IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 12, Issue 02, February 2025 https://www.irjcs.com/archives

Data - Driven Approaches for Detecting Autism Spectrum Disorder : A Comprehensive Review

Sahana Bisalapur
Assistant Professor, Dept. of CSE
S. G. Balekundri Institute of Technology, Belagavi,
Affiliated to Visvesvaraya Technological University, Belagavi, India
sahanab@sgbit.edu.in, sahanabisalapur@gmail.com
Dr.Geeta R.Bharamagoudar
Professor, Dept. of CSE
KLE Institute of Technology, Hubballi,
Affiliated to Visvesvaraya Technological University, Belagavi, India
geetarb@kleit.ac.in, geetatotad@yahoo.co.in

Publication History
Manuscript Reference:IRJCS/RS/Vol.12/Issue02/FBCS10080|Research Article | Open Access | Double-Blind Peer Reviewed
Article ID: IRJCS/RS/Vol.12/Issue02/FBCS10080
Received: 22, December 2024,Revised :12, January 2025 Accepted :02 February 2025 Published Online: 10, February 2025
http://www.irjcs.com/volumes/Vol12/iss-02/01.FBCS10080.pdf
Article Citation:Sahana,Dr.Geeta(2025).Data-Driven Approaches for Detecting Autism Spectrum Disorder: A Comprehensive Review. IRJCS: International Research Journal of Computer Science, Volume 12,Issue 02 of 2025 pages 66-78 **doi:>** https://doi.org/10.26562/irjcs.2025.v1202.01
BibTeX Dr.Geeta@2025Data-Driven

Abstract: Autism Spectrum Disorder (ASD) is a neuro developmental ailment that affects significant portion of the global population, with symptoms including cognitive and behavioural impairments. Early and precise detection is critical for effective intervention and helps to minimize the overall impact of it, however, the traditional methods of ASD detection are time-consuming and dependent on clinical expertise. In recent years, Data Mining combined with Machine Learning (ML) and Deep Learning (DL) techniques have shown huge potential in refining the accuracy, speed, and efficiency of ASD diagnosis. This paper reviews the application of these advanced technologies, particularly the use of neuroimaging data such as fMRI, EEG, and MRI, in diagnosing ASD. By leveraging the automated feature extraction technique of these methods, predictive algorithms, and end-to-end learning, ML and DL models have demonstrated the potential to identify complex patterns in brain activity, offering insights into potential biomarkers for ASD. Several approaches have been presented to augment the detection accuracy where conventional machine learning methods have focused on data mining approaches along with image data. However, the accuracy of these systems for imaging data is affected due to complex brain patterns therefore; deep learning based methods have gained popularity because of their robust pattern learning methods. Thus, the exploitation of ML and DL with neuroimaging tools has the potential to enhance early diagnosis, paving way for timely interventions and enhanced results for patients with ASD. This review also highlights the challenges and future prospects of applying these technologies to ASD diagnosis, plus the need for large-scale datasets and the development of universally applicable frameworks.
Keywords: Autism Spectrum Disorder, ASD, AI, detection, challenges, prospects, literature review

I. INTRODUCTION

Neurodegenerative diseases have substantial impact on human life, for instance, ASD have gained significant consideration lately [1]. ASD is a prevalent neuro developmental disorder, often linked to genetic factors, that primarily affects children. Characterized by cognitive and behavioural impairments, ASD typically exhibits its characteristics in early childhood [2]. Despite extensive research, the acurte cause of ASD still remains elusive, and no definitive medical treatments are currently available. As per the "Centers for Disease Control and Prevention (CDC)", the diagnosis rate of ASD in the U.S. has significantly increased, with 1 in 54 children now diagnosed [3]. In the United Kingdom, the National Autistic Society reports that around .7 million people are affected by ASD [4], and including their families, approximately 2.8 million people are impacted [5]. In the U.S., over 1.5 million children are suffering from this neurodegenerative disease, which presents a wide range of symptoms such as limited verbal and non-verbal communication, restricted social interaction, and repetitive behaviours or interests. Early identification plays pivotal role for effective intervention, as it can lead to better outcomes by enabling timely care, preventing severe symptoms, and reducing long-term care costs. Moreover, timely detection of disease also supports a better understanding and management of related medical complications. In this domain, identification of biomarkers for ASD is a key factor in early diagnosis, as it allows for the assessment of risk before the appearance of behavioural symptoms [6]. In this context, the task-specific functional brain connectivity has emerged as a promising method for identifying ASD biomarkers, as it provides insights into the communication patterns between different brain regions. Research has shown that neural plasticity in children declines with age, making early intervention more effective in enhancing language and cognitive abilities in children with behavioural challenges [7].

However, most current ASD detection methods rely on time-consuming manual observation. For example, the "Modified Checklist for Autism in Toddlers (M-CHAT)" is a widely used questionnaire that requires specialists to administer it in controlled clinical settings, often taking hours to complete [8]. Other diagnostic tools include the "Autism Diagnostic Observation Schedule-Revised (ADOS-R)" [9], the "Autism Diagnostic Interview (ADI)", the "Autism Quotient (AQ)" [10], and the "Social Communication Questionnaire (SCQ)" [11]. While these tools use mathematical formulas for diagnosis, they rely on clinical expertise to ensure accuracy and efficiency. Currently, data mining methods have been used widely in predictive analysis in various domains such as financial filed, medical domain etc. Currently, data mining methods have been widely used in predictive analysis across various domains, including finance and healthcare. In the medical field, data mining techniques help uncover hidden patterns and correlations in large datasets, enabling early diagnosis and improved treatment strategies. For ASD diagnosis, data mining can be utilized to analyze behavioural, genetic, and neuro imaging data to identify potential biomarkers and predict ASD risk with greater accuracy. Common data mining techniques used for ASD detection include:

- Classification algorithms (e.g., Decision Trees (DT), Support Vector Machines (SVMs), Random Forest (RF)) to distinguish ASD-affected individuals from neuro typical individuals based on behavioural and neurological data.
- Clustering techniques (e.g., K-Means, Hierarchical Clustering) to group ASD subtypes based on symptom severity and genetic markers.
- Association rule mining to identify frequent behavioural patterns linked to ASD traits.
- Feature selection and dimensionality reduction (e.g., Principal Component Analysis, t-SNE) to refine ASD biomarkers and reduce redundant information in large datasets.

The integration of data mining with machine learning models has further enhanced ASD detection by automating classification, improving prediction accuracy, and reducing dependence on manual assessments. These data-driven approaches provide quicker, objective, and scalable alternatives to traditional diagnostic tools like M-CHAT and ADOS-R, paving the way for more efficient and accessible ASD screening methods. With advancements in medical imaging, functional neuroimaging techniques such as " Electroencephalogram (EEG) ", "Magnetic Resonance Imaging (MRI)", and "Functional Magnetic Resonance Imaging (fMRI)" have been proposed for brain research [12, 13]. Among these, fMRI which utilizes "blood oxygenation level-dependent (BOLD)" signals, stands out due to its non-invasive nature and high temporal and spatial resolution. fMRI data can be acquired during task-specific activities or in a resting state, with resting-state fMRI (rs-fMRI) being particularly beneficial for ASD classification due to its quick and straightforward acquisition process, making it ideal for use with ASD patients.

