing of the progeny resulting from both types. The Cauchy mutation operator is applied to expand the search space and prevent local optimum. The following is a model of the two sorts of reproduction types: reproduction 1 and reproduction 2. Since worms are zygotes, only one worm may generate progeny. Certain worms that are represented as reproduction system 2 produce multiple offspring, which is a particular worm reproduction method. Three distinct scenarios are considered in this reproduction method by varying the total number of spouses and the number of offspring produced. In order to analyze these three scenarios more thoroughly, several enhanced crossover operations—such as single-point, uniform, and multipoint crossover operations—are altered. Warm optimization provides the best enhancement for consistent crossover operation among all of these enhanced processes.

The best, most optimal features are generated as the output for classification by this method, which takes the input feature data as input for processing. After initializing the parameters, the offspring production is performed as shown in the following equation:

$$X_E = X_{max} + X_{min} - (\rho \times X_E^k),$$

where $X_E^k$ indicates the $k$th element of the worm $X_E$, which indicates the worm $E$, $X_{max}$ and $X_{min}$ are the upper and lower limits of the worm position, respectively, and $\rho$ represents a similarity variable between 0 and 1 that establishes the separation between the worm and its newly formed offspring. As a consequence, reproduction 2 is performed according to the following equation:

$$X_E^k = \begin{cases} X_1^k(k) = P_1^k; & \text{if rand} > 0.5, \\ X_2^k(k) = P_2^k; & \text{else,} \end{cases}$$

where $X_1^k(i)$ and $X_2^k(i)$ are the $k$th component of the two offsprings generated; $P_1^k$, $P_2^k$ are the two selected parents’ $k$th elements for uniform crossover operation. Then, the worms are generated for reproduction 2 as computed in the following equation:

$$X_{E} = \begin{bmatrix} X_1^k \\ X_2^k \end{bmatrix}$$

Moreover, the weighted summation of two offsprings is performed that determines the newly produced worm from reproduction 2 according to the following model:

$$X_{E}^k = w^1 X_1^k + w^2 X_2^k,$$

where $w^1$ and $w^2$ are the weight factors and can be obtained from the two offsprings’ fitness values $X_2^k$. As a consequence, the final position update of the worm in the next generation is also estimated based on the following equation:

$$X_{E}^k = (\rho \times X_{E}) + (1 - \rho) \times X_{E},$$

where $\rho$ represents the proportional factor. The Equivalency constant modifies the proportionate separation between two worms produced by two different methods of reproduction as represented in the following equation:

$$X_{E}^k = X_{E}^k - (\rho \times X_{E}^k) + (1 - \rho) \times X_{E} - X_{E} = Y_{i,k}^i,$$

where $Y_{i,k}^i$ is the arbitrary value drawn in the Cauchy distribution. Then, the differential term in a discrete time implementation is defined based on the fractional derivative as computed below:

$$\Psi_{\theta}[s(t)] = \frac{1}{\Gamma(r + 1)} \int_{r}^{s} (s - \tau)^{r} \Gamma(\theta + 1) \Gamma(\theta - r + 1),$$

where $\Gamma(.)$ is the gamma function, $g$ is the dimension of input features, $T^\theta$ are threshold derivative parameters, and $\theta$ is the parameter value of discrete time. At the end of this process, the final updated position of worm is determined as shown below:

$$X_{E}^k = (\theta - 1) X_{E}^k + \left[ \left( \frac{1}{2} \right) \theta \times X_{E}^k \right] + \left[ \frac{1}{6} \right] \times (1 - \theta) \times X_{E}^k - (\rho \times X_{E}) + (1 - \rho) \times X_{E}.$$
Additionally, the classifier can utilize it to accurately diagnose diseases.

