




## Efficient Deep Learning and Machine Learning Models for Early-stage Identification of Autism Spectrum Disorder in Toddlers: Evidence from Saudi Arabia

Hasan Alkahtani<sup>1,2</sup>, Theyazn H.H. Aldhyani<sup>2,3,\*</sup>, Mohammed Y. Alzahrani<sup>2,4</sup> and Ahmed Abdullah Alqarni<sup>2,4</sup>

<sup>1</sup>Computer Science Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

<sup>2</sup>King Salman Center for Disability Research, Riyadh 11614, Saudi Arabia

<sup>3</sup>College of Applied - in Abqaiq, King Faisal University, Al-Ahsa 31982, Saudi Arabia

<sup>4</sup>Department of Computer Sciences and Information Technology, Albaha University, Saudi Arabia

Correspondence to:

Theyazn H.H. Aldhyani\*, e-mail: [taldhyani@kfu.edu.sa](mailto:taldhyani@kfu.edu.sa)

Received: June 5 2023; Revised: October 2 2023; Accepted: November 1 2023; Published Online: November 22 2023

### ABSTRACT

Autism Spectrum Disorder (ASD) is a type of developmental disorder that can have notable effects on a person's cognitive abilities, language skills, ability to recognize objects, social interactions, and communication skills. The primary etiology of this condition is attributed to genetics, and prompt detection and intervention may mitigate the potential for the individual to face exorbitant healthcare expenses and protracted diagnostic procedures. A machine learning (ML) and deep learning architecture was developed with the capability to effectively analyze datasets of autistic toddlers, accurately classifying and identifying ASD traits. To explore the feasibility of predicting and analyzing ASD characteristics across various age cohorts, we employed multiple supervised ML models, namely support vector machine (SVM), k-nearest neighbors algorithm, and decision tree, and deep learning models, such as long short-term memory (LSTM). In this study, we analyzed the ASD screening dataset of toddlers from Saudi Arabia. The ASD screening datasets of toddlers from Kaggle were used to test these models. The first dataset includes 1054 instances and 19 toddler-related features, while the remaining datasets consist of 16 features, 507 instances, 165 normal, and 141 ASD cases. We report baseline results of behavior classification using ML and DL approaches. The SVM approach achieved 100% accuracy, whereas the LSTM approach attained 100% accuracy in terms of the accuracy metric. The developed system demonstrates the efficacy of the ASD system in detecting ASD toddlers in Saudi Arabia. Furthermore, the ASD system has the potential to assist parents in examining their children at an early stage.

### KEYWORDS

autism spectrum disorder, machine learning, deep learning, diagnosis autism spectrum disorder

## INTRODUCTION

Autism Spectrum Disorder (ASD) is a type of neurodevelopmental disorder that usually becomes apparent during early childhood. To diagnose ASD, two essential areas of impairment must be identified: communication and social interaction, as well as the presence of repetitive or restricted patterns of behavior. This is supported by existing literature (American Psychiatric Association, 2013). The etiology of ASD is widely believed to be multifactorial in nature, with the involvement of various genetic and epigenetic factors (Centers for Disease Control and Prevention, 2022). This condition results in alterations in both the functions and the structure of the brain, ultimately leading to a reduction in social interaction and communication capabilities, as well as the manifestation of repetitive sensory and motor behaviors.

Individuals with ASD characteristics require a distinctive method of acquiring social skills and effective communication, setting them apart from neurotypical individuals. ASD patients may require varying levels of caregiver support, ranging from mild to severe (Centers for Disease Control and Prevention, 2022). ASD exhibits variability in its presentation among individuals, although certain characteristics are shared among those affected.

The body of literature on ASD in Saudi Arabia is comparatively restricted, resulting in the majority of prevailing clinical practices in the country being derived from the findings of research conducted in developed Western nations (Hassan, 2019). According to a study, the number of autism patients in 2002 was 42,500 (Yazbak, 2004). The lower reported

prevalence of ASD in Saudi Arabia may not necessarily indicate a lower occurrence of the disorder. This could be attributed to potential flaws in the methodology employed, limited diagnostic capabilities, and a lower level of awareness among parents regarding ASD symptoms. This lack of awareness may result in a reduced likelihood of recognizing symptoms and seeking appropriate care (Mostafa, 2011; Salhia et al., 2014). The findings of a cross-sectional study comprising 205 individuals diagnosed with ASD revealed a male:female ratio of 4.9:1 in the sample population. The study also revealed that a significant proportion of patients, 65%, exhibited psychiatric comorbidities.

Timely identification of ASD results in improved consequences, such as augmenting the youngster's linguistic, cognitive, and communicative proficiencies, along with their analytical and physical growth (Alnemary et al., 2017; May et al., 2017). However, in Saudi Arabia, the commencement of treatment for ASD may be hindered by various factors such as the parents' educational background, comprehension of the condition, annual income, and geographic location within the country. These factors pose unique considerations that may cause delays in treatment even after the diagnosis has been confirmed (Alnemary et al., 2017).

The majority of ASD screening methods rely on questionnaires or ratings (Thabtah and Peebles, 2019). The utilization of these methods is restricted in areas lacking licensed healthcare professionals and public health resources due to the requirement of trained professional examiners. A disadvantage of these assessments is their time-consuming nature, requiring multiple evaluations at different developmental stages for precise diagnosis. Automated and efficient ASD screening tools can reduce the burden on healthcare infrastructure and children. Machine learning (ML) techniques, particularly deep learning, have been utilized for automated diagnosis of ASD.

Deep learning (DL) and ML approaches have gained considerable attention and have made significant contributions in diverse fields, including computer vision, natural language processing, speech recognition, and robotics. DL techniques can automatically identify relevant features for a given learning task, eliminating the need for manual feature engineering. The diagnosis of ASD can be represented as a classification problem, and the utilization of deep learning techniques can enhance predictive accuracy and generalization to novel data. Prior research has investigated the prediction of ASD using brain imaging data (Heinsfeld et al., 2018), hand-crafted features (Nasser et al., 2019), and tracked gaze data from children watching movies on a tablet (Dawson et al., 2018). As our previous system was developed to detect ASD using a standard dataset (Alkahtani et al., 2023b), in this research, we have used the same algorithms to develop the system that is able to detect ASD using the Saudi Arabian dataset. The system was developed to explore the feasibility and effectiveness of our ML algorithms for ASD detection in Saudi Arabia, as well as the ethical implications of this approach.

The primary aim of our study is to investigate whether toddlers are vulnerable to ASD during its early stages, with the intention of improving the efficiency of the diagnostic procedure. The outcomes of our study indicate that support vector machine (SVM) and long short-term memory (LSTM)

models exhibit the greatest precision when applied to our chosen dataset, which is collected from various regions within Saudi Arabia.

The results of the study indicate the practicality of utilizing a tool based on deep learning and ML for the prediction of ASD in toddlers. Such a tool holds promise for the monitoring of ASD risk in the overall pediatric population and the early identification of high-risk children for focused screening in Saudi Arabia.

