# Machine Learning Applications in Early Identification and Diagnosis of Autism Spectrum Disorder in Children

**Authors:** Hou Tingting

**Date:** 2022-04-07T16:44:18Z

## Abstract

Early detection, early diagnosis, and early intervention constitute a consensus in the educational rehabilitation of children with autism. However, limitations of traditional identification and diagnostic methods, together with a shortage of professionals, frequently result in children with autism missing the optimal intervention period. To ameliorate this situation, machine learning, by virtue of its advantages in objectivity, accuracy, simplicity, and flexibility, has been progressively applied to early prediction, screening, diagnosis, and assessment process management of autism in recent years, accumulating relatively substantial achievements. Nevertheless, machine learning also demonstrates limitations in research subject selection, classification data acquisition, and theoretical model application. Future research should promote the establishment of tracking databases for perinatal and neonatal pathophysiological information and standardized model classification index systems, while continuing algorithmic optimization and accelerating the translation of theoretical achievements in intelligent autism identification and diagnosis into practical applications.

## Full Text

### Application of Machine Learning in Early Identification and Diagnosis of Autism Spectrum Disorder in Children

**HOU Tingting[1], CHEN Xiao[2], KONG Depeng[2], SHAO Xiujun[3], LIN Fengxun[1], LI Kaiyun[1]**

([1] School of Education and Psychology, University of Jinan, Jinan 250022, China)
([2] School of Educational Technology and Science, Zhejiang University of Technology, Hangzhou 310023, China)
([3] Qingdao Chenxing Experimental School, Qingdao 266000, China)

**Abstract:** Early detection, diagnosis, and intervention represent the consensus approach for educational rehabilitation of children with autism spectrum disorder (ASD). However, limitations of traditional identification and diagnostic methods, coupled with a shortage of qualified professionals, often result in missed optimal intervention windows for autistic children. To address this challenge, machine learning has been increasingly applied in recent years to early prediction, screening, diagnosis, and evaluation process management of autism, leveraging its advantages of objectivity, accuracy, simplicity, and flexibility. This approach has yielded substantial research 成果. Nevertheless, machine learning also exhibits limitations in research participant selection, classification data collection, and theoretical model application. Future research should focus on establishing tracking databases for perinatal and neonatal pathophysiological information while developing standardized model classification index systems. Simultaneously, continued algorithm optimization is needed to accelerate the translation of theoretical achievements in intelligent autism identification and diagnosis into practical applications.

**Keywords:** Machine learning, Autism, Early identification and diagnosis, Systematic review

## 1. Introduction

In the era of artificial intelligence, the deep integration of AI and special education has become a significant trend in educational development. Numerous researchers have dedicated themselves to matching AI technologies with the needs of children with different types and severity levels of special educational needs, providing more personalized and precise educational rehabilitation and medical services to improve their quality of life and learning outcomes. As technology matures, autism spectrum disorder (ASD)—a condition with large affected populations and complex etiology—has garnered widespread attention in the intelligent education community. Research has focused on implementing effective interventions for autism, with scientific and accurate early identification and diagnosis serving as the prerequisite for educational intervention, inevitably becoming a critical issue that must be addressed in both research and practice.

According to the U.S. Centers for Disease Control and Prevention, the prevalence of ASD in 2021 was 1 in 44, representing a nearly 240.9% increase from the 1 in 150 rate observed between 2000-2002. Faced with rising prevalence and unclear pathogenesis, autism has evolved into a major global public health challenge. Active early prevention, identification, and intervention constitute essential responses to this challenge. However, large-scale risk screening specifically for ASD populations remains rare, and ASD diagnosis primarily relies on external behavioral observation and subjective judgment by evaluators. The conventional diagnostic process involves assessors using standardized diagnostic tools (such as DSM-IV/DSM-V and ICD) to evaluate children through psychological and educational measurements, medical examinations, parent/guardian interviews, and daily observations. This process is time-consuming and demands extensive professional theoretical knowledge and clinical experience from evaluators. The reality is that domestically, very few professionals possess diagnostic qualifications. Even in countries with relatively advanced special education development, such as the United States, only 8% of pediatricians are qualified to diagnose ASD, creating a severe shortage of professional diagnostic services. Research shows that while behavioral symptoms of autistic children can manifest as early as 12-24 months, the average age of diagnosis for typical autism is after 4 years, with severe autism diagnosed between 5.6-8.6 years and high-functioning autism often delayed until after 11 years, forcing children into a painful "wait-to-fail" situation. Moreover, this entirely external behavior-based and subjective diagnostic approach incorporates limited reference variables, lacks objective and consistent neuroscientific indicators, and frequently results in misdiagnosis and missed diagnoses, causing some children to be incorrectly labeled as autistic, which seriously violates the original purpose of assessment and diagnosis. Most concerningly, these issues severely hinder early intervention for individuals with autism, delay the golden period for educational rehabilitation, and impose tremendous psychological and economic burdens on autistic children and their families. Finally, even after diagnosis is completed, subsequent etiological and symptomatic research based on these patients remains retrospective analysis, making it difficult to comprehensively and objectively summarize ASD behavioral characteristics, extract markers that can represent underlying physiological mechanisms, and screen out features with high clinical predictive relevance. This places autism prevention and treatment efforts in a passive position.

