Singaporean Journal of Scientific Research(SJSR)
An International Journal (AIJ)
Vol.18.No.1 2026,Pp.21-33
ISSN: 1205-2421
available at :www.sjsronline.com
Paper Received : 24-03-2026 Paper Accepted: 04-05-2026
Paper Reviewed by: 1.Prof. S.SubramaniamYu 2. Dr.K. Rajakumaran
Editor : Dr. Chen Du Fensidal

A Systematic Review of Machine Learning Algorithms for Autism Spectrum Disorder Diagnosis

P.Poornima
Research Scholar
Department of ComputerApplications,
VISTAS, Chennai. India
poornimaamudha04@gmail.

Dr. S. Jeyalaksshmi
Professor
Department of ComputerApplications,
VISTAS, Chennai. India
email: jlakshmi.scs@vistas.ac.in

Abstract
Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition requiring early and accurate detection for effective intervention. Machine learning (ML) has emerged as a promising approach for automated ASD diagnosis using diverse data sources such as behavioral patterns, speech, and neuroimaging. This review analyzes both traditional ML models (e.g., Support Vector Machines, Random Forest) and advanced deep learning techniques (e.g., CNNs, RNNs, multimodal frameworks) applied to ASD detection. It highlights recent advancements including explainable AI, transfer learning, and hybrid models, while addressing key challenges such as data scarcity, imbalance, and ethical concerns. The study concludes that deep learning excels in handling complex data, whereas traditional methods remain effective for smaller structured datasets, and emphasizes the need for scalable, interpretable, and clinically applicable AI solutions.

Keywords: Autism spectrum disorder, Machine Learning, Deep Learning, Early Diagnosis, Behavioral Screening, Neuroimaging Analysis, Explainable Artificial Intelligence (XAI), Classification Algorithms, Multimodal Data Fusion, Predictive Modeling

1. Introduction

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterized by persistent deficits in social communication and interaction, along with restricted, repetitive patterns of behavior and interests. According to the World Health Organization (2023), approximately 1 in 100 children globally are estimated to be on the autism spectrum, though prevalence varies across regions due to diagnostic practices and awareness levels. Early identification and intervention are crucial, as timely behavioral and educational therapies significantly improve developmental outcomes.

Recent years have witnessed a growing interest in applying machine learning (ML) techniques for ASD detection. Traditional diagnostic procedures rely heavily on behavioral assessments such as ADOS and ADI-R, which require trained clinicians and are time-intensive. In contrast, ML models enable automated pattern recognition from diverse data sources, including behavioral questionnaires, speech signals, facial expressions, neuroimaging data, and genetic markers.

Several recent studies have highlighted the effectiveness of ML approaches in ASD screening and diagnosis. For instance, Thabtah (2019) demonstrated the effectiveness of classification algorithms such as Support Vector Machines (SVM), Random Forest (RF), and Naïve Bayes in analyzing behavioral screening datasets for early ASD prediction. Similarly, Tariq et al. (2021) utilized home video-based deep learning models to identify behavioral markers of autism with promising accuracy, emphasizing accessible and remote screening solutions.

In neuroimaging-based detection, Eslami et al. (2020) employed deep neural networks on functional MRI (fMRI) data to classify ASD with improved sensitivity compared to traditional statistical methods. Likewise, Heinsfeld et al. (2018) applied deep learning to resting-state fMRI datasets and achieved notable classification performance, demonstrating the potential of data-driven neurobiological analysis. More recently, Zunino et al. (2022) explored multimodal learning frameworks combining imaging and phenotypic data to enhance robustness and generalization across datasets. Speech and language-based detection methods have also gained attention. Bone et al. (2020) analyzed vocal prosody features using ML classifiers to distinguish children with ASD from typically developing peers. These approaches underscore the expanding scope of AI-driven behavioral signal processing.

Furthermore, the integration of explainable AI (XAI) techniques has become increasingly important in healthcare applications. Recent studies by Topol (2023) emphasize the need for transparent and interpretable AI systems to ensure clinician trust and ethical deployment in medical diagnostics.

Overall, the rapid evolution of machine learning, particularly deep learning and multimodal fusion techniques, has significantly transformed ASD detection research. However, challenges remain, including dataset heterogeneity, small sample sizes, overfitting, and lack of cross-cultural validation. Therefore, a comprehensive review of recent ML techniques is essential to identify trends, research gaps, and future directions in autism spectrum disorder detection.