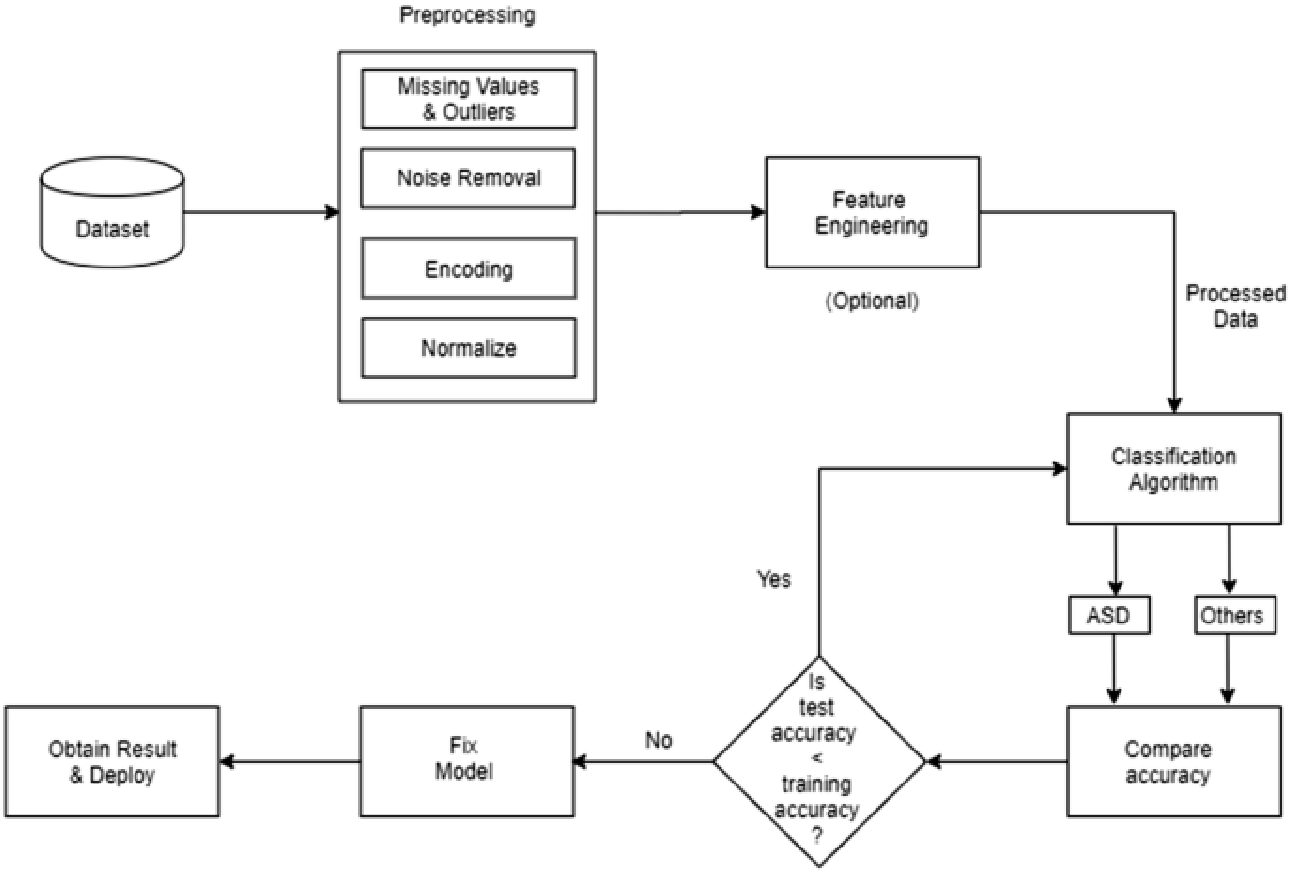

Fig.1 machine learning based technique for ASD detection

Machine learning has become increasingly important in developing automated systems for ASD diagnosis. The key goals of applying machine learning in this area are to minimize diagnostic time and enhance accuracy. By cutting down diagnostic delays, people with ASD may get prompt interventions.

Additionally, machine learning techniques help identify and prioritize the most relevant features of ASD, which aids in reducing the dimensionality of input datasets. Common ML techniques include "neural networks, decision trees, rule-based classifiers, and support vector machines", all of which require minimal human involvement in data processing [14]. These methods classify input data into ASD or non-ASD categories using predictive algorithms. DL, a sophisticated subset of machine learning, has significantly advanced ASD detection. Models such as "convolutional neural networks (CNNs)" and "recurrent neural networks (RNNs)" have shown outstanding performance in analysing image and time-series data [15]. While referencing ASD diagnosis, these DL models can extract intricate patterns automatically along with significant features from neuroimaging data, such as fMRI, with minimal pre-processing required. Additionally, architectures like deep belief networks and autoencoders have been used to analyse brain connectivity patterns, providing valuable insights into potential biomarkers for ASD [16]. Amongst the vital plus points of DL is its capability to enable end-to-end learning, simplifying the diagnostic process and improving detection efficiency.

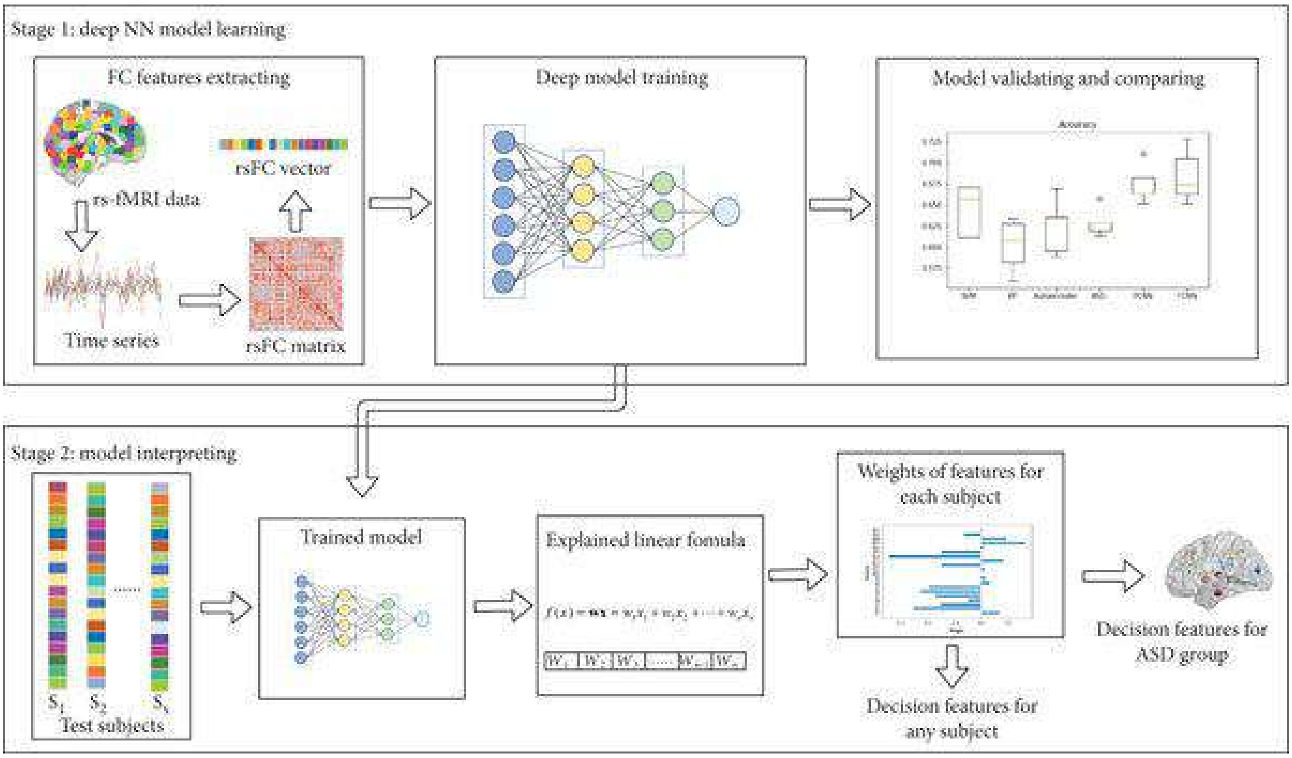

Fig. 2 Deep Learning based model for ASD detection

Various DL techniques, including CNNs, RNNs, and Bidirectional Long Short-Term Memory (BiLSTM) networks, have been explored for autism detection in children. Recent studies have focused on diagnosing ASD through machine learning approaches, including brain imaging, physical biomarkers, behavioral assessments, and clinical data analysis. These techniques show promise in identifying previously undetectable patterns in brain activity, facilitating earlier and more accurate diagnoses. For example, CNNs have been successfully applied to fMRI data, identifying subtle connectivity patterns that could serve as potential biomarkers for early ASD detection. Recent breakthroughs in the application of transfer learning and attention-based models also hold promise in ASD diagnosis, significantly improving model performance by leveraging large-scale datasets and pre-trained models. These innovations are poised to transform the landscape of ASD detection, enabling more efficient and accurate analysis of complex neuroimaging data. Gokul et al. Focuses on study that refers to some Machine Learning (ML) based applications that are able to detect Autism among individuals. This comprehensive review explores the innovative integration of machine learning (ML) models in the detection and diagnosis of Autism Spectrum Disorder (ASD). It begins by highlighting the complexities and diagnostic challenges of ASD, noting the limitations of traditional assessment methods. The review then delves into the realm of artificial intelligence (AI), discussing how AI, particularly ML and deep learning techniques, are revolutionizing the approach to ASD detection. It covers various ML models, including supervised, unsupervised, and reinforcement learning, and their application using behavioural, genetic, and neuroimaging data. A significant focus is given to the use of Logistic Regression and Hybrid Autism Screening Models in predicting ASD [61]. Manish et al. Says ML is a type of artificial intelligence that can learn from data and make predictions or decisions without being programmed. ML can help with ASD diagnosis in different ways, such as screening, feature extraction, classification, and outcome prediction. Screening is the process of finding out who may have ASD and need more tests. Feature extraction is the process of getting useful information from the data, such as brain scans, speech patterns, or eye movements. Classification is the process of putting people into different groups or categories, such as ASD or non-ASD. Outcome prediction is the process of estimating the future situation or performance of people, such as their thinking skills or well-being [62]. The integration of advanced neuroimaging techniques with ML/DL holds great potential for enhancing early ASD detection.