Machine learning (ML), as the core of artificial intelligence, is a technology that automatically analyzes data to obtain patterns and uses these patterns to predict unknown data. The rules automatically generated by this technology are not influenced by individual subjective experience, which is significant for improving classification specificity, sensitivity, accuracy, and efficiency. Furthermore, ML can extract subtle and potentially useful information from large datasets to create predictive or classification models for current problems, making it particularly suitable for gaining insights into the internal patterns of complex problems and helping to identify possible causes early on. Scholars both domestically and internationally have applied ML to the classification, diagnosis, treatment, and prognosis management of various mental and neurological disorders, achieving numerous valuable results in empirical research on prediction, identification, and auxiliary diagnosis for autism populations, though domestic reports in this area remain relatively scarce. This article reviews the application of machine learning in early identification and diagnosis of autistic children, aiming to provide references for domestic research and practice.

## 2. Research Methods

We employed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), an internationally recognized systematic literature review and meta-analysis methodology, to determine the composition of included studies. This method comprises four literature screening stages with 27 review items total. The four stages are identification, screening, eligibility, and inclusion, with review items covering literature titles, abstracts, research objectives, methods, and results. Considering the rapid development and 更新换代 of computer science technology in recent years, this review selected literature from the past five years on machine learning applications in early identification and diagnosis of autistic children, with a specific search date range from January 1, 2016, to April 29, 2021. English search terms included: "Autis," "Prediction," "Early Screening/ Early Detection/ Early Identification/ Early Recognition," "Early Diagnosis," and "Machine Learning/ Artificial Intelligence." We searched five comprehensive foreign databases: EBSCO, Elsevier Science Direct, SAGE, Web of Science, and Springer, limiting the search scope to peer-reviewed journals, yielding 520 articles. An additional 50 articles were obtained through reference citation searching, totaling 570 articles. The authors jointly reviewed titles, abstracts, and full texts, removing duplicates and papers that did not meet established criteria, ultimately obtaining 60 articles for inclusion in the review (see Figure 1).

**Figure 1. PRISMA Literature Screening Flowchart**

[Note: The figure shows the screening process from 570 total records through various exclusion stages to 60 final included studies]

**Inclusion criteria:** (1) English-language empirical journal articles with full text available and 不少于 3 pages; (2) Study participants included at least one child with ASD (in this paper, "children" are defined according to the UN Convention on the Rights of the Child as individuals under 18 years old) with clear statistical variables (age) or "child" designation; (3) Studies contained clear research questions, methods, and conclusions with detailed data support; (4) Research focused on early identification or diagnosis of ASD children, with explicit indication of machine learning technology use for classification in abstracts or keywords.

**Exclusion criteria:** (1) Non-English literature (Korean, German, Spanish, etc.), reviews, book chapters, conference proceedings, patents, and articles 少于 3 pages; (2) Study populations consisting entirely of adults over 18, elderly individuals, or other special groups; (3) Research questions, designs, or conclusions that were ambiguous without data support; (4) Studies on mental disorders unrelated to ASD or those not applying machine learning methods. These criteria were established before literature searching to reduce potential bias during screening.

Finally, based on the "categorization consistency formula" proposed by Xu Jianping and Zhang Houcan (2005), we assessed coding reliability. The first two authors jointly calculated that inter-rater reliability for included literature ranged from 0.50 to 1.00, with overall categorization consistency at 0.812; coding reliability coefficients ranged from 0.75 to 0.80, with an overall coding reliability coefficient of 0.759, indicating fair to good consistency levels.

## 3. Data Types and Collection for Machine Learning-Based Autism Identification and Diagnosis

Machine learning can discover hidden patterns in data through abstraction, making data the foundational factor in model construction. The first step in modeling is data collection.