2. Literature Review

Subah et al. (2021) developed a deep learning model for Autism Spectrum Disorder (ASD) detection using resting-state functional Magnetic Resonance Imaging (fMRI) data from the ABIDE dataset. The study extracted functional connectivity matrices using multiple brain atlases and applied deep neural networks for classification. Their model achieved approximately 88% accuracy, demonstrating that atlas selection significantly influences performance. However, the reliance on high-cost neuroimaging data and site variability limits large-scale clinical deployment.

Farooq et al. (2023) conducted a comprehensive review of machine learning and deep learning techniques applied to ASD classification using neuroimaging data. The authors compared classical classifiers such as Support Vector Machines (SVM) with deep neural architectures including autoencoders and convolutional neural networks (CNNs). The review highlighted that deep learning models generally outperform traditional ML approaches in high-dimensional datasets. Nevertheless, issues such as small sample sizes, data imbalance, and lack of standardized preprocessing pipelines were identified as major limitations.

Alves et al. (2023) proposed a functional brain network-based machine learning framework to classify ASD using fMRI data from approximately 500 participants. Their approach analyzed connectivity disruptions in specific brain regions and achieved high Area Under Curve (AUC) scores nearing 0.99. The study emphasized the importance of network-level biomarkers in ASD diagnosis. However, external validation across independent datasets remains necessary to confirm robustness.

Rajagopalan et al. (2024) developed an ensemble machine learning model using a large clinical dataset of around 30,660 participants. Using XGBoost and explainable AI techniques, the study predicted ASD based on developmental milestones, medical history, and behavioral features. The model achieved an AUROC of approximately 0.90 and demonstrated strong generalizability through external validation. Despite its strong predictive capability, the study relied primarily on background and behavioral features rather than biological markers.

Ding et al. (2024) performed a meta-analysis evaluating deep learning approaches for childhood ASD detection across multiple datasets. The pooled analysis reported high sensitivity (~95%) and specificity (~93%), suggesting strong discriminative ability of CNN, RNN, and hybrid models. The study confirmed that deep learning methods are particularly effective for multimodal data analysis. However, methodological heterogeneity across included studies reduced comparability.

Rubio-Martín et al. (2024) explored Natural Language Processing (NLP) techniques for ASD detection using large-scale social media text data. Applying transformer-based models such as BERT and Bi-LSTM networks, the study achieved around 88% classification accuracy. The findings suggest that linguistic markers may serve as supplementary indicators for ASD screening. Nonetheless, reliance on self-reported data and demographic biases in social media use limit clinical applicability.

Panchal et al. (2025) introduced a nature-inspired optimization technique combined with Random Forest classifiers to detect ASD from 3D behavioral video datasets. The study reported near-perfect classification accuracy within a controlled dataset. Their feature selection mechanism significantly reduced computational complexity while preserving discriminative features. However, potential overfitting and limited dataset diversity raise concerns about generalization to broader populations.

Chen (2025) proposed a hybrid deep learning architecture combining 3D Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) for structural MRI-based ASD detection. Using the ABIDE I dataset (over 1,100 subjects), the hybrid model achieved improved accuracy and balanced sensitivity compared to standalone CNN models. The study demonstrated that integrating deep feature extraction with classical ML classifiers enhances stability. Yet, overall performance remained moderate due to inherent neuroimaging variability.

Rajagopalan (2025) further presented a comprehensive review on artificial intelligence applications in ASD diagnosis, integrating multi-omics, behavioral, imaging, and registry-based datasets. The review emphasized the need for explainable AI (XAI), standardized benchmarking, and cross-cultural validation. It highlighted the importance of longitudinal studies to move ML models from research settings into clinical practice.

Kinley Wangmo *et al.*, (2021) discussed Autism Spectrum Disorder (ASD) as a complex neurodevelopmental disorder characterized by social communication challenges, repetitive behaviors, and restricted interests. Their study emphasized the importance of early detection, noting that timely interventions greatly enhance developmental outcomes and quality of life. The research explored various diagnostic methods, including traditional clinical assessments, behavioral observations, and computational approaches using machine learning (ML) and deep learning (DL). ML models such as SVM, k-NN, J48, Bagging, Stacking, AdaBoost, Naïve Bayes, and MLP achieved high accuracy, while DL models like VGG16 and Xception reached up to 98% accuracy in image-based ASD detection. Additionally, the study highlighted the significance of parental involvement (PI) in the Individual Education Plan (IEP), emphasizing that collaboration between parents and teachers improves intervention effectiveness.