## BACKGROUND OF THE STUDY

The utilization of neural networks in ML for the purpose of deductive reasoning may pose certain challenges, particularly in cases where the model is responsible for making decisions related to medical matters (Rocchetti et al., 2021). The efficacy of a neural network can be significantly impacted by the quality of the input data. Therefore, when developing a neural network aimed at addressing a health-related concern, it is crucial to conduct a thorough assessment of the data's characteristics and quality. As a result, the synthetic minority oversampling technique was utilized to balance the autism dataset. Numerous research endeavors have sought to diagnose and manage ASD utilizing diverse ML methodologies. Bala et al. (2022) introduced a ML framework aimed at enhancing the identification of ASD across various age cohorts. Various classification techniques were applied to the aforementioned datasets. The results indicate that SVMs exhibited superior performance compared to other classifiers when applied to datasets related to the autism spectrum. Finally, the Shapley Additive Explanations (SHAP) method was utilized to identify the feature sets that yielded the highest level of accuracy. Hasan et al. (2022) presented a successful approach for assessing ML techniques in the context of early detection of ASD. The aforementioned system employed four distinct attribute scaling methodologies in conjunction with eight fundamental yet effective ML algorithms to accurately classify datasets that had undergone feature scaling. ASD was accurately identified by the linear discriminant analysis algorithms with a high degree of precision. Specifically, identification in the toddler and child datasets showed accuracies of 99.25% and 97.95%, while that in the adult and teenager datasets showed accuracies of 99.03% and 97.12%, respectively. Rodrigues et al. (2022) utilized functional magnetic resonance imaging and ML techniques to identify potential indicators of ASD prevalence in a separate study. The autism diagnostic observation schedule score was employed as a metric of severity. The results of the study suggest that there is a functional differentiation among ASD subclasses, as evidenced by the attainment of 73.8% accuracy in the cingulum regions. Raj and Masood, (2020) introduced a framework that incorporates various ML algorithms. The convolutional neural network-based predictive models for ASD screening demonstrated high levels of accuracy in datasets pertaining to adults, adolescents, and children, with respective rates of 99.53%, 96.68%, and 98.30%.

Hossain et al. (2021) endeavored to enhance the diagnostic process by identifying the most crucial features and

automating early diagnosis through the utilization of pre-existing classification algorithms. Based on their observations, it has been determined that the sequential minimum optimization algorithm, when applied to SVMs, outperforms all other ML techniques in terms of accuracy. The study demonstrated that the relief quality approach is the optimal technique for discerning the crucial attributes within datasets pertaining to ASD. Akter et al. (2019) gathered autism datasets that were initially identified for individuals in various age groups, including toddlers, children, adolescents, and adults. They utilized several feature transformation techniques to analyze these datasets. Subsequently, the efficacy of various classification methodologies was assessed by utilizing the modified ASD datasets. The SVM algorithm exhibited superior performance on the toddler dataset, while the Adaboost algorithm demonstrated optimal performance on the child dataset. The generalized linear model boosting algorithm outperformed the others on the teenage dataset, and Adaboost once again exhibited the best performance on the adult dataset. The researchers identified pivotal characteristics that show a strong correlation with ASD and achieved a 98.77% level of accuracy. Thabtah and Peebles (2019) proposed a novel ML-based architecture for screening adults and adolescents for autism. The architecture incorporates crucial elements and employs logistic regression for predictive analysis, thereby revealing significant insights into autism screening. In addition, diverse datasets underwent a comprehensive examination of features through the application of chi-square testing and information gain techniques to identify significant characteristics. The findings suggest that the employment of ML methodologies yielded predictive models with acceptable levels of efficacy. Pietrucci et al. (2022) collected 959 data samples from eight distinct projects and employed ML techniques such as random forest (RF), gradient boosting machine, and SVM to forecast the distinction between individuals with and without ASD. The researchers investigated the potential importance of gut microbiota in individuals with ASD. Their results indicate that three distinct algorithms consistently recognized *Parasutterella* and *Alloprevotella* as genera of significant importance. Furthermore, Omar et al. (2019) introduced a ML-driven model for predicting ASD and a mobile application suitable for individuals of all ages. The research yielded a prognostic framework for autism and a mobile software application by combining RF with Classification and Regression Tree, as well as RF with Iterative Dichotomizer 3. The efficacy of the model was evaluated through experimentation with a sample size of 250 authentic datasets, encompassing both autistic and nonautistic individuals. The proposed predictive model exhibited superior performance in terms of evaluation metrics when compared to both datasets. Akter et al. (2021) proposed a ML methodology to differentiate between subgroups of individuals with and without ASD, while also identifying the distinct characteristics of those with ASD. The researchers incorporated records pertaining to autism and employed the k-means clustering technique to discern distinct subcategories. The silhouette score was utilized to select the most suitable autism dataset. Subsequently, the principal dataset and its equitably distributed subclasses

were subjected to classification procedures utilizing classifiers. The SHAP method was utilized to rank features and assess discriminatory variables.

The utilization of ML has the capacity to enhance the prompt detection of ASD in young children, thereby resulting in earlier interventions and improved outcomes for children diagnosed with ASD. It is noteworthy to acknowledge that ML algorithms cannot substitute for clinical judgment and should be employed in tandem with conventional assessment techniques. The efficacy of ML models is contingent upon the quality of the data used for their training, and they are also vulnerable to potential inaccuracies and partialities. Consequently, it is imperative to guarantee that the data utilized for the purpose of training ML models is of high quality and encompasses diversity. Additionally, it is crucial to subject the models to rigorous validation and testing prior to their implementation in clinical settings. Furthermore, it is recommended that ML models serve as a supplementary aid for clinicians in the process of diagnosing and detecting children who are at risk rather than completely replacing clinical judgment (Alkahtani et al., 2023a).

## MATERIALS AND METHODS

### Proposed system

The diagram depicted in Figure 1 illustrates the overall functionality and progression of our system. The initial step involves dataset preprocessing, which entails the elimination of missing values and noise as well as the encoding of categorical attributes. Following the preprocessing of the data set, the classification algorithms of SVM, k-nearest neighbors (KNN), decision tree (DT), and LSTM classifiers are employed to detect the output label of ASD or normal. The precision of each classifier is evaluated and contrasted. In addition, the classifiers have been evaluated using metrics such as the F1 score and precision-recall values to enhance their assessment. If the classifier exhibits satisfactory performance, the accuracy of the training set will surpass that of the test set. Subsequently, this particular model may be considered the optimal model and employed for training and classification purposes.

### Dataset

The present data collection encompasses screening data on toddlers aged 12 to 36 months residing in various regions of Saudi Arabia, comprising both those diagnosed and those not diagnosed with ASD. Data were gathered through an online survey utilizing Google Forms. The survey incorporates an Arabic version of the Q-CHAT-10 queries, along with supplementary demographic data pertaining to the participants' age, gender, geographical location, and familial background of ASD. Identifying the entity responsible for test administration is essential. The questions of the ASD dataset are presented in Figure 2. The class of data is shown in Figure 3.