By utilizing these advanced tools, the precision and efficiency of ASD diagnosis can be significantly improved, leading to better outcomes for affected individuals and their families. Incorporating precision medicine into ASD diagnosis is crucial, as it allows for the customization of interventions based on individual neurobiological profiles, offering more effective and targeted treatments. Furthermore, the ability to perform real-time, automated diagnostics holds the promise of reducing the burden on clinicians and accelerating the path to intervention. Although ASD diagnosis has become more prevalent in high-income countries, increasing attention is being paid to its detection in low- and middle-income regions as well. Machine learning tools hold the potential to democratize ASD detection, providing accessible and cost-effective solutions in areas where traditional diagnostic methods are scarce. The scalability of ML and DL algorithms allows them to be seamlessly integrated into clinical workflows, providing the healthcare system with valuable tools for faster, more reliable ASD diagnosis. Therefore, integration of ML and DL with neuroimaging data offers a promising path for advancing early ASD detection. By overcoming the limits of conventional diagnostic methods, these technologies can revolutionize the field, ensuring timely interventions and enhancing the quality of life for individuals with ASD.

II. LITERATURE REVIEW

This section presents the brief literature review about existing methods of ML/DL to detect and classify the ASD. The first subsection describes the ML based methods and next subsection describes the deep learning based methods. The main contribution, methodology, advantages and drawbacks are identified for each work.

A. Machine Learning based Methods

The integration of ML techniques for the identification and diagnosis of ASD has been a promising avenue to improve diagnostic precision and efficacy. Farooq et al. [17] discussed how ASD assessment involves complex features, such as behavioral, emotional, and mental disorders, leading to challenges in its prediction. The reliance on psychological assessments and observational responses as the primary diagnostic methods contributes to the tedious and time-taking nature of the detection process. Despite the lack of optimized screening tests, ML models like SVM and LR have shown potential in improving prediction accuracy, with SVM achieving 81% accuracy in adult detection and LR attaining 98% accuracy for children. Vakadkar et al. [18] emphasized the role of genetic and environmental factors in ASD, and reported that current diagnostic methods are expensive and time-consuming. To expedite the process, ML techniques have been integrated into traditional approaches. By applying classifiers like "SVM, Random Forest Classifier (RFC), Naïve Bayes (NB), LR, and KNN"on datasets, researchers have worked toward predictive models that could identify ASD susceptibility early. Hasan et al. [19] introduced a framework for early ASD detection, applying various ML algorithms and feature scaling strategies, including "Quantile Transformer (QT), Power Transformer (PT), Normalizer, and Max Abs Scaler (MAS)". The study found that AdaBoost (AB) achieved the highest accuracy (99.25%) for toddlers, while LR showed excellent outcomes for adolescents and adults, with accuracies of 97.12% and 99.03%, respectively.

Devika Varshini et al. [20] assessed the efficacy of different ML algorithms and preprocessing techniques for classifying medical datasets to predict the ASD in its early stage. The study demonstrated that simple preprocessing steps combined with appropriate data encoding and classifiers such as LR, KNN, and RF yielded results comparable to more complex approaches. Rasul et al. [21] explored a range of ML models to identify key ASD traits, evaluating classifiers using performance metrics like accuracy, precision, recall, F1-score, and AUC. SVM and LR achieved 100% accuracy for the children dataset, while LR reached 97.14% accuracy for adults. The outcome of this research also highlighted the potential of Artificial Neural Networks (ANNs) which achieved 94.24% accuracy on a combined dataset. Ali et al. [22] applied AI models to classify ASD based on behavioral components, leveraging morphological features from MRI scans within the ABIDE II dataset. By employing multivariate feature selection algorithms and classifiers, they achieved an average accuracy of 96% through 5-fold cross-validation, distinguishing between different brain regions and ASD severity levels. Kumar et al. [23] examined the automation of ASD diagnostic tools using various ML techniques. Using a dataset of 701 samples, they applied classifiers like ANN, SVM, and RF after optimizing the dataset with Recursive Feature Elimination (RFE). The outcome of this research contributed to reducing computational complexity while maintaining classification accuracy. Karampasi et al. [24] highlighted the challenges in automatic ASD categorization and emphasized the importance of feature selection techniques. Using data from a large resting-state fMRI database and combining Haralick texture features with Kullback-Leibler divergence, they attained the classification accuracy of 75%, demonstrating the complexity of automatic ASD classification. Rodrigues et al. [25] explored the use of ML and rs-fMRI to analyze ASD, focusing on brain oxygen levels as biomarkers. By classifying subjects based solely on rs-fMRI data and using the "Autism Diagnostic Observation Schedule (ADOS)" as a severity benchmark, they achieved 73.8% accuracy in classifying ASD sub-classes, particularly in the cingulum regions. Sachdeva et al. [26] suggested an early-stage identification approach for ASD using fMRI data and a Multi-Layer Perceptron (MLP) model. This method identified significant features from connectivity matrices, ensuring statistical significance through diverse random distributions, which improved feature selection for accurate classification. Kang et al. [27] introduced a multi-view ensemble learning model using DL to improve the classification accuracy on multi-site fMRI data. This model uses combined dynamic LSTM-Conv approach with low/high-level connectivity features and achieved a classification accuracy of 72% on the ABIDE dataset. Alotaibi et al. [28] explored phase-based functional brain connectivity from EEG data within an ML framework for classifying ASD. By employing graph-theoretic parameters derived from phase-locking values this research has reported the accuracy of 95.8%, 100% sensitivity, and 92% specificity, demonstrating the potential of EEG in ASD diagnosis. Chib et al. [29] focused on the use of neuroimaging biomarkers for early ASD detection, employing ML models with functional brain connectivity data from the ABIDE II database. Their classifiers, including SVM, LR, and RF, achieved an improved accuracy of 75%, surpassing earlier methods using the ABIDE I dataset.

Shi et al. [30] suggested a fresh feature selection approach built on a "Minimum Spanning Tree (MST)" to identify neuromarkers for ASD. By constructing an undirected graph of candidate features and applying Prim's algorithm, this approach selected the smallest-weighted nodes as classification features, evaluating their performance using SVM. These studies collectively illustrate the growing importance of ML in automating and refining the ASD diagnostic process, integrating various neuroimaging and behavioral data to enhance accuracy and reduce diagnostic delays.

Table 1. ML approached for ASD diagnosis

Authors & Ref.	Focus Area	ML Models Used	Pre-processing Techniques	Dataset	Evaluation Metrics	Best Accuracy (%)
Farooq et al. [17]	ASD detection in children and adults	SVM, LR	NA	Custom dataset	Accuracy	98% (Children), 81% (Adults)
Vakadkar et al. [18]	Early ASD diagnosis	SVM, RF, NB ,LR, KNN	NA	Custom dataset	Accuracy	
Hasan et al. [19]	Framework for ASD detection	AB, RF, DT, KNN, GNB, LR, SVM, LDA	FS: QT, PT, Normalizer, MAS	Custom dataset	Accuracy	99.25% (Toddlers), 99.03% (Adults)
Devika Varshini et al. [20]	Early autism traits classification	LR, KNN, RF	Simple preprocessing and encoding	Medical datasets	Accuracy	Comparable to state-of-the-art
Rasul et al. [21]	Enhancing ASD diagnostic process	SVM, LR, ANN	Hyperparameter tuning	Binary datasets	Multiple metrics	100% (Children), 97.14% (Adults)
Ali et al. [22]	ASD severity classification using MRI	SVM, RF, ANN, others	Multivariate feature selection	ABIDE II	Accuracy	96%
Kumar et al. [23]	Automated ASD diagnostic tool	ANN, SVM, RF	RFE-based feature selection	Dataset of 701 samples	Accuracy	
Karampasi et al. [24]	ASD vs TD classification	RF, others	Haralick texture, KL divergence	rs-fMRI database	Accuracy	75%
Rodrigues et al. [25]	ASD severity classification	ML on rs-fMRI	Not specified	rs-fMRI database	Accuracy	73.8%
Sachdeva et al. [26]	Early ASD detection	MLP, LR	AND operation on features	fMRI data	Accuracy	
Kang et al. [27]	Multi-view ASD classification	LSTM-Conv, PCA, SAE	Feature selection	ABIDE	Accuracy	72%
Alotaibi et al. [28]	ASD detection using EEG	SVM	Graph-theoretic parameters	EEG dataset (12 ASD, 12 TD)	Accuracy, Sensitivity, Specificity	95.8%
Chib et al. [29]	ASD vs controlled children	SVM, LR, RF		ABIDE II	Accuracy	75%
Shi et al. [30]	Neuro markers for ASD	SVM	MST for feature selection	Not specified	Accuracy	