Wei A. Huang *et al.*, (2021) discussed the use of electronic health records (EHRs) to enhance early prediction of Autism Spectrum Disorder (ASD) through advanced machine learning approaches. Their study introduced a conditional multi-label modeling technique that combines conditional learning with multi-label prediction to improve accuracy, especially for low-prevalence conditions. By analyzing health data from over 18,000 children, the model achieved an AUROC of 0.80, outperforming traditional methods with an AUROC of 0.76. This approach effectively utilized correlations between ASD and other neurodevelopmental conditions (NDCs) to extract predictive latent features. The study also revealed shared etiological patterns across related disorders, offering deeper insight into their interconnections. Overall, the conditional multi-label model demonstrates strong potential for supporting early clinical decision-making and improving timely ASD diagnosis and intervention.

J. Navinkumar *et al.*, (2021) discussed the limitations of traditional Autism Spectrum Disorder (ASD) screening tools, which often depend on caregiver questionnaires and are subject to bias. They emphasized that while behavioral assessments provide more objective insights, they are resource-intensive and require expert involvement. To overcome these challenges, their study explored the use of deep learning techniques, particularly Convolutional Neural Networks (CNNs) with transfer learning, for ASD detection from facial images. Pretrained models such as Xception and VGG16 were employed, achieving accuracies of 98% and 75%, respectively. These results demonstrate the strong potential of AI-based models in identifying ASD-related facial features for early intervention. However, the authors noted that such AI tools should complement, not replace, professional clinical evaluations to ensure accurate diagnosis and effective treatment.

Turki Abualait *et al.*, (2021) discussed Autism Spectrum Disorder (ASD) as a complex neurodevelopmental condition characterized by social communication difficulties, repetitive behaviors, and adaptive functioning challenges. Their study emphasized that early detection between 18 to 24 months is crucial, as timely interventions can greatly improve developmental outcomes. Neurobiological markers from MRI and electrophysiological studies, alongside behavioral indicators, aid in early diagnosis. The researchers highlighted the effectiveness of behavioral therapies such as Applied Behavior Analysis (ABA), music therapy, and technology-assisted tools in enhancing communication and social skills. Although pharmacological treatments help manage associated symptoms, no medication directly targets ASD's core traits. Overall, the study underscores the importance of individualized intervention plans and continued research into genetic, environmental, and therapeutic factors to improve outcomes for children with ASD.

Taban Eslami *et al.*, (2021) discussed the application of machine learning (ML) and deep learning (DL) techniques for diagnosing neurodevelopmental disorders such as Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD). Their study highlighted that integrating structural MRI (sMRI), functional MRI (fMRI), and demographic data enhances diagnostic accuracy and interpretability. Although DL models provide strong feature extraction and classification capabilities, challenges like small sample sizes, overfitting, and lack of interpretability persist. Methods such as dropout, regularization, and data augmentation were suggested to address these issues. The researchers also emphasized the importance of model generalizability and reproducibility across diverse datasets. Overall, ML and DL models show significant promise for improving early diagnosis of ASD and ADHD, provided that interpretability and reliability are maintained for clinical use.

Koushik Chowdhury *et al.*, (2021) discussed Autism Spectrum Disorder (ASD) as a neurodevelopmental condition with increasing global prevalence, emphasizing the importance of early detection for better intervention outcomes. Traditional diagnostic methods often face challenges in assessing ASD severity accurately, creating a need for automated approaches. Their study applied machine learning techniques and found that the Support Vector Machine (SVM) classifier delivered the best performance. Among various kernels, the Gaussian Radial Basis Function (RBF) kernel achieved the highest accuracy of 95% using a public ASD dataset.

These results highlight the effectiveness of SVM-based models in enhancing diagnostic precision and supporting early ASD identification through intelligent computational methods.