### 3.1 Data Composition

Children with ASD exhibit high heterogeneity and weak data structure, so classification model raw data generally consists of binary classification data. In terms of participant composition, researchers typically establish both experimental and control groups, selecting children with ASD and matched children with other developmental disorders or typically developing (TD) children matched for chronological age and mental age. Common types of children with disorders include Attention Deficit Hyperactivity Disorder (ADHD), Developmental Delay (DD), Childhood Apraxia of Speech (CAS), and Other Neurodevelopmental Disorders (OND). Some studies also categorize children into High Risk (HR) and Low Risk (LR) groups based on future autism risk levels. Regarding age range, studies cover various developmental stages, with the earliest screening and diagnosis timing basically within 3 years of age, during infancy and early childhood. Autism risk prediction research occurs even earlier, generally before 1 year of age, with some studies extending to the perinatal and neonatal periods, though such research remains rare. In terms of participant sources, studies involve dozens of countries and regions including the United States, United Kingdom, France, China, Iran, New Zealand, India, Italy, South Korea, Spain, and Australia. Common data collection channels include public autism research databases (see Table 1), hospitals, special education schools or rehabilitation centers, and a small number of online search engines (such as Yahoo Answers, Google Search).

Sample sizes vary dramatically across studies, ranging from dozens to thousands, with studies exceeding 300 participants typically drawing data from public databases. Public autism research databases are open resources established by researchers to facilitate collaborative research across different disciplines and professions. Their large-scale data storage provides tremendous convenience for autism identification and diagnosis using machine learning. Taking the most frequently used ABIDE database as an example, it integrates brain structure and functional imaging research data on ASD and TD children collected from multiple laboratories worldwide, freely available to researchers. Two phases of data have been collected, with over 2,200 participants, and registered researchers from North America, Europe, Africa, Asia, and other regions.

**Table 1. Public Autism Research Databases**

| No. | Chinese Name | English Name (Abbreviation) |
|---|---|---|
| 1 | Autism Brain Imaging Data Exchange Database | Autism Brain Imaging Data Exchange (ABIDE) |
| 2 | Autism Genetic Resource Exchange Database | Autism Genetic Resource Exchange (AGRE) |
| 3 | Boston Autism Consortium | Boston autism consortium (AC) |
| 4 | Autism Treatment Network | Autism Treatment Network (ATN) |
| 5 | Simons Simplex Collection Database | Simons Simplex Collection (SSC) |
| 6 | Simons Variation in Individuals Project Database | Simons Variation in Individuals Project (Simons VIP) |
| 7 | University of California Irvine Repository | University of California Irvine Repository (UCI) |

The application of machine learning technology has overcome the limitations of traditional retrospective research, enabling autism identification and diagnosis at younger ages. Hidden patterns extracted through complex computations can also be applied to perinatal and neonatal examinations to help identify high-risk factors early, buying time for early intervention. Unfortunately, such research remains rare. Additionally, machine learning technology enables large-scale, multi-indicator, cross-cultural data collection, moving from independent small-sample data to cross-regional public shared databases, effectively improving research external validity and providing evidence for deeper exploration of autism causes and successful prevention. However, it should be noted that current public databases mostly originate from individual or institutional research with inconsistent data collection purposes and standards, lacking homogeneity in sample information integration. This may create bias between training and test sets, and model generalization issues require further verification.

### 3.2 Data Collection

After sample selection, data collection can commence. In recent years, data collection methods and types have become more diverse due to technological updates, such as physiological signal data collection based on EEG signals and imaging technology, and behavioral data collection through scales, observations, and interviews. Machine learning identification and diagnosis are based on features reflecting autistic tendencies and various neurological markers. Currently, these mainly include demographic data such as age, gender, and ethnicity; typical symptom manifestations like eye contact, social smiling, and imitation; clinical data including personal and family medical history; physiological signal data such as functional magnetic resonance imaging, structural magnetic resonance imaging, electroencephalography, and ultrasound; and multi-modal data combining different data types (see Table 2).

**Table 2. Data Types for Machine Learning-Based Autism Identification and Diagnosis**

| Data Type | Specific Indicators |
|---|---|
| Imaging | Functional magnetic resonance imaging (fMRI), structural magnetic resonance imaging (sMRI), electroencephalogram (EEG), ultrasound (UT) |
| Scales/Questionnaires | Autism Diagnostic Observation Schedule (ADOS), Autism Diagnostic Interview-Revised (ADI-R), Diagnostic and Statistical Manual of Mental Disorders (DSM-IV/DSM-V), International Classification of Diseases (ICD-10), Autism Observational Scale for Infants (AOSI), Modified Checklist for Autism in Toddlers (M-CHAT), Mullen Scales of Early Learning (MSEL), Vineland Adaptive Behavior Scales (VABS), Social Responsiveness Scale (SRS), Autism-spectrum Quotient-10 (AQ-10), Child Behavior Checklist (CBCL), Social Communication Questionnaire (SCQ), Parent Questionnaire |
| Behavior | Eye movement behavior, social behavior, stereotyped behavior, postural control, upper limb movement, etc. |
| Clinical Data | Parental medical history, child personal information (age, gender, handedness, IQ, ethnicity, medical history, diet, sleep) |
| Genetics | DNA, RNA, metabolites, speech |