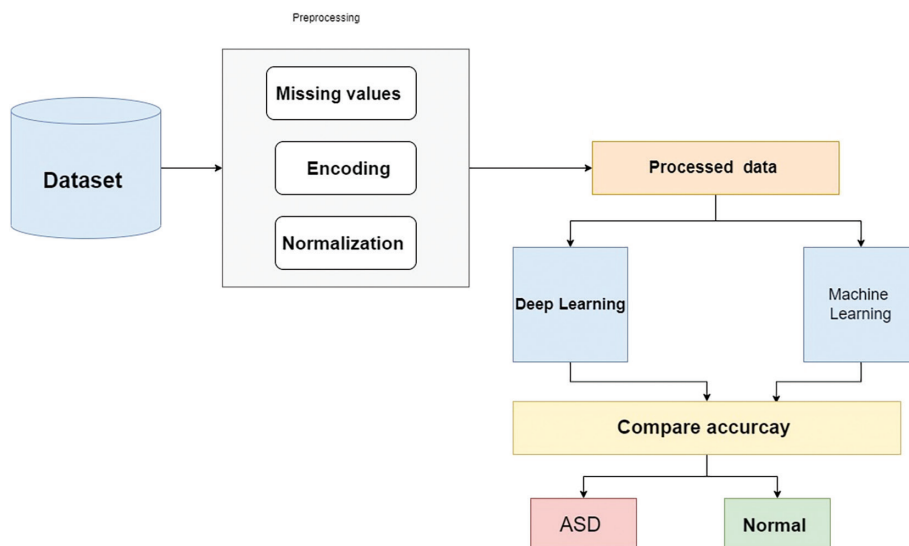


Figure 1: Framework of the ASD system. Abbreviation: ASD, Autism Spectrum Disorder.

|    | Question   |
|----|--|
| 1  | Does your child look at you when you call his/her name?  |
| 2  | How easy is it for you to get eye contact with your child?   |
| 3  | Does your child point to indicate that she/he wants something (e.g. toy that is out of reach)?   |
| 4  | Does your child point to share interest with you (e.g. pointing at an interesting sight)?  |
| 5  | Does your child pretend (e.g. care for dolls, talk on a toy phone)?  |
| 6  | Does your child follow where you are looking?  |
| 7  | If you or someone else in the family is visibly upset, does your child show signs of wanting to comfort them (e.g. stroking hair, hugging them)? |
| 8  | Would you describe your child's first word.  |
| 9  | Does your child use simple gestures (e.g. wave goodbye)?   |
| 10 | Does your child stare at nothing with no apparent purpose?   |

Figure 2: Features of the dataset. Abbreviation: ASD, Autism Spectrum Disorder.

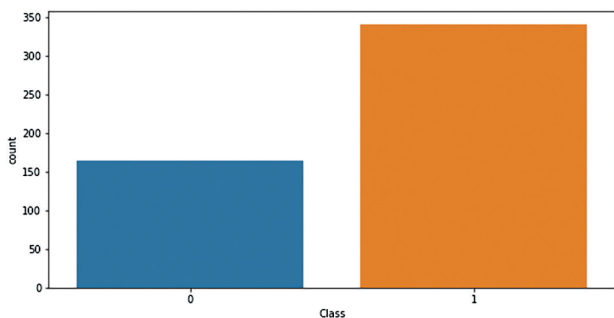


Figure 3: Class of data: 0, normal and 1, ASD. Abbreviation: ASD, Autism Spectrum Disorder.

### Preprocessing

Feature scaling is a method used to normalize independent features within a dataset to a consistent range. Data preprocessing involves normalization to address variations in magnitudes, values, or units. Failure to perform feature scaling can cause a ML algorithm to assign disproportionate weight to larger values and treat smaller values as relatively insignificant, irrespective of their units.

One-hot encoding represents categorical features—namely Region, Family member with ASD history, Who is completing the test, Age, Gender, and Screening Score—as binary



Figure 4: Preprocess step.

vectors consisting of zeros and ones. Categorical variables are represented using dummy variables, where each category is assigned a separate dummy variable. The value of the dummy variable is set to 1 if the observation belongs to that category and 0 otherwise. Figure 4 shows the preprocess step.

## Classification algorithms

### KNN algorithm

The KNN algorithm is a basic ML approach that relies on the principles of supervised learning. The KNN algorithm operates under the assumption of similarity between a novel data point and existing cases, thereby assigning the novel data point to the category that exhibits the highest degree of similarity with the available categories. The KNN algorithm retains all the available data and categorizes a novel data point by assessing its similarity to the existing data points. The KNN algorithm facilitates the efficient classification of newly acquired data by assigning it to a suitable category. The KNN algorithm is applicable for both regression and classification tasks, although it is predominantly utilized for the latter.

The KNN algorithm is founded on the calculation of the Euclidean distance between the newly introduced data point and the preexisting data points. The Euclidean distance is widely regarded as the most prevalent distance metric. It is likely that individuals utilize it frequently in their daily discourse when referencing the distances between their residence and workplace, among other things. The mathematical expression for this can be represented as follows:

$$E_i = \sqrt{(x_1 - x_2)^2 + (x_3 - x_4)^2} \quad (1)$$

where  $x_1$ ,  $x_2$ ,  $x_3$ , and  $x_4$  are training features.

### Support vector machines

SVMs possess a robust mathematical foundation and exhibit a close association with certain established statistical theories. In addition to accurately categorizing the training data, they also aim to optimize the margin to enhance the overall generalization performance. The aforementioned formulation results in the generation of a separating hyperplane that solely relies on the support vectors (i.e. the data points that are situated on the margin), which typically constitute a

minor proportion of the entire dataset. Therefore, the complete computational procedure is referred to as an SVM. Furthermore, given that real-world data analysis predicaments frequently encompass nonlinear dependencies, SVMs can be readily expanded to represent such nonlinearity through positive semi-definite kernels.

$$K(X, X') = \exp\left(-\frac{\|X - X'\|^2}{2\sigma^2}\right) \quad (2)$$

The feature vector utilized for training an algorithm on a dataset is commonly denoted as  $X, X'$  in the academic literature. The aforementioned feature vector is utilized for the purpose of assessing the dataset. Furthermore, the quantity  $(X - X')^2$  represents the squared Euclidean distance between two feature vectors, and it is a modifiable variable.

### Decision tree

A DT is a supervised learning algorithm frequently employed in ML for the purpose of modeling and forecasting results on the basis of input data. The structure in question is reminiscent of a tree, wherein internal nodes correspond to decisions or tests pertaining to a particular feature or attribute, branches correspond to the result of the said decision, and leaf nodes correspond to the ultimate decision or prediction. Presented below is an illustration of a DT.

$$Entropy = (S) = \sum_{i=1}^C p_i \log_2 p_i \quad (3)$$

$$entropy(S|B) = \sum_{j=1}^j \frac{|S_j|}{|S|} entropy(S_j) \quad (4)$$

$$Gain(S|B) = entropy(S) - entropy(S|B) \quad (5)$$

The training dataset is represented by  $S$ , and the class of the ASD dataset is denoted as  $C$ , which includes both attack and normal data. The probability of simple data, indicating class  $S_j$ , is denoted as  $P_j$ , and it pertains to the subsets of class in features  $B$ .

### Long short-term memory

The LSTM is a neural network architecture that can process inputs of varying lengths. Its implementation can be adapted

to meet specific needs in a flexible manner. The present study employs an LSTM architecture that comprises several layers, as depicted in Figure 5. The input layer of the multilayer LSTM model is trained on a flow-based dataset by receiving a specific number of sequential packets per flow (10, 64, or 100). The initial packet of a given flow is fed into the initial cell of the LSTM layer in a flow-based dataset. The output generated by the initial LSTM cell is utilized as an input during the subsequent packet’s arrival at the input. Consequently, the outcome of the initial cell has an impact on the functioning of the subsequent cell. Additionally, the output of the initial cell is utilized as the input for the subsequent LSTM layer. Simultaneously, the outcome of the initial cell within the second LSTM layer serves as the input for the subsequent cell within the same layer. Next, the output of the second cell in the third layer of the LSTM is utilized as the input for the third cell in the same layer. The ultimate objective of the sequence is to generate a label that categorizes the eight applications via the output layer.

$$f_t = \sigma(W_f \cdot X_t + W_f \cdot h_{t-1} + b_f) \tag{6}$$

$$i_t = \sigma(W_i \cdot X_t + W_i \cdot h_{t-1} + b_i) \tag{7}$$

$$S_t = \tanh(W_c \cdot X_t + W_c \cdot h_{t-1} + b_c) \tag{8}$$

$$C_t = (i_t * S_t + f_t * S_{t-1}) \tag{9}$$

$$o_t = \sigma(W_o \cdot X_t + W_o \cdot h_{t-1} + V_o \cdot C_t + b_o) \tag{10}$$

$$h_t = o_t + \tanh(C_t) \tag{11}$$

The provided equation comprises various components, such as the forget gate  $f$  that employs the sigmoid function  $\sigma$ , and the weight  $W_f$  between the forget gate and input gate. The previous hidden state is denoted as  $h_t$ , while the input at the current timestamp is denoted as  $i_t$ , which represents the

input gate at time  $t$ . The respective gate weight is denoted as  $W_f$ , and the tangent function is represented by  $\tanh$ . The weight between the cell state  $C_t$  and the network output  $W_c$  is included, along with the output gate weight  $W_o$  and input/output biases  $b_f$ ,  $b_i$  and  $b_o$ . Lastly, the output of the LSTM model is represented by  $h_t$ .