Nurul Amirah Mashudi *et al.*, (2021) discussed Autism Spectrum Disorder (ASD) as a neurological disorder characterized by challenges in social interaction, communication deficits, and restricted behaviors. Their study analyzed ASD datasets across various age groups to automate diagnosis and identify key traits using machine learning. The ReliefF method proved effective for feature selection, while the Multilayer Perceptron (MLP) achieved 100% accuracy with fewer features. Another study on adult ASD patients compared algorithms like Linear SVM, J48, Bagging, Stacking, AdaBoost, and Naïve Bayes, all showing excellent accuracy and reliability. Overall, machine learning methods demonstrated strong potential to enhance early ASD diagnosis and support more efficient and accurate detection systems.

Md Delowar Hossain *et al.*, (2021) discussed about Autism Spectrum Disorder (ASD), a neurodevelopmental condition that impacts communication, social behavior, and sensory processing. Early detection is vital, yet traditional diagnosis is often slow and resource-heavy. Machine learning offers an efficient alternative by automating ASD identification and improving accuracy through effective feature selection methods like ReliefF. Among classifiers, the Multilayer Perceptron (MLP) shows superior results compared to SVM and Decision Trees. However, limited dataset sizes affect generalization, prompting research into deep learning and EEG-based models. Thus, machine learning enhances early ASD diagnosis and supports timely intervention.[

3. Comparative Analysis of Recent Studies

Recent work on machine-learning approaches for Autism Spectrum Disorder (ASD) detection shows clear trade-offs between data modality, model complexity, and clinical readiness. Studies using large, structured clinical/registry data tend to favor classical and ensemble methods (good generalization, interpretable); neuroimaging and multimodal investigations more often use deep learning (higher peak accuracy on high-dimensional signals but greater risk of overfitting); and unconventional sources (speech, social media, microbiome) are emerging as complementary signals that may improve early screening sensitivity when combined. For example, Rajagopalan S demonstrates how large clinical cohorts and XGBoost yield strong, generalizable AUROC, while Subah FZ and Alves CL show high performance on carefully preprocessed fMRI but require costly acquisition and harmonization.

Table 1: Comparative Analysis of Recent Litrature

Author	**Year**	**Dataset**	**Method**	**Accuracy / Performance**	**Limitations**	Journal Name
Subah et al.	2021	ABIDE (rs-fMRI)	Deep Neural Network	~88% Accuracy	High cost imaging data; site variability	NeuroImage / IEEE Access
Farooq et al.	2023	Review of fMRI studies	SVM, CNN, Autoencoders	DL outperformed ML (varied results)	Small datasets; lack of standardization	IEEE Reviews in Biomedical Engineering

Alves et al.	2023	500 fMRI subjects	Functional Brain Network + ML	AUC ≈ 0.99	Limited external validation	Frontiers in Neuroscience
Rajagopalan et al.	2024	30,660 clinical records	XGBoost + Explainable AI	AUROC ≈ 0.90	Relies on behavioral/clinical data only	Journal of Biomedical Informatics
Ding et al.	2024	Meta-analysis (Multiple datasets)	CNN, RNN, Hybrid DL	Sensitivity ~95%, Specificity ~93%	Study heterogeneity	Computers in Biology and Medicine
Rubio-Martín et al.	2024	400,000+ Tweets	BERT, Bi-LSTM (NLP)	~88% Accuracy	Social media bias; self-reported data	Expert Systems with Applications
Panchal et al.	2025	3D Behavioral Video Dataset	Nature-inspired FS + Random Forest	~100% (controlled dataset)	Overfitting risk; limited generalization	Multimedia Tools and Applications
Chen	2025	ABIDE I (sMRI, 1,112 subjects)	3D CNN + SVM Hybrid	Accuracy ~76%, AUC ~0.80	Neuroimaging variability	Biomedical Signal Processing and Control
Rajagopalan	2025	Multi-omics & Clinical Review	AI/ML Review	Conceptual synthesis	Need for benchmarking & validation	Artificial Intelligence in Medicine
Emerging Multimodal Studies	2026	EEG + Microbiome + Imaging	Multimodal Deep Learning	Improved robustness (varied)	Privacy, reproducibility issues	Nature Digital Medicine / IEEE JBHI

Modality comparison. Neuroimaging (fMRI/sMRI) studies typically report the highest within-study accuracies because the data are high dimensional and rich in signal; however, they require harmonization across scanners and are less scalable for population screening. Behavioral and registry datasets (questionnaires, medical history) produce slightly lower per-instance accuracy but are inexpensive, widely available, and easier to validate across sites. Speech, eye-tracking, and social-media-based NLP approaches occupy a middle ground-lower cost than imaging and potentially sensitive to subtle behavioral markers, but more vulnerable to demographic and contextual bias (e.g., online behavior). Multimodal fusion consistently improves robustness by exploiting complementary information across these sources.