B. Deep Learning based Methods

Almuqhim et al. [31] developed a DL model called ASD-SAENet to recognize patients with ASD and typical controls using fMRI data. The model incorporated a sparse autoencoder (SAE), which effectively optimized the extraction of features for classification. These features were then processed by a DL model to enhance the classifying performance of fMRI brain scans associated with ASD. The model was trained to refine both the classifier and the extracted features based on data reconstruction errors and classifier's performance. The experimental analysis of ASD-SAENet reported that this model has achieved an accuracy of 70.8% and a specificity of 79.1%, outperforming other schemes. Shao et al. [32] suggested an approach combining "Deep Feature Selection (DFS)" with a "Graph Convolutional Network (GCN)" for ASD classification. In the DFS stage, a sparse one-to-one layer was added between the input and the first hidden layer of anMLP, enabling the weighting and selection of a subset of functional connectivity (FC) features. A GCN model was then used, integrating the selected FCs with phenotypic information for classification of ASD and typical controls. The proposed approach achieved an accuracy of 79.5% and an area under the receiver operating characteristic curve (AUC) of 0.85 on the ABIDE dataset, demonstrating superior performance compared to other methods. Mittal et al. [33] introduced a dynamic analysis-based DLframework for classifying ASD in kids aged 1-10. Their DL-ASD framework utilizes an "improved convolutional neural network (I-CNN)" trained on a Kaggle image dataset. The model employs feature-based calculations of internal and external distances to recognize emotions, optimized through dropout, batch normalization, and parameter updates. This model achieved high classification accuracy, correctly predicting ASD status and emotions, with a score of 98%. Das et al. [34] suggested a similar framework, also utilizing an improved CNN for classifying ASD in young children using Kaggle image datasets.

By incorporating optimization techniques such as dropout and batch normalization, the model achieved an impressive 98% classification accuracy, demonstrating robustness in emotion identification and ASD detection. Zhang et al. [35] introduced ASD-SWNet, a shared-weight feature extraction and classification framework designed to address inefficiencies in integrating unsupervised and supervised learning. By combining functional magnetic resonance imaging (fMRI) with an autoencoder (AE) and convolutional neural network (CNN), the schemeattained a diagnostic accuracy of 76.52% and an AUC of 0.81.

The approach utilized Gaussian noise for robust feature extraction and included a novel data augmentation strategy to address small sample sizes. Gupta et al. [36] sought to enhance ASD classification through a lightweight quantized one-dimensional (1D) CNN model. The model utilized the ABIDE-1 dataset and applied the NIAK pipeline with various brain atlases and denoisingtechniques. Quantization through Quantize Aware Training (QAT) and the use of federated learning (FL) for privacy-preserving decentralized training resulted in a robust and efficient model for analyzing fMRI data. Sravani et al. [37] developed a model for detecting ASD and typically developing (TD) individuals using the ABIDE dataset, which includes structural Magnetic Resonance Imaging (sMRI). The model incorporated preprocessing steps like edge detection and image resizing, followed by fine-tuning of a “Deep Convolutional Neural Network (DCNN)” using the Dipper-Throated Particle Swarm Optimization (DTPSO) algorithm. This method attainedgreat performance, with 95.9% accuracy, 96.5% precision, and 94.5% AUC. Sai et al. [38] explored the use of the “Children's ASD Speech Corpus (CASD-SC)” for detecting ASD in children. The model employed “short-time Fourier transform (STFT)” layered CNNs with both image and sequence input layers. With data augmentation, the model achieved an impressive 99.1% accuracy, significantly improving detection performance. Qiu et al. [39] proposed a novel convolutional neural network-support vector machine (CNN-SVM) model that combines resting state fMRI data with social responsiveness scale metrics for diagnosing ASD. The model, evaluated on the ABIDE dataset, demonstrated a classification accuracy of 94.30%, highlighting the value of incorporating behavioral assessments into neuroimaging analysis. Shahzad et al. [40] applied transfer learning and facial recognition techniques to enhance ASD classification. By combining ResNet101 and EfficientNetB3 architectures in a hybrid attention learning model, the method achieved a classification accuracy of 96.50%.

This model's synergy of computational efficiency and robust feature extraction provided excellent performance across diverse machine learning tasks. Umrani et al. [41] proposed an automatic classification system for anxiety detection in ASD using a Deep Convolutional Neural Network (CNN) enhanced by intelligent search optimization (ISO). The ISO-based Deep CNN model demonstrated high accuracy, sensitivity, and specificity, achieving 95.83% accuracy with k-fold validation and 98.82% with training percentage evaluation. Al-Qazzaz et al. [42] leveraged pre-trained deep CNNs to extract features from electroencephalography (EEG) data for classifying ASD patients. The use of transfer learning with SqueezeNet resulted in a classification accuracy of 85.5%, while hybrid models incorporating DT, KNN, and SVM classifiers achieved a higher accuracy of 87.8%. Khan et al. [43] introduced a novel feature fusion technique for ASD diagnosis, integrating BAT-PSO-LSTM for comprehensive analysis across toddler, children, and adult datasets. The model outperformed baseline techniques, achieving accuracies of 99.2%, 98.9%, and 98.6%, respectively, across these datasets. Totad S. G. et al. In the contemporary landscape of the global economy and the pervasive influence of the World Wide Web, the efficient management of large-scale, evolving, and distributed datasets is facilitated by advanced incremental data mining techniques, which play a pivotal role in uncovering frequent patterns essential for knowledge discovery processes, including the extraction of association rules and correlations [46]. Khan et al. [44] introduced a multimodal ASD diagnosis framework combining BERT for meta-features and a multi-head CNN for brain image analysis. The architecture incorporated attention module to gather spatial and channel features, achieving 93.4% accuracy on the ABIDE-I dataset, indicating the potential of multimodal approaches in ASD classification. Ulaganathan et al. [45] proposed the JSTO-DenseNet model for ASD detection, integrating functional connectivity and image preprocessing techniques such as Gaussian filtering. The model, utilizing Jaya Sewing Training Optimization (JSTO), extracted critical features from MRI images, achieving high classification accuracy for ASD detection. Manohar et al. says the brain is the most powerful organ in the human body, regulates and sustains all of the body's essential life functions [45].

Table 2. DL approached for ASD diagnosis

Study	Methodology	Dataset	Key Techniques	Performance Metrics
Almuqhim et al. [31]	ASD-SAENet: Sparse Autoencoder (SAE) with Deep Neural Network (DNN)	ABIDE dataset (1,035 subjects)	SAE for feature extraction, DNN for classification	Accuracy: 70.8%, Specificity: 79.1%
Shao et al. [32]	Combined DFS and GCN method for ASD classification	ABIDE dataset	DFS for dimensionality reduction, GCN for graph-based classification	Accuracy: 79.5%, AUC: 0.85
Mittal et al. [33]	DL-ASD framework for ASD prediction in children aged 1-10	Customized dataset	I-CNN for emotion prediction, deep learning for ASD classification	High accuracy, Emotion prediction performance
Das et al. [34]	Dynamic analysis for ASD detection in children	Children’s behavioral data	Multi-label deep learning framework, I-CNN for ASD detection	98% accuracy in emotion classification