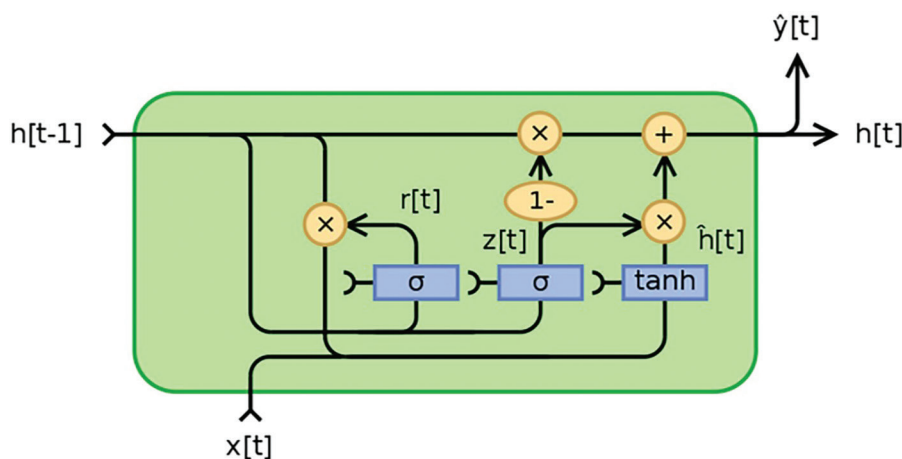
The inclusion of a cell state feature layer within an LSTM cell enables the determination of weight value maintenance by the LSTM model. The LSTM model possesses the capacity to selectively modify the cell state by means of specialized mechanisms known as gates, thereby regulating the flow of information. Gates serve as a mechanism to selectively permit the flow of information and are tasked with the responsibility of incorporating or removing prior information, thereby endowing the LSTM model with a high degree of persistence. The LSTM model exhibits superior performance in various tasks due to its ability to regulate long-term memory and output.

## EXPERIMENTS

This section presents the outcomes of the ML and deep learning models and highlights the noteworthy achievements of the developed system for detecting ASD.

### Environment of the ASD system

The configuration of an ASD system may exhibit variability contingent upon the particular system under consideration. Typically, such a system encompasses both software and hardware components. The development of ASD systems is dependent on the utilization of programming languages such as Python (specifically Jupyter Notebook) for the implementation of the proposed ASD system. The utilization of the TensorFlow library for composing LSTM models and the Sklearn library for ML were among the essential requirements. The hardware configuration for the development of



**Figure 5:** The LSTM structure for developing the ASD system. Abbreviations: ASD, Autism Spectrum Disorder; LSTM, long short-term memory.

an ASD system may comprise an Intel(R) Core i5-CPU and 4 GB of RAM.

## Splitting dataset

The efficacy of individual classification models was assessed by utilizing an 80% training set for their training and a 20% test set for evaluation. Both sets were employed in combination with the training set. The training phase is employed to train the ASD model, while the testing set is used to evaluate its performance for detecting ASD. The optimal split ratio is

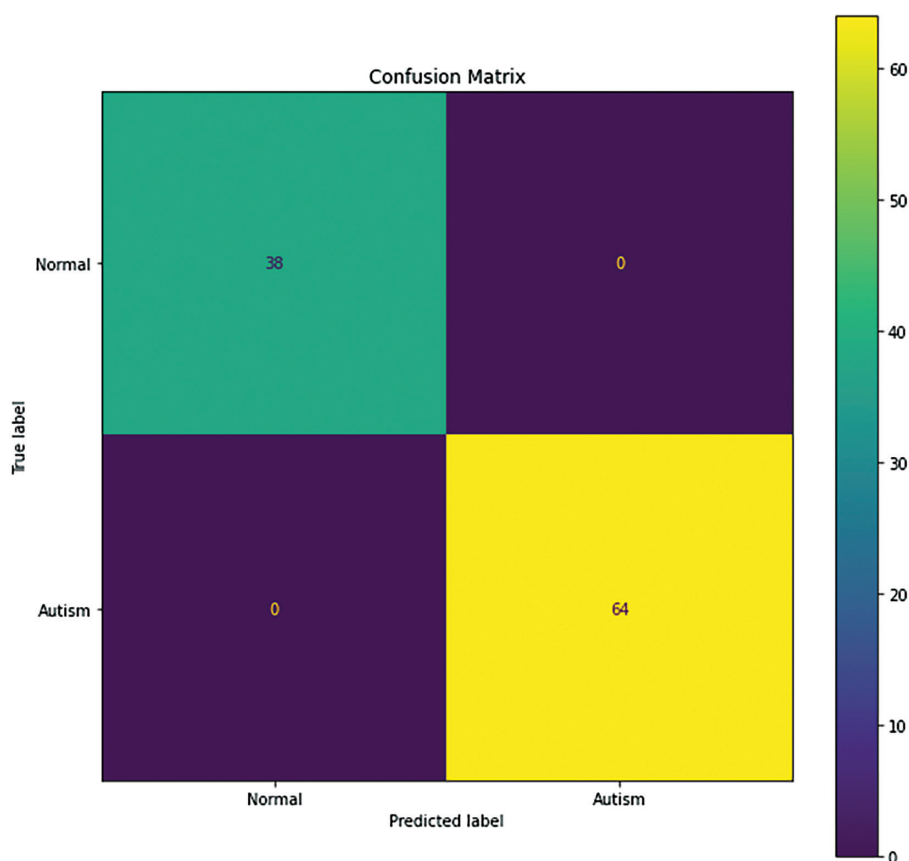
contingent upon the magnitude and intricacy of the dataset as well as the intended objective. However, it is crucial to ascertain that the training and testing subsets are reflective of the initial dataset.

The study involved training individual classification models using 80% of the training set, followed by an evaluation of their effectiveness using the remaining 20% of the test set. Both sets were employed in combination with the training set. To evaluate the performance of the classifiers, various evaluation metrics such as accuracy, area under the sensitivity, specificity, precision, and F1 score were calculated.

**Table 1:** SVM results.

|              | #Precision % | #Recall % | #F1 score % | #Support |
|--------------|--------------|-----------|-------------|----------|
| Normal class | 100          | 100       | 100         | 38       |
| Autism class | 100          | 100       | 100         | 64       |
| Accuracy (%) |              | 100       |             |          |
| Macro Avg    | 100          | 100       | 100         | 102      |

Abbreviation: SVM, support vector machine.