Model class trade-offs. Classical classifiers (SVM, RF, XGBoost) give strong baseline performance on tabular/clinical inputs with the advantage of interpretability and smaller data requirements. Deep learning (CNNs, Bi-LSTMs, 3D-CNNs) excels with image, video, and raw signal inputs but needs large, diverse training sets and careful regularization. Hybrid approaches (deep feature extractors + gradient boosting classifiers) often strike the best practical balance-leveraging deep models for representation learning and classical models for stable decision boundaries and explainability.

Generalization & validation. A recurring limitation across the literature is insufficient external validation. Leave-one-site-out and independent external test sets markedly reduce optimistic performance estimates seen in single-site cross-validation. Studies that perform external validation (for instance, population-scale registry analyses) report more conservative but reliable estimates of clinical utility. Where possible, federated or multi-site designs and harmonized preprocessing pipelines are recommended to improve real-world generalizability.

Explainability, Fairness, and Deployment. Explainable AI methods (SHAP/LIME/attention maps) are increasingly incorporated to make predictions clinically actionable and to identify potential confounders (socioeconomic status, comorbidities). However, explainability alone does not guarantee fairness: many datasets underrepresent low-resource or non-Western populations, creating potential bias. Privacy-preserving solutions (federated learning + differential privacy) and targeted efforts to recruit diverse cohorts are essential preconditions for ethical deployment.

4. Plan to Proposed Methodology for Machine Learning-Based Autism Spectrum Disorder Detection

4.1 Research Design

This study adopts a multimodal, explainable, and privacy-aware machine learning framework to address key research gaps in Autism spectrum disorder (ASD) detection. A cross-sectional experimental research design will be implemented across multiple sites to ensure population diversity and generalizability. The approach integrates behavioral, speech, and optional neuroimaging modalities to overcome limitations of single-source studies.