Algorithm 1: Warm optimized feature selection (WOFS)

Input: Preprocessed medical data;
Output: Optimal features;
Procedure:
Step 1: Reproduction 1: The offspring \( X_1 \) is produced from the reproduction using Equation (1);
Step 2: Reproduction 2: For this uniform crossover operation, the two offsprings \( X_2^1(k) \) are generated using Equation (2);
Step 3: Generate worms \( X_3 \) generated for the reproduction 2 calculated using Equation (3);
Step 4: Estimate the weighted summation of the two offsprings \( X_3 \) using Equation (4) that determines the newly produced worm from Reproduction 2;
Step 5: Update the final position of the worm in the next generation \( X_3^t \) using Equation (5);
Step 6: Consequently, the proportionality constant adjusts the proportional distance between two worms generated from two kinds of reproduction systems \( X_3^t - X_2^t \) based on Equation (6);
Step 7: The differential term \( \psi(t) \) in a discrete-time implementation can be represented by fractional derivative as illustrated in Equation (7);
Step 8: Update the final position \( X_{opt}^t \) of the worm based on Equation (8);
Step 9: Return optimal feature set \( F_{opt} = X_{opt}^t \);

Auto-encoded term memory equilibrium learning model

Following feature selection, AET-MEL—a novel classification technique based on intelligence learning—has been utilized to correctly classify ASD from the provided medical data. A range of ML, DL, and hybrid algorithms were created for ASD detection in previous research efforts. However, a few of the more modern approaches have issues with reduced efficiency, complicated training and testing procedures, and poor prediction accuracy. Therefore, the goal of this research project is to combine long short-term memory (LSTM) approaches with the auto encoder (AE) mechanism to create a clever and lightweight categorization system. In this model, the EO technique is used to adjust the hyperparameters. In order to ensure an accurate ASD prediction, the proposed study intends to utilize the novel and intelligent classification technique, AET-MEL, by combining all these techniques together. The AET-MEL model is used for both ASD recognition and classification in order to ensure a reliable disease diagnosis. The DL model’s fundamental structure, the deep RNN model, is used to identify features and patterns from the medical data. However, unlike the RNN approach, the LSTM has memory cells for pattern recognition that rely on long- as well as short-term input datasets. These are useful for identifying and predicting outliers in time-series information sets. Three memory gates—forget, input (update), and output—are part of an LSTM cell.

Tanh and sigmoid functions have been included in the state vector of the LSTM model, in the course of time. The derivative components of the sigmoid and tanh functions become additive, while the algorithm examines the gradient error at each time point. It keeps the simulation from experiencing issues with gradient disappearing. This classifier uses a small-scale batch technique to train the data samples, in contrast to conventional gradient descent models. This model typically consists of one single layer of cells; however, the length of the layer can be increased by utilizing an ensemble and a mixture of many LSTM models. As a result, the model performs better and is more accurate during training. It also aids in determining the existence of large datasets and recurring trends over the short and long terms. Furthermore, the most significant features and long-term patterns are trained using the AET-MEL technique. It is an unsupervised kind of network architecture that uses encoding and decoding techniques to discover the fundamental hidden representation of the datasets.

To assess the variations, the input and output datasets have been compared. This indicates that the reconstruction loss is greater when there are significant disparities. It can be inferred from this result that the model is capable of interpreting the reconstructed dataset. As a result, it is acknowledged that the data are erratic. The encoder-decoder layer of the AET-MEL uses LSTM cells as part of its AE implementation. The benefits of both approaches are combined in this configuration for time-series data. Since it has several advantages over standard AEs, the AET-MEL has been used in this study. For example, it can receive sequenced information as input, but a normal AE is unable to take a consecutive sample as the input sample. Additionally, unlike the regular AE which only requires a preset dimension of the input dataset, the AET-MEL models accept a larger variety of input dimensions. The long- and short-term dependence on time in the prior data impacts the present information, which causes the data dimension to rise and the computation to become more complex. Finally, the AET-MEL model’s hyperparameter tuning uses the equilibrium optimizer. The position update equation for this discrete meta-heuristic includes