**Figure 6:** Confusion metrics of the SVM approach. Abbreviation: SVM, support vector machine.

**Table 2:** KNN results.

|              | #Precision % | #Recall % | #F1 score % | #Support |
|--------------|--------------|-----------|-------------|----------|
| Normal class | 97           | 97        | 97          | 38       |
| Autism class | 98           | 98        | 98          | 64       |
| Accuracy (%) |              | 98        |             |          |
| Macro Avg    | 98           | 98        | 98          | 102      |

Abbreviation: KNN, k-nearest neighbors.

### Evaluation metrics

Performance measurement is utilized to evaluate the efficacy of developing algorithms. The evaluation of the proposed system’s ability to detect ASD was conducted using four key performance indicators: precision, accuracy, recall, and F1 score.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \tag{12}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{13}$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{14}$$

$$Fscore = \frac{2 * precision * Sensitivity}{precision + Sensitivity} \times 100\% \tag{15}$$

### Results

#### Results of ML

Table 1 shows the results of the SVM used to detect ASD in toddlers. The SVM had a score accuracy of 100%, indicating that the SVM correctly classified all of the samples in the test set.

The confusion metrics for binary classification using the SVM approach are presented in Figure 6. The SVM model exhibited encouraging outcomes, with a true negative rate of 38 and a false-positive rate of 0. The SVM approach exhibited a misclassification rate of 0.

The efficacy of binary classification was evaluated through the application of the proposed KNN algorithms. The dataset was categorized into two groups, namely normal and ASD. The outcomes of the KNN algorithm utilized for the identification of ASD in Saudi Arabia are presented in Table 2. The KNN algorithms demonstrated a high level of accuracy,

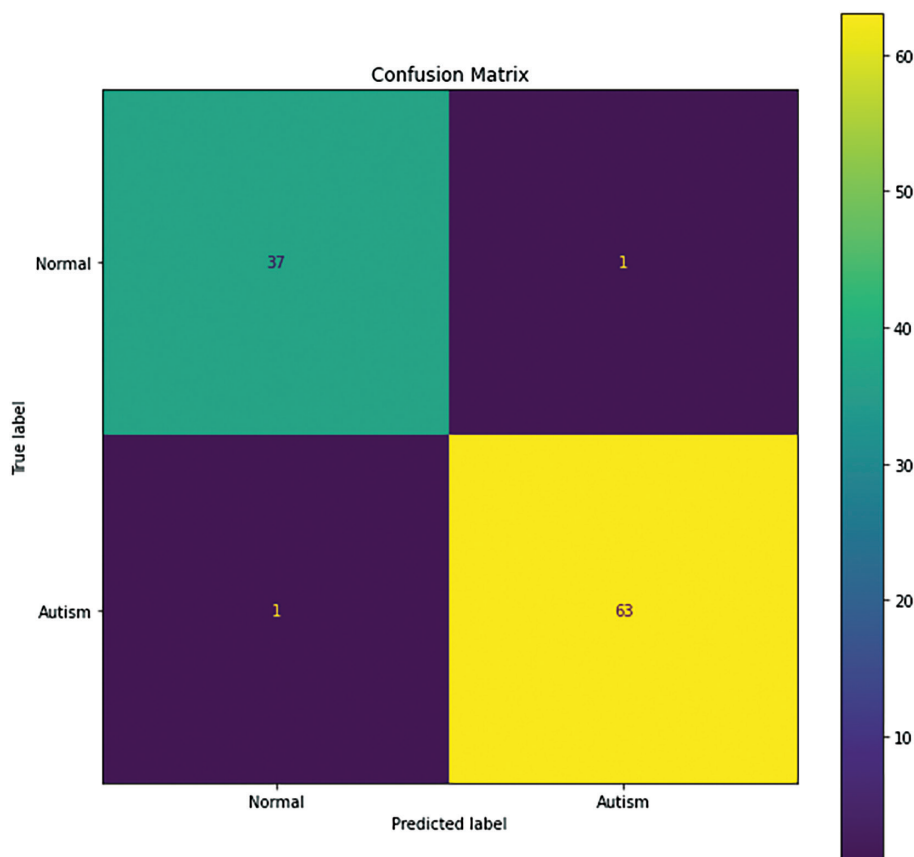


Figure 7: Confusion metrics of the KNN approach. Abbreviation: KNN, k-nearest neighbors.

Table 3: DT results.

|              | #Precision % | #Recall % | #F1 score % | #Support |
|--------------|--------------|-----------|-------------|----------|
| Normal class | 92           | 95        | 94          | 38       |
| Autism class | 97           | 95        | 96          | 64       |
| Accuracy (%) |              | 95        |             |          |
| Macro Avg    | 95           | 95        | 95          | 102      |

Abbreviation: DT, decision tree.



with the majority of them achieving a 98% accuracy rate, as evidenced by the empirical results.

The confusion metrics indicators, including the actual negative false-positive rate, valid positive rate, and false-negative rate, are depicted in Figure 7. According to the KNN model, the corrected normal was determined to be 37, while the model accurately identified and classified a score of 63 as belonging to the ASD category. The false-positive rate was accurately identified as 1, indicating that there were no instances of false positives. The KNN model yielded a 98% accuracy rate based on the obtained results.

The superior performance achieved by the DT algorithm is demonstrated in Table 3. The DT model achieved a 95% accuracy rate across all classes. The DT achieved a weighted average performance of 95% across two classes.

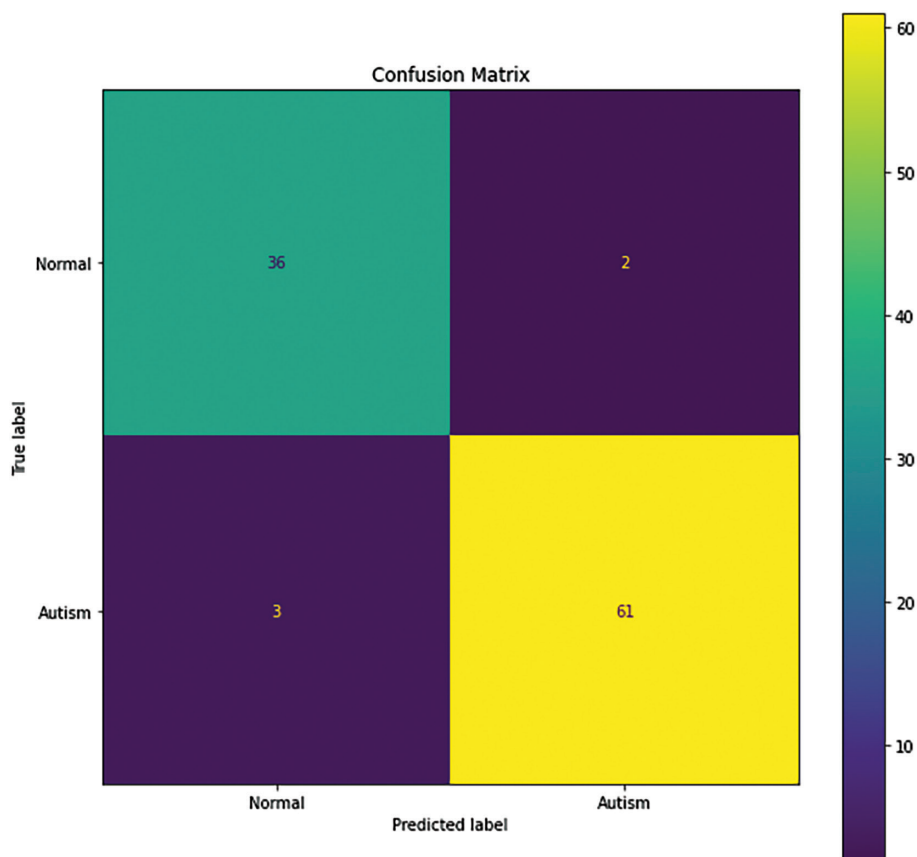
The confusion metrics of the DT model utilized for the purpose of ASD detection in the ASD system are presented in Figure 8. The DT graphical depiction exhibited an accurate classification of normalcy (TN=61). Additionally, 36 of

the classifications were correctly identified as normal. The high false-positive rate was more 3.