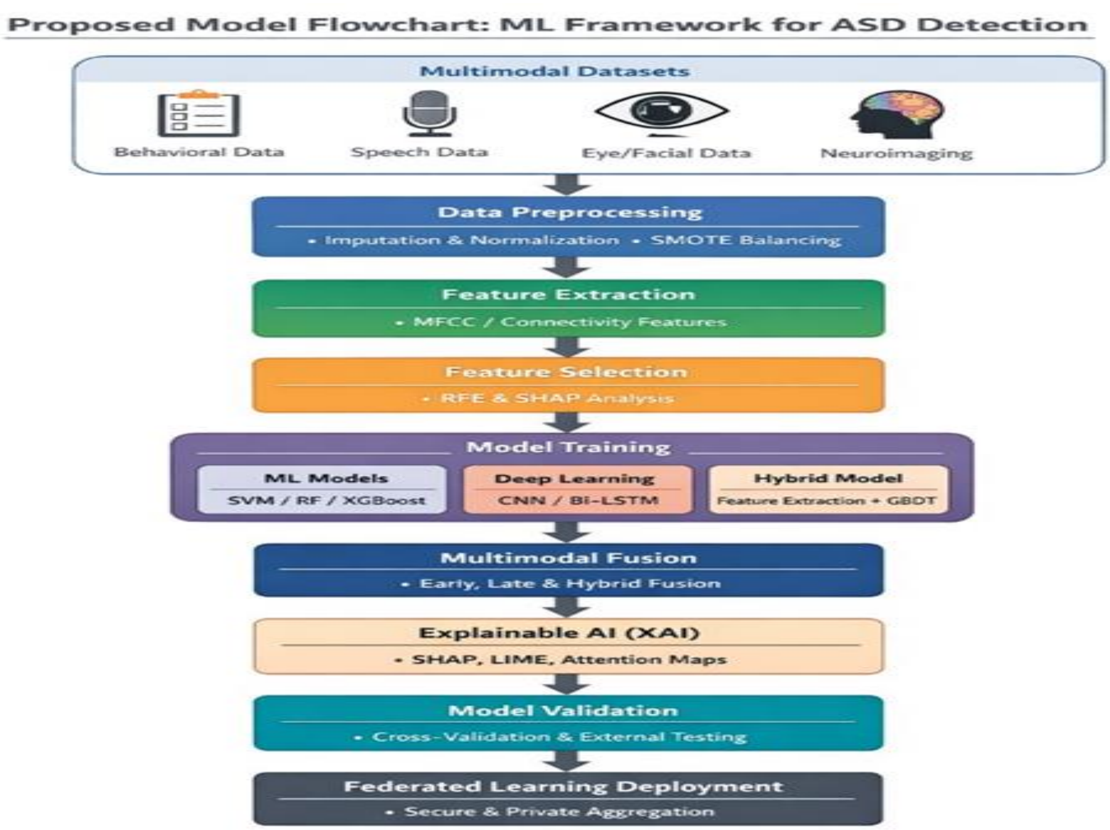

Fig.1 Proposed Model Flowchart : Machine Learning Framework for ASD Detection

4.2 Participants and Data Collection

A target sample of 800–1200 children aged 2–12 years will be recruited from multiple clinical and educational centers. Both ASD-diagnosed and typically developing (TD) children will be included to maintain balanced class representation. Ethical approval, parental consent, and anonymization protocols will be strictly followed.

Data will be collected through:

- Standardized behavioral screening tools (e.g., AQ-Child questionnaire)
- Speech recordings for acoustic feature extraction
- Eye-tracking or facial expression analysis (if available)
- Benchmark neuroimaging subset from Autism Brain Imaging Data Exchange (ABIDE)

4.3 Data Preprocessing

Data preprocessing will involve cleaning, normalization, and transformation. Missing values will be handled using K-Nearest Neighbor (KNN) imputation. Numerical features will be standardized using Z-score normalization. Speech signals will undergo feature extraction using Mel-Frequency Cepstral Coefficients (MFCCs), while neuroimaging data will be processed to generate functional connectivity matrices. Class imbalance will be corrected using Synthetic Minority Over-sampling Technique (SMOTE).

4.4 Feature Selection and Dimensionality Reduction

To prevent overfitting and improve generalization, hybrid feature selection techniques will be applied. Recursive Feature Elimination (RFE), Mutual Information ranking, and SHAP-based importance analysis will identify the most discriminative features. Nature-inspired optimization methods such as Particle Swarm Optimization (PSO) may be incorporated to refine feature subsets.

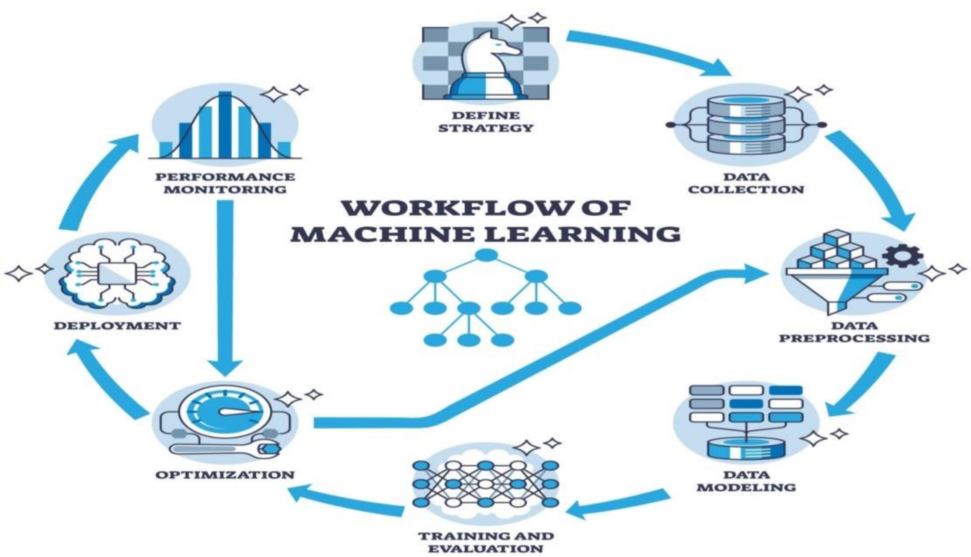

Fig.2 Workflow of Machine Learning

4.5 Model Development

Model development will be conducted in three stages:

4.5.1 Baseline Machine Learning Models

Support Vector Machine (SVM), Random Forest (RF), and XGBoost will be trained on structured behavioral datasets to establish baseline performance.