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Data information</th>
<th>Type of attributes</th>
<th>No. of samples</th>
<th>Yes or no</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ASD screening for children</td>
<td>Binary, categorical, and continuous</td>
<td>292</td>
<td>Y: 141; N: 151</td>
</tr>
<tr>
<td>2</td>
<td>ASD screening for adults</td>
<td>Binary, categorical, and continuous</td>
<td>704</td>
<td>Y: 189; N: 515</td>
</tr>
<tr>
<td>3</td>
<td>ASD screening for combined information</td>
<td>Binary, categorical, and continuous</td>
<td>996</td>
<td>Y: 330; N: 666</td>
</tr>
</tbody>
</table>

Abbreviation: ASD, autism spectrum disorder.
discrete valued operators (addition, multiplication, subtraction, and division). From the provided medical data, the AET-MEL model accurately identifies and classifies the ASD illness.

RESULTS AND DISCUSSION

This section verifies the efficacy and outcomes of the proposed AE$_{L}$ system using the commonly used ASD dataset and evaluation measures. To successfully execute and validate the suggested configuration, a Windows 10 PC with specs including a CPU speed of 2.9 GHz core i7, GPU of Intel 620, RAM of 12 GB, and a spare hard drive space of at least 5 GB has been used for the testing. Even though ASD is becoming more commonplace worldwide, there are not many publicly available statistics about the disorder’s study. Currently, the majority of resources are genetics-focused, and there are not enough clinical screening datasets for autism. The ASD used in this investigation has been gathered from a public UCI repository (Hossain et al., 2021; Sujatha et al., 2021). It was present in the adult and children’s datasets, which included 704 instances and 21 characteristics in the adult dataset and 292 instances and 21 attributes in the children’s dataset. There are 292 observations in this children’s dataset, in the age range 4 to 11 years. Due to the large number of respondents, working with the ASD dataset comes with several obstacles. Table 1 provides details about the dataset.

Accuracy

The probability of the classifier’s number of accurate predictions is known as accuracy. Stated differently, it represents the portion of accurate forecasts among all the forecasts.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%.
\]

Precision

The precision metric quantifies the degree of accuracy of our positive predictions, meaning that it indicates the proportion of presented beneficial outcomes that really occurred.

\[
\text{Precision} = \frac{TP}{TP + FP} \times 100\%.
\]

Recall

Recall quantifies the percentage of positives that our model detected, or the number of positively predicted points among those that were labelled as positive. Sensitivity and recall are synonymous.

\[
\text{Recall} = \frac{TP}{TP + FN} \times 100\%.
\]
F1 score

The weighted median of the precision and recall values has then been computed using the above information to determine the F1 score. The value here might range from 0 to 1. An F1 score of 1 indicates the best model; a value close to 1 indicates that a model is superior.

\[
F1\text{-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\%.
\]  

(12)

where, \(TP\) indicates the true positives, \(TN\) indicates true negatives, \(FP\) represents the false positives, and \(FN\) represents the false negatives. The aforementioned parameters have been employed in this study to validate the effectiveness of the suggested AE\(_2\)L model, which aids in determining the overall outcomes and results of the proposed ASD diagnosis system. The fitness plot for the WOFS approach is validated in this investigation, as Figure 3 illustrates. The WOFS is primarily employed in the suggested framework to assess an optimization’s effectiveness based on its fitness value. Based on the findings, it can be concluded that the suggested AE\(_2\)L models function effectively and offer enhanced optimization efficiency. Following that, as seen in Figure 4, the receiver operating characteristics of the suggested AE\(_2\)L framework are estimated for the provided data. The precision and recall values of the proposed AE\(_2\)L model are validated with respect to changing threshold as shown in Figure 5. The total efficacy of the prediction or classification techniques is often ascertained using the precision and recall metrics. The results of this study show that the suggested AE\(_2\)L model works well and offers better precision and recall values by accurately identifying ASD.