## Results of deep learning

The Keras API library was used to help the training of LSTM learning models. Matplotlib, Sklearn, and Pandas are widely used tools for data visualization and analysis in order to assess the efficacy of models. The research used Adam optimizers with a batch size of 120 and a learning rate of 0.001 across a span of 50 epochs. The dataset consisted of 507 occurrences, which were used for both training and testing purposes. The results of the LSTM model are shown in Table 4.

Figure 9 is an illustration of the accuracy performance and loss shown by the LSTM technology in the process of ASD detection. The findings suggest that there was a significant rise, from 70 to 100%, in both the accuracy of training and the volition of the participants. In addition, the accuracy

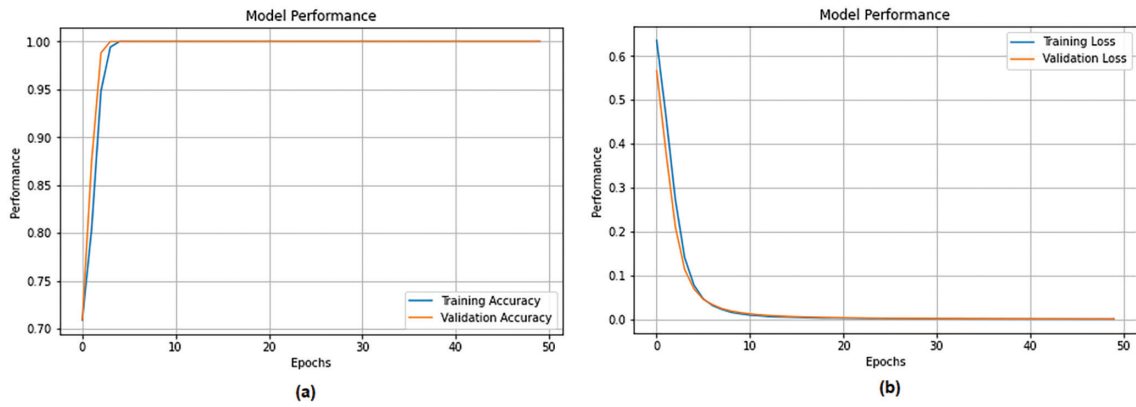


**Figure 8:** Confusion metrics of DT. Abbreviation: DT, decision tree.

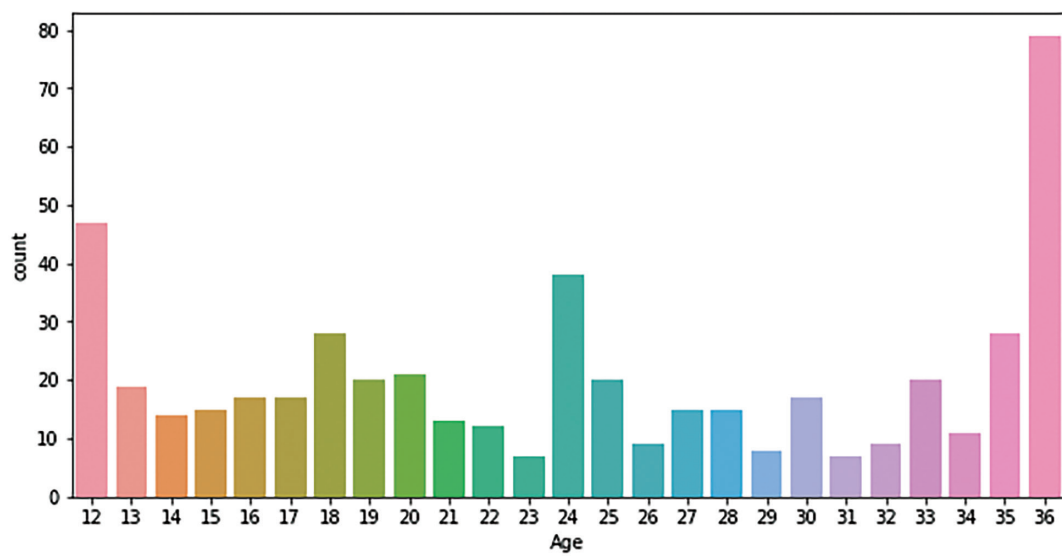
**Table 4:** LSTM results.

|              | #Precision % | #Recall % | #F1 score % | #Support |
|--------------|--------------|-----------|-------------|----------|
| Normal class | 100          | 100       | 100         | 38       |
| Autism class | 100          | 100       | 100         | 64       |
| Accuracy (%) |              | 100       |             |          |
| Macro Avg    | 100          | 100       | 100         | 102      |

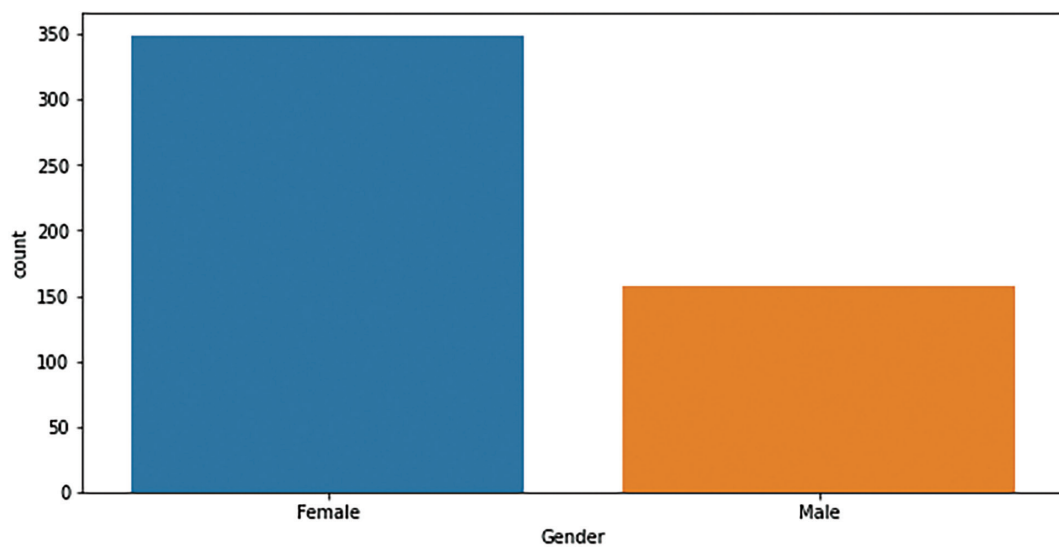
Abbreviation: LSTM, long short-term memory.