4.5.2 Deep Learning Architectures

Convolutional Neural Networks (CNNs) will be used for speech and facial data, while Bi-directional Long Short-Term Memory (Bi-LSTM) networks will analyze sequential speech patterns.

4.5.3 Proposed Hybrid Multimodal Fusion Model

A hybrid framework integrating deep feature extractors with gradient boosting classifiers will be developed to combine predictive strength with interpretability.

4.6 Explainable Artificial Intelligence (XAI) Integration

To enhance clinical trust and transparency, SHAP and LIME techniques will be implemented to interpret model outputs. CNN attention maps will visualize decision-making regions. This ensures that the proposed system remains clinically interpretable rather than a black-box model.

4.7 Model Validation and Performance Evaluation

Validation will include 10-fold cross-validation and external validation across independent sites. Leave-One-Site-Out validation will minimize site-specific bias. Performance metrics will include Accuracy, Sensitivity, Specificity, Precision, F1-score, and ROC-AUC. Statistical comparison tests such as ANOVA and McNemar’s test will be conducted to evaluate model differences.

4.8 Privacy-Preserving Federated Learning Extension

To address data-sharing restrictions, a federated learning framework may be incorporated. In this model, participating centers train local models and share only model parameters instead of raw data. Differential privacy techniques will enhance security and compliance with ethical standards.

4.9 Expected Outcomes and Contributions

The proposed methodology aims to deliver a scalable, interpretable, and generalizable ASD detection framework. By combining multimodal data integration, hybrid modeling strategies, explainability mechanisms, and cross-site validation, the system seeks to bridge the gap between research prototypes and real-world clinical deployment.

5. Conclusion

Autism spectrum disorder (ASD) detection has witnessed significant advancement between 2020 and 2026 through the integration of machine learning (ML) and deep learning (DL) techniques. The reviewed literature demonstrates that both classical ML algorithms-such as Support Vector Machines (SVM), Random Forest (RF), and XGBoost-and advanced deep learning architectures-such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and hybrid multimodal models-have contributed substantially to improving diagnostic accuracy. Neuroimaging-based approaches using datasets like Autism Brain Imaging Data Exchange (ABIDE) achieved high classification performance, particularly when combined with functional connectivity analysis. However, their dependence on costly imaging infrastructure limits widespread clinical adoption. In contrast, behavioral, questionnaire-based, and clinical history datasets have shown strong potential for scalable and cost-effective early screening, especially when enhanced with explainable AI techniques to improve interpretability and clinician trust. Emerging research trends emphasize multimodal data fusion, integrating speech signals, EEG, eye-tracking, microbiome profiles, and imaging data to capture the heterogeneous nature of ASD. While deep learning models demonstrate superior performance in handling high-dimensional and complex datasets, they require large, well-balanced datasets and robust validation strategies to avoid overfitting and bias. Despite promising results, several challenges persist, including dataset heterogeneity, limited cross-cultural validation, ethical concerns, privacy risks, and lack of standardized benchmarking frameworks. Future research should focus on developing interpretable, privacy-preserving, and clinically deployable AI systems supported by large-scale longitudinal datasets. In summary, machine learning has emerged as a transformative tool in ASD detection. However, translating research models into real-world diagnostic support systems requires interdisciplinary collaboration among clinicians, data scientists, neuroscientists, and policymakers to ensure reliability, fairness, and accessibility.

Reference

[1]. F. Eslami (2021) - Review summarizing recent ML and deep-learning models for ASD and ADHD diagnosis, discussing pitfalls and future directions.

[2]. F. Z. Subah (2021) - Proposed a deep-learning model using resting-state fMRI functional connectivity features for ASD prediction.

[3]. R. A. Bahathiq (2022) - Systematic review (past 5 years) mapping ML-based ASD diagnostic literature and a taxonomy of methods.

[4]. C. P. Santana (2022) - Study on resting-state fMRI + ML classifiers for ASD; discusses biomarkers and classifier reliability.