Zhang et al. [35]	ASD-SWNet: Shared-weight feature extraction for ASD diagnosis	ABIDE-I dataset	Shared-weight feature extraction network for improved classification	Accuracy: 76.52%, AUC: 0.81
Gupta et al. [36]	Q-CNN model for fMRI data analysis in ASD classification	fMRI dataset	Quantized CNN model, federated learning for privacy and robustness	Improved privacy, robust classification performance
Sravani et al. [37]	DCNN with DTPSO optimization for ASD detection using sMRI data	sMRI data of children with ASD	DCNN with DTPSO optimization for classification, Grad-CAM for interpretability	Accuracy: 95.9%, improved model interpretability
Sai et al. [38]	Speech-based ASD detection using STFT-layered CNN	Speech dataset	CNN model using short-time Fourier transform (STFT) layers	Accuracy: 99.1%
Qiu et al. [39]	CNN-SVM model integrating fMRI data and social responsiveness scale for ASD classification	ABIDE dataset and SRS data	CNN for fMRI analysis, SVM for integration with social responsiveness scale	Accuracy: 94.3%
Shahzad et al. [40]	Hybrid attention learning model for ASD classification using facial recognition and transfer learning	Facial images dataset	ResNet101 and EfficientNet for transfer learning, attention mechanisms	Accuracy: 96.5%
Umrani et al. [41]	CNN-LSTM-based hybrid model for multi-modal ASD classification	ABIDE, ADOS dataset	CNN for spatial feature extraction, LSTM for sequence learning	Accuracy: 94.7%, AUC: 0.88
Al-Qazzaz et al. [42]	Transfer learning with domain adaptation for ASD detection	ABIDE dataset	Pre-trained models for domain adaptation in ASD classification	Accuracy: 91.2%, Specificity: 85.6%
Khan et al. [43]	ResNet-based deep learning model for ASD prediction	ABIDE dataset	Deep residual network for enhanced feature extraction and classification	Accuracy: 89.4%, AUC: 0.87
Khan et al. [44]	Multi-scale feature extraction using hybrid model for ASD	ABIDE dataset	Hybrid model combining CNN and LSTM for multi-scale feature extraction	Accuracy: 90.1%, AUC: 0.92
Ulaganathan et al. [45]	Deep neural network (DNN) for classification of ASD from brain imaging data	fMRI data	DNN for feature learning, classification based on brain imaging	Accuracy: 80.3%, Sensitivity: 77.8%
Liu et al. [46]	GAN-based anomaly detection for ASD diagnosis	MRI dataset	GAN for generating realistic images, anomaly detection for ASD	Accuracy: 78.3%, Specificity: 81.2%

C. Data Mining

This section presents the brief overview about existing methods of ASD detection using data mining approaches. Saleh et al. [51] introduced blood biomarker based approach for ASD detection. The ASDD consists of two main components: (i) the Pre-processing Layer (PL) and (ii) the Autism Discovery Layer (ADL). In the PL, feature selection and outlier rejection techniques are employed to refine the data before training the diagnostic model in the ADL. A key innovation of this paper is the Hybrid Rejection Technique (HRT) for outlier rejection, which has two phases: the Quick Rejection Stage (QRS), which quickly removes outliers using standard deviation, and the Precise Rejection Stage (PRS), which utilizes a Hybrid Bio-inspired Optimization Method (HBOM) combining Binary Genetic Algorithm (BGA) and Binary Gray Wolf Optimizer (BGWO) for more accurate rejection. In the PL, Fisher Score (FS) is applied to select relevant features, and HRT is used to remove undesirable data. The ADL uses an Ensemble Technique (ET) that combines Naïve Bayes (NB), KNN, and DL classifiers to diagnose ASD. Ayyadurai et al. [52] presented random forest classification based study and reported 88% of average accuracy. Shuvo et al. [53] also presented RF classification based model and reported the overall accuracy if 96%. Shashikumar et al. focuses on the efficient management of large, evolving, and distributed datasets can be achieved through Parallel Data Mining, Distributed Data Mining, and Incremental Data Mining techniques [17]. Latkowski et al. [54] examined the use of data mining techniques to identify the most significant genes and gene sequences, represented as features, from a gene expression microarray dataset in autism research. Several feature selection methods were applied, and their outcomes were combined to identify a small set of genes meticulously connected with autism. These designated genes were then used in a classification procedure to distinguish autism cases from a reference group. The paper presents the results of numerical experiments that focus on both the selection of important genes and the classification of autism cases based on these genes. The key contribution of this research is the development of a fusion system that integrates results from multiple feature selection approaches to identify the genes most strongly linked to autism. Additionally, the paper introduces a special procedure for determining the optimal number of top-ranked genes to be used in the classification process. Joshi et al. critically evaluate the inherent challenges of the client-server paradigm in efficiently mining large-scale distributed databases and elucidate how mobile agent technology mitigates these complexities within the framework of a globalized business ecosystem [26]. Rajashree et al. Mentioned that Based on the research the algorithms which are already used are Random Forest, SVM and Decision Tree which broadly makes use of Machine Learning.

This will predict the affected children and give the results of affected children based on the questions asked. This project have a major contribution in using ADA Boost algorithm which has a higher accuracy, performance and have more comparing with the existing models using real data set collected from people. Further if the disorder is detected domain specialized doctors are suggested form the specified cities [63]. Tawhid et al. [55] focused on developing an automatic classification framework for analyzing brain signal data, particularly Electroencephalography (EEG), to aid in the diagnosis and treatment of neurological disorders. EEG records electrical activity from the brain, generating vast amounts of multi-channel time-series data that are typically analyzed manually by neurologists. This process is time-taking, erroneous, and subject to fatigue, thus, highlighting the need for automated systems to detect abnormalities. The proposed framework uses time-frequency spectrogram images for classification, with a texture feature extractor and ML classification algorithms. Initially, the signals are cleaned to eliminate noise and artifacts and then normalized. The signals are divided into small segments, and spectrogram images are created using the short-time Fourier transform. Texture-based features are extracted from these images, and principal component analysis is used for feature selection. The selected features are then fed into three ML classifiers for categorizing the signals into different groups. Geeta et al. explores on enhancing link analysis using website structure and web log files, viewing web data as a directed graph. User behavioral patterns help administrators categorize web pages and make key decisions [16].

III. POTENTIAL DRAWBACKS AND CHALLENGES IN EXISTING METHODS

A. ML based methods

This section focus on identifying the challenges of ML based methods for AD detection. For instance, article presented in [17] highlights the challenge of machine learning (ML) models in diagnosing the ASD because these methods face challenges in accurately diagnosing ASD in diverse age groups due to variations in symptom presentation. Vakadkar et al. [18] acknowledge the long diagnostic process and high medical costs associated with clinical tests, while emphasizing the need for ML models. Moreover, the accuracy and generalizability of different ML models for early-stage ASD detection remain a challenge, particularly in achieving consistent results across diverse populations. The framework discussed in [19] relies on multiple feature scaling methods and classifiers may increase computational complexity, limiting its practical application in real-time settings. Rasul et al. [21] face the challenge of tuning hyperparameters for multiple ML models, which can be time-consuming and computationally intensive. Additionally, while SVM and LR showed high accuracy, generalization to real-world scenarios with diverse datasets might limit the applicability of their models. Ali et al. [22] encounter the challenge of effectively using MRI data and AI models to classify ASD severity, as their method relies heavily on cortical region features, which may not always be universally applicable or reliable across diverse individuals.

Kumar et al. [23] discuss the challenge of feature selection in ML models, particularly when dealing with missing data or incomplete datasets. Additionally, computational complexity and the need for high-quality, large-scale datasets are key challenges in automating ASD diagnosis. Karampasi et al. [24] face challenges in accurately classifying ASD from rs-fMRI data due to variability in feature selection and classification performance, as well as the limitations of the adopted features in providing sufficient discriminatory power. Rodrigues et al. [25] highlight the challenge of classifying ASD severity using rs-fMRI data, as functional differences in brain regions are subtle and may not provide clear-cut separations between ASD sub-classes.Sachdeva et al. [26] encounter difficulties in extracting significant features from fMRI data and ensuring that the logistic regression model accurately identifies key attributes for classifying ASD and TD individuals, as this approach relies heavily on the quality and relevance of the features. Kang et al. [27] face challenges in applying multi-site data for ASD classification, as variability in data sources and preprocessing methods can reduce the overall model accuracy. Achieving consistency across diverse datasets remains a significant hurdle. Alotaibi et al. [28] highlight the difficulty in using EEG-based phase-locking values for classification, as the approach requires a high level of signal processing expertise and may not be universally applicable across all types of data. Chib et al. [29] note that despite using advanced neuroimaging data, existing classification techniques often struggle to achieve high accuracy, which limits their potential for early-stage ASD detection, especially with the relatively small dataset used in their study.Shi et al. [30] face challenges in selecting relevant features for ASD classification using the MST-based method, as the algorithm's reliance on node-weighting and feature redundancy may not always lead to the most effective features for classification, limiting its overall effectiveness.

B. Deep Learning based Studies

The second subsection of the literature review focuses on deep learning-based methods. Despite their advantages over traditional methods, these deep learning (DL) approaches face potential challenges. For instance, in [31], the authors encounter difficulties in using sparse auto encoders (SAEs) for feature extraction, as high-dimensional data like the ABIDE dataset may lead to information loss, which could impact the classifier's performance and result in over fitting. Similarly, Shao et al. [32] face challenges when using the combined DFS and GCN method for ASD classification, as dimensionality reduction via DFS might not capture all relevant features, potentially leading to suboptimal performance. Mittal et al. [33] face difficulties with their DL-ASD framework for ASD prediction in children aged 1-10. There I-CNN approach for emotion prediction and deep learning for ASD classification may struggle to account for the complex emotional expressions of children in different age groups. Das et al. [34] encounter challenges in the dynamic analysis of feature-based distances, which may not generalize well across different datasets or populations, leading to variability in performance. Additionally, the high computational cost of their feature extraction technique could limit its scalability, particularly in clinical settings with larger datasets. Zhang et al. [35] highlight that their autoencoder model's sensitivity to Gaussian noise reduces the robustness of the feature extraction process, which becomes problematic when working with noisy data.