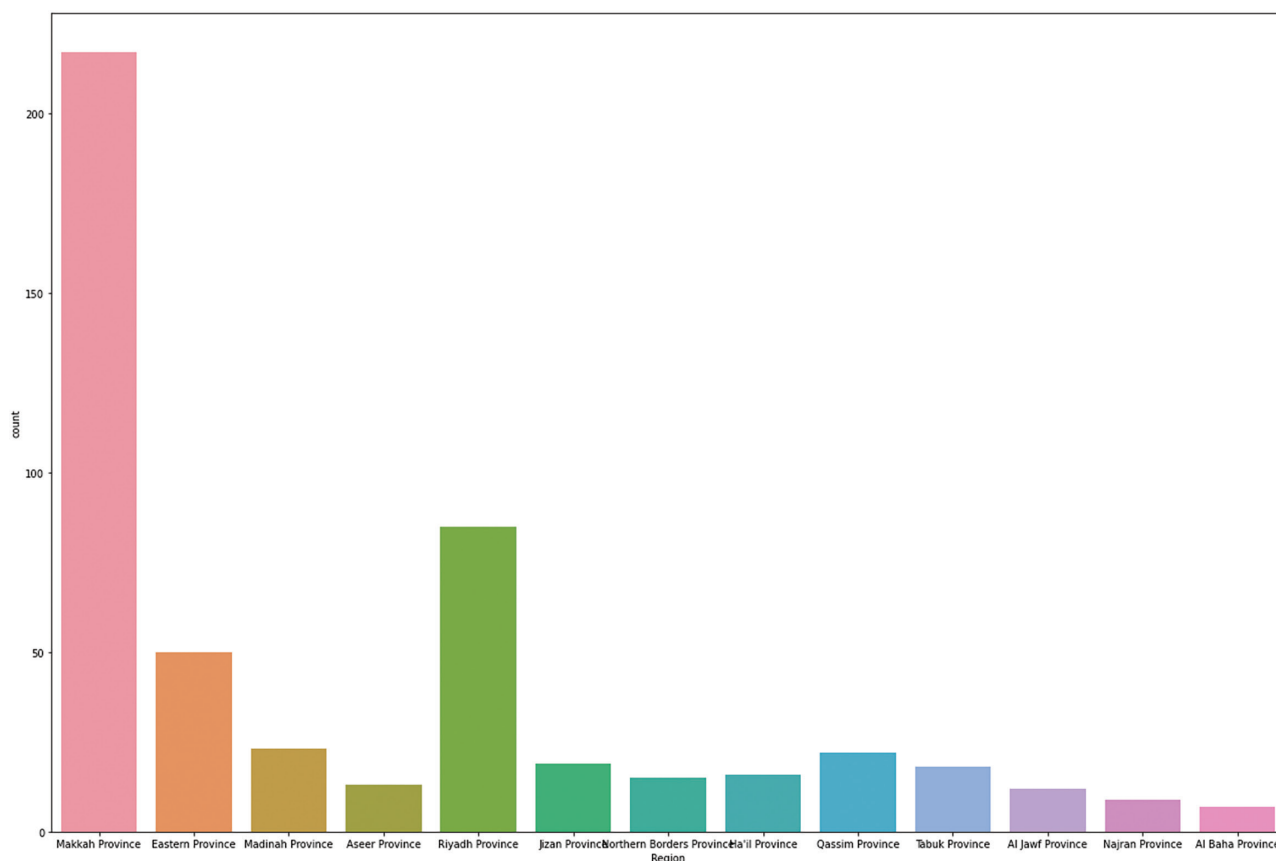
**Figure 9:** Performance of LSTM for detecting ASD using Saudi Arabia dataset. Abbreviations: ASD, Autism Spectrum Disorder. LSTM, long short-term memory.



**Figure 10:** Age of Toddlers.



**Figure 11:** Gender of dataset.



**Figure 12:** Region of Saudi Arabia where the dataset was collected.

loss observed throughout the testing phase decreased from 0.6 to 0.00 during the duration of the 50 epochs that were conducted.

Early identification of ASD in toddlers is crucial to facilitate their timely access to intervention and enhance their long-term prospects. This study showcases the viability of utilizing health claims data, as well as ML and deep learning models, to forecast the diagnosis of ASD during the early developmental stages. The study revealed that the SVM and LSTM models demonstrated a 100% accuracy rate in predicting ASD diagnosis across various time intervals ranging from 12 to 36 months. Figure 10 displays age categories.

Gender is a significant predictor for detecting ASD, with male and female sexes being key factors. The available dataset predominantly comprises female subjects, suggesting a higher prevalence of ASD among female toddlers. This is illustrated in Figure 11, which displays the gender distribution of the dataset.

The dataset was collected from different regions of Saudi Arabia, and Figure 12 shows the region of Saudi where the data have been collected.

Features were chosen based on their strong correlations with the respective classes. The outcomes of Pearson's correlation coefficient approach for identifying significant elements are illustrated in Figure 13. The majority of features in the dataset have an almost similar correlation.

## CONCLUSION AND FUTURE WORK

The present investigation examines the manifold applications of ML and deep learning methodologies for the identification and classification of ASD in toddlers in Saudi Arabia, along with their associated advantages and disadvantages. This is accomplished through a thorough assessment and scrutiny of current research endeavors. The present investigation employed various ML techniques, including SVM, KNN algorithm, DT, and deep learning models such as LSTM, to identify ASD in publicly available nonclinical ASD screening datasets from Saudi Arabia. These datasets were obtained from the Kaggle ML repository. The aforementioned datasets are associated with toddlers residing in various regions of Saudi Arabia. Multiple performance assessment criteria were employed to evaluate the efficacy of the models developed for the identification of ASD.

The evaluation of behavioral characteristics associated with ASD is a laborious undertaking that is further complicated by the presence of overlapping symptoms. At present, there is a dearth of diagnostic tests that can expeditiously and precisely ascertain the presence of ASD or a refined and comprehensive screening instrument that is specifically tailored to detect the emergence of ASD. An automated model for predicting ASD has been developed, utilizing a minimal set of behavioral characteristics selected from diagnostic datasets for each individual. We evaluated