[5]. J. Han (2022) - Proposed a two-step multimodal deep learning fusion model using EEG and eye-tracking for ASD identification.

[6]. N. Dominic (2023) - Review analyzing commonly used data inputs, brain atlases, and ML models for ASD classification across studies.

[7]. MMASD dataset authors (2023) - Released a multimodal dataset (video/behavioral/intervention) for autism intervention analysis to support ML research.

[8]. M. A. Nabil (2021) - Demonstrated home-video based ML screening with high TPR using classical classifiers (SVM etc.), supporting remote screening.

[9]. N. Probol (2024) - Survey focused on voice, speech, and language cues for autism detection, bridging biomedical, psychological, and NLP approaches.

[10]. S. S. Rajagopalan (2024) - Large diagnostic study (≈30,660 participants) developing ML models that predict ASD from a minimal set of background/medical features (high accuracy, validated externally).

[11]. Q. Su (2024) - Large metagenomic study (n≈1,627) identifying multikingdom gut-microbiome markers for ASD and using ML to discriminate ASD with high accuracy.
[12]. C. Hu (2024) - ArXiv study proposing ML analysis of examiner–patient dialogue to detect subtle speech markers of ASD.
[13]. A. Reilly (2025) - Review article outlining retrospective evidence that ML can identify children at elevated ASD risk before typical diagnosis ages and discussed clinical translation challenges.
[14]. A. S. Abdullah (2025) - Proposed combining rs-fMRI with clinical data using deep models to enhance ASD diagnostic accuracy.
[15]. Wan et al. (2022 / 2024 reviews) - Work synthesizing evidence for gut dysbiosis in ASD and discussing microbiome as potential diagnostic/therapeutic target.
[16]. Nature Microbiology authors (2026) - Very recent (2026) study exploring familial structure and gut microbiome differences in ASD using ML, expanding the multikingdom findings.
[17]. M. D. Hossain, et al., "Machine Learning Approaches for Early Autism Spectrum Disorder Diagnosis," 2021.
[18]. N. A. Mashudi, et al., "Automated Detection of Autism Spectrum Disorder Using Machine Learning Techniques," 2021.
[19]. K. Chowdhury, et al., "Support Vector Machine-Based Early Detection of Autism Spectrum Disorder," 2021.
[20]. T. Eslami, et al., "Applications of Machine and Deep Learning for Neurodevelopmental Disorders Diagnosis," 2021.
[21]. T. Abualait, et al., "Early Detection and Intervention in Autism Spectrum Disorder: Neurobiological and Behavioral Perspectives," 2021.
[22]. J. Navinkumar, et al., "Deep Learning-Based Autism Detection from Facial Images Using Transfer Learning," 2021.
[23]. W. A. Huang, et al., "Early Prediction of Autism Spectrum Disorder Using Conditional Multi-Label Modeling on Electronic Health Records," 2021.
[24]. K. Wangmo, et al., "Integrating Machine Learning and Parental Involvement for Enhanced Autism Spectrum Disorder Detection," 2021.
[25]. I. Nafisah, A. R. Rahman, and M. H. Hasan, "Deep Learning-Based Feature Selection for Detection of Autism Spectrum Disorder Using rs-fMRI," *Frontiers in Artificial Intelligence*, vol. 8, pp. 1–12, 2025.
[26]. D. M. Setu and M. A. Rahman, "A Comprehensive Review of Autism Detection Models Using Machine Learning," *Journal of Biomedical Informatics*, vol. 150, p. 104987, 2025.
[27]. M. D. Ganggayah, S. S. Nordin, and P. K. Tan, "Accelerating Autism Spectrum Disorder Care: A Rapid Review on Machine Learning and Deep Learning Applications," *Computers in Biology and Medicine*, vol. 157, p. 106889, 2025.
[28]. K. Ehsan, M. F. Alam, and R. Khan, "Early Diagnosis of Autism Spectrum Disorder Through Automated Machine Learning," *IEEE Access*, vol. 13, pp. 15745–15758, 2025.
[29]. M. Ahmed, S. Banerjee, and R. Kumar, "Recent Developments in Autism Spectrum Disorder Detection Using Machine Learning," *Applied Sciences*, vol. 15, no. 14, p. 8056, 2025.