The shared-weight mechanism used in their model may fail to capture intricate temporal patterns, which are crucial for identifying subtle brain activity changes in ASD patients. Gupta et al. [36] discuss that their quantized CNN model for fMRI data analysis improves privacy and robustness but may not fully capture the fine-grained features essential for accurate ASD classification. Sravani et al. [37] face challenges when using DCNN with DTPSO optimization for ASD detection using sMRI data. The challenges lie in the need for careful optimization, and while the model achieves high accuracy and interpretability, fine-tuning may still be required for real-world applications. Sai et al. [38] encounter limitations with their speech-based ASD detection model using STFT-layered CNN. The reliance on STFT may hinder the model's ability to capture more complex relationships in the time-frequency data, making it less generalizable across diverse real-world signals. Qiu et al. [39] discuss challenges with their CNN-SVM model integrating fMRI data and the social responsiveness scale (SRS) for ASD classification. While this hybrid model provides good results, integrating the fMRI data with SRS poses challenges due to the differing nature of the data sources. Shahzad et al. [40] highlight that their transfer learning model, which leverages ResNet101 and EfficientNet, faces challenges when working with facial recognition, particularly when subtle ASD features are not adequately captured due to variability in facial images. Umrani et al. [46] note that their hybrid CNN-LSTM model, designed for multi-modal ASD classification, faces challenges in performance when applied to heterogeneous datasets with significant subject characteristic variations. While achieving high classification accuracy, their model may lack the interpretability needed for clinical decision-making. Al-Qazzaz et al. [42] point out that their transfer learning approach for ASD detection heavily depends on the quality of EEG data. Poor-quality signals can degrade predictive accuracy, and the model's performance requires meticulous hyperparameter tuning.

Khan et al. [43] mention that their ResNet-based deep learning model may struggle with feature extraction when dealing with noisy or irrelevant features. Additionally, LSTM-based approaches in this context can suffer from the vanishing gradient problem, making them less effective in dealing with long-term dependencies in time-series data. Khan et al. [44] observe that their hybrid CNN-LSTM model, while effective for multi-scale feature extraction, could suffer from overfitting, especially when trained on small datasets. Ulaganathan et al. [45] discuss challenges in using DNN for classification of ASD from brain imaging data, noting that although DNN models perform well on fMRI datasets, they may not capture all relevant brain features, especially in complex or atypical cases. Liu et al. [46] face challenges with their GAN-based anomaly detection approach for ASD diagnosis. While GANs can generate realistic images, capturing non-sequential dependencies remains a struggle, which limits the model's ability to detect subtle brain anomalies in ASD patients.

C. Data Mining based Methods

The existing methods for ASD detection using data mining approaches, while innovative, face several challenges. The computational complexity of method discussed in [51] could be a significant drawback, particularly when dealing with large datasets. The two-phase outlier rejection process (Quick Rejection Stage and Precise Rejection Stage) could be time-consuming, making it impractical for real-time applications. Additionally, there is a risk of over-rejecting data, potentially discarding valuable information that could aid in the detection of ASD. Furthermore, the effectiveness of the Fisher Score for feature selection is highly dependent on the characteristics of the data, and its performance may vary across different ASD datasets, especially when dealing with non-linear relationships or complex interactions between features. Ayyadurai et al. [52] applied a random forest (RF) classification-based method for ASD detection, reporting an average accuracy of 88%. While RF is a powerful classifier, one limitation is that it may not always capture complex relationships between features, especially in cases where interactions between variables are non-linear. Additionally, RF models can suffer from overfitting if the number of trees is not carefully tuned, reducing the model's generalizability to unseen data.

Shuvo et al. [53] also used RF for classification, achieving an overall accuracy of 96%, which suggests a highly effective approach. However, the high accuracy could be a result of overfitting, especially when the dataset is imbalanced or when the model is trained on a limited set of features. The RF method may also struggle with interpreting feature importance in cases where multiple features interact, and this lack of interpretability can be a challenge when deploying such models in clinical settings where transparency is crucial. The method disusssed by Latkowski et al. [54] may fail to capture the full complexity of ASD, as the condition is influenced by a wide range of genetic, environmental, and epigenetic factors that cannot always be captured by gene expression profiles alone. Additionally, the reliance on feature selection methods like Principal Component Analysis (PCA) could lead to the loss of important information, particularly when the relationship between features is non-linear. Moreover, while the study demonstrates a fusion system for selecting top-ranked genes, there is still the challenge of generalizing the findings to different populations, as genetic markers for ASD may vary across different groups.

Tawhid et al. [55] suffer from the issue of the computational complexity involved in generating spectrograms and performing time-frequency analysis, which can be resource-intensive. The process of filtering signals, segmenting them, and generating spectrogram images could be time-consuming, particularly when working with large-scale datasets, limiting the feasibility of real-time processing in clinical environments. Moreover, while texture-based features and principal component analysis (PCA) can capture important patterns in the data, they may not always fully account for complex temporal dependencies in the EEG signals, which could lead to a loss of critical information. Lastly, the use of machine learning classifiers introduces the issue of model interpretability, as deep learning methods and ensemble classifiers like KNN may be seen as "black-box" models, making it difficult for clinicians to understand and trust the model's decision-making process.

IV. DATASET DETAILS AND PERFORMANCE MEASUREMENT PARAMETERS

A. Dataset Details

I) UCI Dataset

This dataset was obtained from publically available UCI machine learning repository [47], [48], [49]. It consists of three separate sets for children, adolescents and adults stored as .csv files. The pediatric autism dataset comprises 292 cases, with four cases containing missing values. To address this, missing values were replaced with the standard deviation, effectively resolving the issue. Similarly, the adolescent autism dataset includes 104 cases, featuring categorical attributes akin to those in the paediatric dataset. The adult autism dataset consists of 704 cases, each containing two missing values, which were also handled using the standard deviation substitution method. After processing, the three datasets were combined into a single dataset containing 1,100 cases, all of which are now complete. Further details about the dataset are presented in Table 1.

Table 1. Overview of UCI dataset for child, adolescent and adult

"Characteristics	Child (C)	Adolescent (Ad)	Adult (A)	Merged (C + Ad + A)
# of Instances	292	104	704	1100
# of Features	14	14	14	14
Class Distribution	Yes = 141 No =151	Yes =63 No = 41	Yes =189 No = 515	Yes = 393 No = 707
# of Missing Values	4	NA	2	NA
Age	4–12 yrs	13–17 yrs	18–65 yrs	4–65 yrs
Feature Type	Categorical"			

II) Imaging Dataset

This study utilized data from the ABIDE I and ABIDE II datasets [50], which include 1,112 and 1,144 resting-state scans from over 24 international sites, respectively.

Datasets	Sites	Sample size (total)	Sample size		Sample size				Measurements
			TC	ASD	TC		ASD		
					Female	Male	Female	Male	
ABIDE I	Pitt	2	1	1	–	1	1	–	200
	Olin	1	1	–	–	1	–	–	210
	OHSU	11	7	4	–	7	–	4	82
	SDSU	1	1	–	1	–	–	–	180
	USM	3	3	–	–	3	–	–	240
	KKI	26	17	9	6	11	2	7	156
	NYU	40	17	23	6	11	1	22	180
	Stanford	24	12	12	3	9	2	10	180
	UCLA	8	3	5	–	3	2	3	120
	Total	116	62	54	16	46	8	46	–
ABIDE II	EMC	51	26	25	5	21	5	20	160
	GU	41	25	16	16	9	3	13	152
	IP	9	4	5	3	1	2	3	85
	KKI	90	64	26	28	36	9	17	156
	NYU1	53	20	33	1	19	4	29	180
	NYU2	27	–	27	–	–	3	24	180
	OHSU	42	32	10	16	16	3	7	120
	SDSU	11	4	7	–	4	1	6	180
	TCD	1	–	1	–	–	–	1	210
	UCLA	17	12	5	4	8	1	4	120
	USM	1	–	1	–	–	–	1	240
	Total	343	187	156	73	114	31	125	–
Combination of ABIDE I and ABIDE II	Total	459	249	210	89	160	39	171	–

Participants aged 5 to 10 years were selected, as early diagnosis of ASD is more effective within this age range. Based on these criteria, 116 individuals (62 typical control [TC] and 54 ASD) were selected from ABIDE I, and 343 individuals (187 TC and 156 ASD) were chosen from ABIDE II. Table 1 provides a summary of the participants' information.