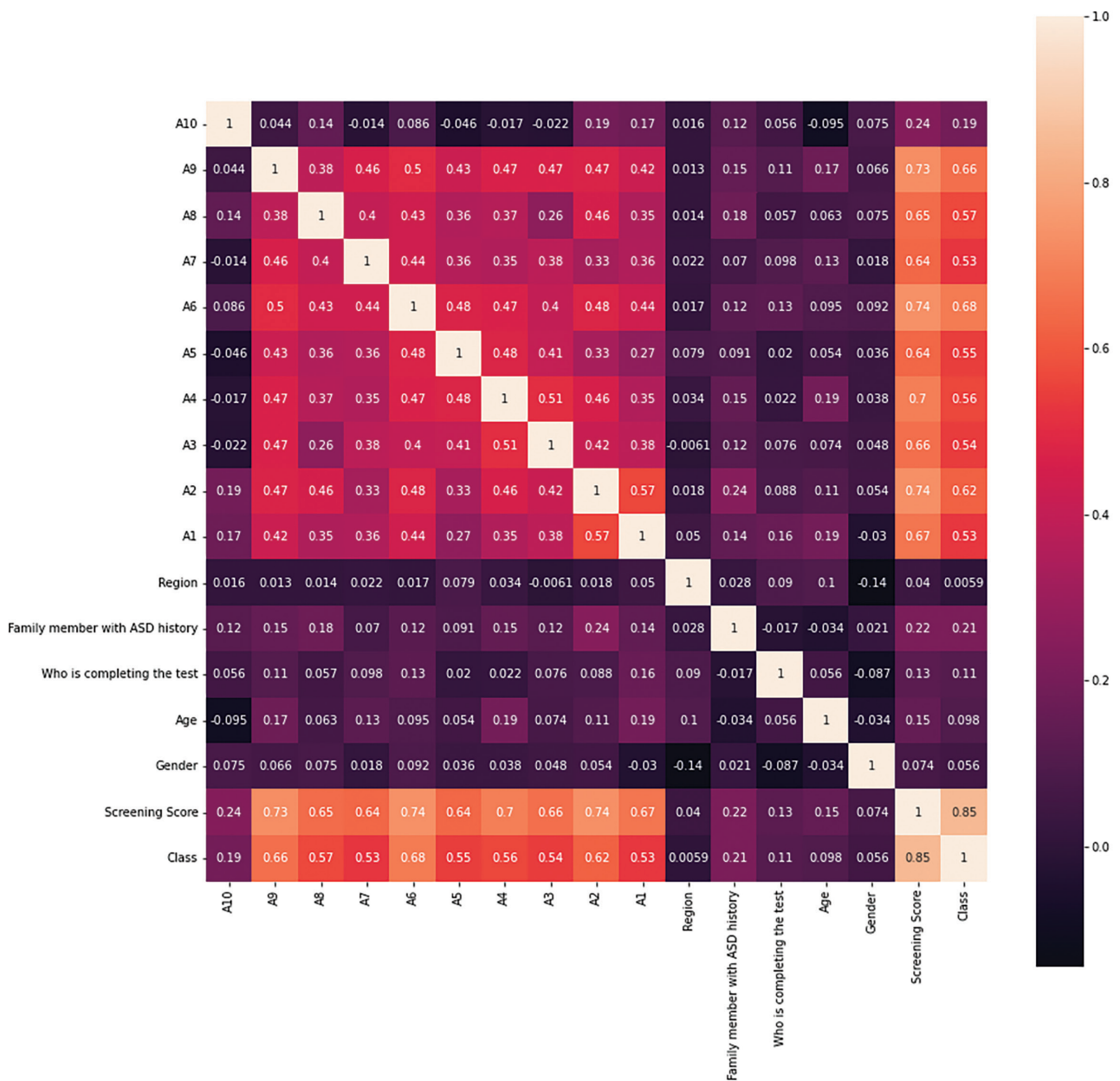


Figure 13: Correlation of features of the ASD dataset. Abbreviation: ASD, Autism Spectrum Disorder.

a set of ML and deep learning models on a dataset collected from Saudi Arabia to determine their efficacy. The system under consideration attained a notable level of precision, reaching 100%, through the utilization of both SVM and LSTM models.

Subsequent research endeavors may investigate the dynamic predicting of ASD diagnosis through the development of a mobile application designed to assist healthcare professionals, parents, and psychiatrists. The examination of these categorization models has the potential to serve as a fundamental framework for other scholars to delve deeper into this particular dataset or other datasets related to ASD.

## ACKNOWLEDGMENTS

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through research group No: KSGR-2022-013.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest in association with the present study.

## REFERENCES

- Akter T., Ali M.H., Satu M., Khan M. and Mahmud M. (2021). Towards autism subtype detection through identification of discriminatory factors using machine learning. In: *International Conference on Brain Informatics*, Springer, Cham, Switzerland; pp. 401-410.
- Akter T., Shahriare Satu M., Khan M.I., Ali M.H., Uddin S., Lió P., et al. (2019). Machine learning-based models for early stage detection of autism spectrum disorders. *IEEE Access*, 7, 166509-166527.
- Alkahtani H., Aldhyani T.H.H. and Alzahrani M.Y. (2023a). Deep learning algorithms to identify autism spectrum disorder in children-based facial landmarks. *Appl. Sci.*, 13, 4855.
- Alkahtani H., Aldhyani T.H.H. and Alzahrani M.Y. (2023b). Early screening of autism spectrum disorder diagnoses of children using artificial intelligence. *J. Disabil. Res.*, 2, 1, 14-25.
- Alnemary F.M., Aldhalaan H.M., Simon-Cerejido G. and Alnemary F.M. (2017). Services for children with autism in the Kingdom of Saudi Arabia. *Autism*, 21, 592-602.
- American Psychiatric Association. (2013). Autism spectrum disorder. In: *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; p. 50, American Psychiatric Association, Arlington, VA, USA.
- 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.
- Centers for Disease Control and Prevention. (2022). *Autism Spectrum Disorder (ASD)*. Available online: <https://www.cdc.gov/ncbddd/autism/facts.html> (accessed on 26 May 2023).
- Dawson G., Campbell K., Hashemi J., Lippmann S.J., Smith V., Carpenter K., et al. (2018). Atypical postural control can be detected via computer vision analysis in toddlers with autism spectrum disorder. *Sci. Rep.*, 8, 1, 1-7.
- 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.
- Hassan A. (2019). Arab views on autism. In: *Encyclopedia of Autism Spectrum Disorders*, pp. 302-305, Springer, New York, NY, USA.
- Heinsfeld A.S., Franco A.R., Craddock R.C., Buchweitz A. and Meneguzzi F. (2018). Identification of autism spectrum disorder using deep learning and the abide dataset. *Neuroimage Clin.*, 17, 16-23.
- Hossain M.D., Kabir M.A., Anwar A. and Islam M.Z. (2021). Detecting autism spectrum disorder using machine learning techniques. *Health Inf. Sci. Syst.*, 9, 386.
- May T., Sciberras E., Brignell A. and Williams K. (2017). Autism spectrum disorder: updated prevalence and comparison of two birth cohorts in a nationally representative Australian sample. *BMJ Open*, 7, e015549.
- Mostafa A. (2011). Addressing autism in the Arab world. *Nat. Middle East*, 21, 147.
- Nasser I.M., Al-Shawwa M. and Abu-Naser S.S. (2019). Artificial neural network for diagnose autism spectrum disorder. *Int. J. Acad. Inf. Syst. Res.*, 3, 27-32.
- Omar K.S., Mondal P., Khan N.S., Rizvi M.R.K. and Islam M.N. (2019). A machine learning approach to predict autism spectrum disorder. In: *Proceedings of the 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Cox's Bazar, Bangladesh, 7-9 February 2019; pp. 1-6.
- Pietrucci D., Teofani A., Milanese M., Fosso B., Putignani L., Messina F., et al. (2022). Machine learning data analysis highlights the role of parasutterella and alloprevotella in autism spectrum disorders. *Bio-medicines*, 10, 2028.
- Raj S. and Masood S. (2020). Analysis and detection of autism spectrum disorder using machine learning techniques. *Procedia Comput. Sci.*, 167, 994-1004.
- Rocchetti M., Delnevo G., Casini L. and Mirri S. (2021). An alternative approach to dimension reduction for pareto distributed data: a case study. *J. Big Data*, 8, 39.
- Rodrigues I.D., de Carvalho E.A., Santana C.P. and Bastos G.S. (2022). Machine learning and rs-fMRI to identify potential brain regions associated with autism severity. *Algorithms*, 15, 195.
- Salhia H.O., Al-Nasser L.A., Taher L.S., Al-Khathaami A.M. and El-Metwally A.A. (2014). Systemic review of the epidemiology of autism in Arab Gulf countries. *Neurosciences*, 19, 291-296.
- Thabtah F. and Peebles D. (2019). Early autism screening: a comprehensive review. *Int. J. Environ. Res. Public Health*, 16, 18, 3502.
- Yazbak F.E. (2004). Autism seems to be increasing worldwide, if not in London. *BMJ*, 328, 226-227.