Furthermore, the training and validation accuracy of the proposed AE\(_2\)L model are estimated with respect to changing epochs, as shown in Figure 6. Similarly, the training and validation loss is also estimated for the proposed AE\(_2\)L method, as illustrated in Figure 7. According to the findings, the proposed AE\(_2\)L model improves the accuracy of the training and validation operations with reduced loss value. Adopting the WO technique is one of the primary factors for improving accuracy in the proposed mechanism because it reduces the size of the features needed for effective disease classification. Furthermore, the accuracy, precision, recall, and F1 score values for the suggested and standard classification methodologies utilized for ASD diagnosis are validated in
Figure 10: Log loss value. Abbreviations: ANN, artificial neural network; DT, decision tree; KNN, K-nearest neighbor; LR, logistic regression; NB, Naive Bayes; RF, random forest; SVM, support vector machine; XGB, XGBoost.

Figure 11: Comparison based on AUC. Abbreviations: ANN, artificial neural network; AUC, area under the receiver operating characteristic curve; DT, decision tree; KNN, K-nearest neighbor; LR, logistic regression; NB, Naive Bayes; RF, random forest; SVM, support vector machine; XGB, XGBoost.

Figure 12: Kappa coefficient. Abbreviations: ANN, artificial neural network; DT, decision tree; KNN, K-nearest neighbor; LR, logistic regression; NB, Naive Bayes; RF, random forest; SVM, support vector machine; XGB, XGBoost.

Figure 13: Overall performance analysis of the proposed AE\textsubscript{2}L model with and without WOFS. Abbreviations: AE\textsubscript{2}L, auto-encoded warm equilibrium automated learner; EWO, earthworm optimization; WOFS, warm optimized feature selection.

Figures 8 and 9. Consequently, as shown in Figure 10, the log loss value is also verified and contrasted with the ML techniques. The comparison data demonstrate that the proposed AE\textsubscript{2}L yields an ASD prediction result with a reduced log loss value compared to all other existing methods, as the main strategy for improving classification results is implementing WOFS and EO.

Furthermore, Figures 11 and 12 show the estimated area under the receiver operating characteristic curve and kappa values for the suggested and traditional learning approaches employed for ASD diagnosis. Next, as illustrated in Figure 13, an overall performance study is carried out for the suggested AE\textsubscript{2}L approach, both with and without WO. Based on all of these assessments, it is clear that the suggested AE\textsubscript{2}L method performs better than any of the current methodologies, as improved results are primarily due to integrating unique intelligence algorithms like EO, AET-MEL, and WO.

The authors have encountered challenges due to the limited availability of labeled image data for ASD, especially data containing subtle manifestations of the disorder. To address this, they have utilized data augmentation, transfer learning, or semi-supervised learning to augment the training dataset and improve model performance. The authors have conducted extensive validation studies on diverse datasets to assess the model’s robustness and generalization performance. They have employed domain adaptation or adversarial training techniques to improve cross-population generalization.

CONCLUSION

The AE\textsubscript{2}L model presents a promising approach for the automatic recognition and classification of ASD from images. By integrating advanced DL techniques, such as auto-encoded warm equilibrium learning, the AE\textsubscript{2}L model addresses critical challenges in identifying subtle patterns associated with ASD, offering potential solutions for improving early diagnosis and intervention. The AE\textsubscript{2}L
model’s ability to extract discriminative features from heterogeneous image data, adapt to variability in ASD manifestations, and maintain equilibrium dynamics during learning demonstrates its robustness and effectiveness in capturing complex patterns indicative of ASD. Furthermore, by incorporating interpretability mechanisms, the AE_L model enhances clinicians’ understanding of its decision-making processes, fostering trust and confidence in its diagnostic capabilities. While further research and validation are necessary to realize the clinical utility of the AE_L model fully, its development represents a significant step toward advancing the state-of-the-art technique in ASD diagnosis. By harnessing the power of automated learning and image analysis, the AE_L model holds promise for facilitating early detection, personalized intervention, and improved outcomes for individuals living with ASD. As we continue to refine and expand upon the AE_L model, we move closer to achieving our shared goal of enhancing the lives of individuals affected by ASD through timely and accurate diagnosis.

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


Qiang N., Gao J., Dong Q., Li J., Zhang S., Liang H., et al. (2023). A deep L model, we move