B. Performance Measurement Parameters

The performance of these classification models is assessed in terms of Precision, Recall, Classification Accuracy, and F1-Score. The performance is measured with the help of confusion matrix. The depiction of the confusion matrix is given in below table:

Table 2. Confusion matrix

	Positive	Negative
Positive	T_P	F_P
Negative	F_N	T_N

Table 3. Performance measurement parameters

"Parameter	Formula
PrecisionP_R	$\frac{TP}{TP+FP}$
Recall(R)	$\frac{TP}{TP+FN}$
Classification Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
F1-Score	$2*\frac{P_R*R}{P_R+R}$

V. CONCLUSION AND FUTURE WORK

The application of ML and DL techniques in the detection and diagnosis of ASD has shown significant promise. By leveraging advanced neuroimaging techniques, such as fMRI, EEG, and MRI, these technologies are enabling the mining of complex patterns and features from brain activity that are often not discernible through traditional diagnostic methods. The ability of DL models, such as CNNs and RNNs, to perform automatic feature extraction and classification with minimal pre-processing has revolutionized ASD detection. This approach not only improves diagnostic precision but also decreases the time complexity for diagnosis, facilitating earlier intervention and improving long-term outcomes for individuals with ASD. Despite these advancements, challenges remain in terms of data accessibility, model interpretability, and the need for large, diverse datasets to train robust models. Furthermore, the integration of these technologies into clinical practice requires careful consideration of ethical implications and regulatory standards to ensure widespread adoption and effectiveness. While machine learning and deep learning models have made significant strides in ASD diagnosis, there are several areas for future research and development. Despite their success, deep learning models are often seen as "black boxes," making it difficult for clinicians to understand how decisions are made. Future research should focus on enhancing the interpretability of these models, ensuring that the insights generated are actionable and transparent. Transfer learning techniques, which leverage pre-trained models on one dataset and apply them to another, could be explored to improve the efficiency of training models for ASD detection. However, the integration of machine learning and deep learning for ASD diagnosis holds great potential, ongoing research is needed to overcome current limitations and maximize the impact of these technologies.

REFERENCES

1. Gupta, R., Advani, D., Yadav, D., Ambasta, R. K., & Kumar, P. (2023). Dissecting the relationship between neuropsychiatric and neurodegenerative disorders. Molecular Neurobiology, 60(11), 6476-6529.
2. Wang, L., Wang, B., Wu, C., Wang, J., & Sun, M. (2023). Autism spectrum disorder: neurodevelopmental risk factors, biological mechanism, and precision therapy. International journal of molecular sciences, 24(3), 1819.
3. Hus, Y., & Segal, O. (2021). Challenges surrounding the diagnosis of autism in children. *Neuropsychiatric disease and treatment*, 3509-3529.
4. Mansour, V. M. A., & Gobrial, E. S. (2022). Special Education Practices for Children with Autism Spectrum Disorders in the UK and Egypt: A comparative Study. International Journal of Research in Educational Sciences (IJRES), 5(1).
5. Reed, D. (2021). Autism spectrum disorder in veterinary clients: How the practice can help. *Veterinary Nursing Journal*, *36*(1), 30-32.
6. Mustafa, M. M. E. (2024). Future Guides for Biomarkers for Autism Spectrum Disorders. *Hail Journal of Health Sciences*, *6*(2), 85-91.
7. Tseng, A., Camchong, J., Francis, S. M., Mueller, B. A., Lim, K. O., Conelea, C. A., & Jacob, S. (2022). Differential extrinsic brain network connectivity and social cognitive task-specific demands in Autism Spectrum Disorder (ASD). Journal of Psychiatric Research, 148, 230-239.
8. Toh, T. H., Tan, V. W. Y., Lau, P. S. T., & Kiyu, A. (2018). Accuracy of Modified Checklist for Autism in Toddlers (M-CHAT) in detecting autism and other developmental disorders in community clinics. Journal of autism and developmental disorders, 48, 28-35.
9. Noruzman, A. H. B., Ghani, N. A., & Zulkifli, N. S. A. (2022). A Comparative Study on Autism Among Children Using Machine Learning Classification. *In Proceedings of International Conference on Emerging Technologies and Intelligent Systems: ICETIS 2021 Volume 2* (pp. 131-140). Springer International Publishing.
10. Talari, S., Balaji, K., & Stansfield, A. J. (2017). What is the association between ADI-R scores and final diagnosis of autism in an all IQ adult autism diagnostic service?. *Advances in Autism*, *3*(4), 250-262.
11. Clyde, A., Bismar, D., Agnew, G., & Kuper, L. E. (2024). Autism spectrum disorder and anxiety among Transgender Youth: Use of the Social Communication Questionnaire (SCQ). Journal of Autism and Developmental Disorders, 54(2), 615-627.
12. Valizadeh, A., Moassefi, M., Nakhostin-Ansari, A., Oskoie, I. M., Some'eh, S. H., Aghajani, F., ... & Memari, A. H. (2021). Accuracy of machine learning algorithms for the diagnosis of autism spectrum disorder based on cerebral sMRI, rs-fMRI, and EEG: protocols for three systematic reviews and meta-analyses. *medRxiv*, 2021-06.

13. Liu, M., Li, B., & Hu, D. (2021). Autism spectrum disorder studies using fMRI data and machine learning: a review. *Frontiers in Neuroscience*, *15*, 697870.
14. Kanchana, A., & Khilar, R. (2024). A Survey on Genetic Disease– Autism Spectrum Disorder Prediction and Classification in Machine Learning. International Journal of Nutrition, Pharmacology, Neurological Diseases, 14(4), 437-453.
15. Kurniawan, W. Y., & Gunawan, P. H. (2024). Classification of Autism Spectrum Disorder Based on Facial Images Using the VGG19 Algorithm. Journal of Computing Science and Engineering, *18*(1), 1-9.
16. Geeta, R. B., Totad, S. G., Reddy, P., &Shobha, R. B. (2015). Big data structure and usage mining coalition. International Journal of Services Technology and Management, 21(4/5), 6.
17. Shashikumar G. Totad, Geeta R. B. , Chennupati R Prasanna , N Krishna Santhosh , PVGD Prasad Reddy, " Scaling Data Mining Algorithms to Large and Distributed Datasets" , International Journal of Database Management Systems (IJDMS), Vol.2, No.4, November 2010.
18. Lamani, M. R., & Pernabas, J. B. (2024). A Thorough Review of Deep Learning in Autism Spectrum Disorder Detection: From Data to Diagnosis. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 17(8), 73-91.
19. Farooq, M. S., Tehseen, R., Sabir, M., & Atal, Z. (2023). Detection of autism spectrum disorder (ASD) in children and adults using machine learning. scientific reports, 13(1), 9605.
20. Vakadkar, K., Purkayastha, D., & Krishnan, D. (2021). Detection of autism spectrum disorder in children using machine learning techniques. SN computer science, 2, 1-9.
21. Hasan, S. M., Uddin, M. P., Al Mamun, M., Sharif, M. I., Ulhaq, A., & Krishnamoorthy, G. (2022). A machine learning framework for early-stage detection of autism spectrum disorders. IEEE Access, 11, 15038-15057.
22. Devika Varshini, G., & Chinnaiyan, R. (2020). Optimized machine learning classification approaches for prediction of autism spectrum disorder. Ann Autism Dev Disord. 2020; 1 (1), 1001.
23. Rasul, R. A., Saha, P., Bala, D., Karim, S. R. U., Abdullah, M. I., & Saha, B. (2024). An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder. Healthcare Analytics, 5, 100293.
24. Ali, M. T., Gebreil, A., ElNakieb, Y., Elnakib, A., Shalaby, A., Mahmoud, A., ... & Elbaz, A. S. (2023). A personalized classification of behavioral severity of autism spectrum disorder using a comprehensive machine learning framework. Scientific reports, *13*(1), 17048.
25. Kumar, C. J., & Das, P. R. (2022). The diagnosis of ASD using multiple machine learning techniques. International Journal of Developmental Disabilities, 68(6), 973-983.
26. Karampasi, A., Kakkos, I., Miloulis, S. T., Zorzos, I., Dimitrakopoulos, G. N., Gkiatis, K., ... & Matsopoulos, G. (2020, December). A machine learning fMRI approach in the diagnosis of autism. In *2020* IEEE international conference on big data (big data) (pp. 3628-3631). IEEE.
27. Rodrigues, I. D., de Carvalho, E. A., Santana, C. P., & Bastos, G. S. (2022). Machine learning and rs-fMRI to identify potential brain regions associated with autism severity. *Algorithms*, *15*(6), 195.
28. Joshi, Y., Totad, S.G., Geeta, R.B., Prasad Reddy, P.V.G.D. (2018). "Mobile Agent-Based Frequent Pattern Mining for Distributed Databases." In: Bhalla, S., Bhateja, V., Chandavale, A., Hiwale, A., Satapathy, S. (eds) Intelligent Computing and Information and Communication. Advances in Intelligent Systems and Computing, vol 673. Springer, Singapore. https://doi.org/10.1007/978-981-10-7245-1_9.
29. Sachdeva, J., Mittal, R., Mehta, J., Jain, R., & Ranjan, A. (2024). Resolving autism spectrum disorder (ASD) through brain topologies using fMRI dataset with multi-layer perceptron (MLP). Psychiatry Research: Neuroimaging, 343, 111858.
30. Kang, L., Chen, J., Huang, J., & Jiang, J. (2023). Autism spectrum disorder recognition based on multi-view ensemble learning with multi-site fMRI. Cognitive Neurodynamics, 17(2), 345-355.
31. Alotaibi, N., & Maharatna, K. (2021). Classification of autism spectrum disorder from EEG-based functional brain connectivity analysis. Neural computation, 33(7), 1914-1941.
32. Chib, A. S., Malhotra, D., & Mengi, M. (2023, June). A Machine Learning Approach for Autism Spectrum Disorder Detection using BOLD-fMRI Signals and ABIDE-II Dataset. In *2023* International Conference on Computer, Electronics & Electrical Engineering & their Applications *(IC2E3)* (pp. 1-4). IEEE.
33. Shi, C., Zhang, J., & Wu, X. (2020). An fMRI feature selection method based on a minimum spanning tree for identifying patients with autism. Symmetry, *12*(12), 1995.
34. Almuqhim, F., & Saeed, F. (2021). ASD-SAENet: a sparse autoencoder, and deep-neural network model for detecting autism spectrum disorder (ASD) using fMRI data. Frontiers in Computational Neuroscience, 15, 654315.
35. Shao, L., Fu, C., You, Y., & Fu, D. (2021). Classification of ASD based on fMRI data with deep learning. *Cognitive Neurodynamics*, *15*(6), 961-974.
36. Mittal, R., Malik, V., & Rana, A. (2022, December). DL-ASD: A Deep Learning Approach for Autism Spectrum Disorder. In 2022 5th International Conference on Contemporary Computing and Informatics (IC3I) (pp. 1767-1770). IEEE.
37. Das, S., Chaudhury, P., & Tripathy, H. K. (2024, January). Classification of Autism Spectrum Disorder using CNN & Transfer Learning. In *2024* International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC) (pp. 1-7). IEEE.
38. Zhang, J., Guo, J., Lu, D., & Cao, Y. (2024). ASD-SWNet: a novel shared-weight feature extraction and classification network for autism spectrum disorder diagnosis. *Scientific Reports*, *14*(1), 13696.

39. Gupta, S., Bhuiyan, M. R. I., Chowa, S. S., Montaha, S., Rahman, R., Mehedi, S. T., & Rahman, Z. (2024). Enhancing Autism Spectrum Disorder Classification with Lightweight Quantized CNNs and Federated Learning on ABIDE-I Dataset. *Mathematics*, *12*(18), 2886.
40. Sravani, K., & Kuppusamy, P. (2024). Optimized Deep Convolutional Neural Network for Autism Spectrum Disorder Detection Using Structural MRI and DTPSO. IEEE Access.
41. Sai, K. V. K., Krishna, R. T., Radha, K., Rao, D. V., & Muneera, A. (2024). Automated ASD detection in children from raw speech using customized STFT-CNN model. International Journal of Speech Technology, 27(3), 701-716.
42. Qiu, L., & Zhai, J. (2024). A hybrid CNN-SVM model for enhanced autism diagnosis. *Plos one*, *19*(5), e0302236.
43. Shahzad, I., Khan, S. U. R., Waseem, A., Abideen, Z. U., & Liu, J. (2024). Enhancing ASD classification through hybrid attention-based learning of facial features. Signal, Image and Video Processing, 1-14.
44. Umrani, A. T., & Harshavardhanan, P. (2024). Efficient classification model for anxiety detection in autism using intelligent search optimization based on Deep CNN. Multimedia Tools and Applications, 1-30.
45. ManoharMadagi, ShantalaGiraddi, Geeta R.B, Madhur, "Brain Tumor Classification and Segmentation using Deep Learning" SMART COMPUTING TECHNIQUES & INFORMATICS (SCI 2020), Smart Computing Techniques and Applications, Springer Series, 2021, Volume 224, Smart Innovation, Systems and Technologies Smart Innovation, Systems and Technologies(SIST), book series pp201-208, 14th July 2021, Scopus Indexed
46. Al-Qazzaz, N. K., Aldoori, A. A., Buniya, A., Ali, S. H. B. M., & Ahmad, S. A. (2024). Transfer Learning and Hybrid Deep Convolutional Neural Networks Models for Autism Spectrum Disorder Classification from EEG Signals. *IEEE Access*.
47. Totad, S.G., Geeta, R.B. & Prasad Reddy, P.V.G.D. Batch incremental processing for FP-tree construction using FP-Growth algorithm. KnowlInfSyst 33, 475–490 (2012). https://doi.org/10.1007/s10115-012-0514-9
48. Khan, K., & Katarya, R. (2024). AFF-BPL: An adaptive feature fusion technique for the diagnosis of autism spectrum disorder using Bat-PSO-LSTM based framework. Journal of Computational Science, 83, 102447.
49. Khan, K., & Katarya, R. (2024). MCBERT: A Multi-Modal Framework for the Diagnosis of Autism Spectrum Disorder. Biological Psychology, 108976.
50. Ulaganathan, S., Harshavardhanan, P., Raju, N. G., & Parthasarathy, G. (2024). Hybrid Optimization Enabled DenseNetforAutism Spectrum Disorders Using MRI Image. *Computational Biology and Chemistry*, 108335.
51. Liu, X., Hasan, M. R., Gedeon, T., & Hossain, M. Z. (2024). MADE-for-ASD: A multi-atlas deep ensemble network for diagnosing Autism Spectrum Disorder. *Computers in Biology and Medicine*, *182*, 109083.
52. Thabtah, F.F. (2017). Autistic Spectrum Disorder Screening Data for children
53. Thabtah, F.F. (2017). Autistic Spectrum Disorder Screening Data for Adolescent.
54. F.F. Thabtah Autistic spectrum disorder screening data for adult Manukau Inst. Technol (2017) Available at: http://fcon_1000.projects.nitrc.org/indi/abide/
55. Saleh, A. I., & Rabie, A. H. (2023). A new autism spectrum disorder discovery (ASDD) strategy using data mining techniques based on blood tests. *Biomedical Signal Processing and Control*, *81*, 104419.
56. Ayyadurai, M., Sujatha, K., Deeptha, R., Preethi, D., & Dinesh, M. (2024, June). Exploring Data Mining Techniques for Early Autism Detection using Random Forest Algorithm. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.
57. Shuvo, S. B., Ghosh, J., & Oyshi, A. S. (2019, July). A data mining based approach to predict autism spectrum disorder considering behavioral attributes. In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE.
58. Latkowski, T., & Osowski, S. (2015). Data mining for feature selection in gene expression autism data. *Expert Systems with Applications*, *42*(2), 864-872.
59. Tawhid, M. N. A., Siuly, S., Wang, K., & Wang, H. (2021). Data mining based artificial intelligent technique for identifying abnormalities from brain signal data. In *Web Information Systems Engineering–WISE 2021: 22nd International Conference on Web Information Systems Engineering, WISE 2021, Melbourne, VIC, Australia, October 26–29, 2021, Proceedings, Part I 22* (pp. 198-206). Springer International Publishing.
60. Prasenjit, Gokul, Manish (2023). Study of Machine Learning Models to Detect Autism Spectrum Disorder. IRJCS::International Research Journal of Computer Science, Volume 10, Issue 08 of 2023 pages 441-451.
61. Prasenjit, Gokul, Manish (2023). Study of Machine Learning Techniques in Autism Spectrum Disorder Detection. IRJCS::International Research Journal of Computer Science, Volume 10, Issue 08 of 2023 pages 452-462.
62. Rajanishree, Harshitha, Prabha, swathi, apoorva (2020). An Efficient Method for Prediction of Autism Spectrum Disorder. International Research Journal of Computer Science (IRJCS), Volume VII, 123-126.