**3.2.1 Image-Based Classification** Using biological markers such as brain imaging, ultrasound imaging, and retinal imaging as input features for machine learning technology can provide objective evidence for autism classification, making it the most widely applied approach in research. fMRI' s high temporal resolution facilitates precise localization of cortical activity in specific brain regions. sMRI offers high spatial resolution, objectively recording differences in brain tissue composition among individuals. Some scholars propose that combining both technologies yields the best results in autism classification studies. For example, Rakić et al. (2020) collected sMRI and resting-state fMRI (rs-fMRI) images from 817 participants aged 7-64 years in the ABIDE-I database. They used the Fisher algorithm to reduce dimensionality of feature vectors, which then served as input for classifiers, finally employing stacked autoencoders and multilayer perceptron algorithms for identification, achieving accuracy up to 85.06%. Sen et al. (2018) also found that using both technologies as feature inputs produced optimal ASD vs. TD identification results. However, other studies have shown that using fMRI or sMRI images alone can also achieve accurate identification of individuals with autism (Eugenia et al., 2020; Katuwal et al., 2016). Guo et al. (2017) collected rs-fMRI data from individuals with ASD and TD, finding that a deep neural network with feature selection achieved classification accuracy of 86.36%. EEG, being less expensive and more convenient, can provide evidence of neurophysiological activity for autism identification in children. Studies have used EEG technology to identify autistic children as young as 3 months (Dickinson et al., 2020) and aged 3-14 years (Grossi et al., 2017; Abdolzadegan et al., 2020) with good performance.

**3.2.2 Scale/Questionnaire-Based Classification** Traditional standardized measurement tools, which incorporate rich theoretical knowledge and solid practical experience from clinical experts, serve as the "gold standard" for autism screening and diagnosis, providing strong evidence for machine learning model construction. There are two forms of identification and diagnosis based on traditional assessment tools. The first involves using machine learning technology to analyze data from traditional standardized tests to explore classification accuracy. In studies from 2018 and 2020, Abbas et al. used behavioral modules commonly employed in ADOS and ADI-R as classifier input information, proposing a machine learning-based multi-module autism assessment method. This method includes three modules: parent questionnaire, key behaviors marked in home videos, and clinical questionnaires. Using random forest, L2-regularized logistic regression, and gradient boosted decision trees to classify children aged 18-72 months, they found that when using modules one and two simultaneously with an inconclusive results variant ( 25%), classification accuracy exceeded 98%, reaching 92% when all three modules were combined. Even with three modules, total testing time was only about 15 minutes. Other researchers have designed mobile screening applications, ASD Tests (Thabtah, 2019) and Autism AI (Shahamiri & Thabtah, 2020), based on AQ-10 (child/adolescent/adult versions) and the Quantitative Checklist for Autism in Toddlers (Q-CHAT) for autism prediction in children, with ASD Tests achieving identification accuracy up to 99.85% in just 3-5 minutes.

The second approach involves machine learning algorithms screening out the most representative questions from traditional standardized scales or questionnaires to build classifiers (Levy et al., 2017). For example, Duda et al. (2016b) found through logistic regression and other algorithms that when 7 key items were selected from ADI-R' s 93 items or when ADOS dimensions 2 and 3 were used alone as classification criteria, classification accuracy for autistic children reached 99%, reducing testing time to just over ten minutes. Duda and Ma et al. (2016a) and Duda and Haber et al. (2017) also built classifiers based on 5 typical behaviors and 15 derived questions from the SRS scale, with classification performance meeting expectations (AUC > 93%, AUC = 89%). Notably, nearly half of the literature on scale-based classification simultaneously employed two or more tools, hoping to achieve better classification results through integration of different assessments.

**3.2.3 Behavior-Based Classification** Typical behavioral manifestations of autistic children can be directly observed in early development, providing explicit indicators for classification model construction. Social impairment and repetitive stereotyped behaviors are core diagnostic criteria for ASD and serve as important bases for behavioral indicator classification. Qiu et al. (2020) used the Still-Face Paradigm (SFP) to model classification indicators including frequencies and durations of non-social/social smiling, resistant behavior, eye contact, and positive social engagement in HR-ASD and TD children. Negin et al. (2021) constructed an Expanded Stereotype Behavior Dataset (ESBD) for HR-ASD children, achieving differentiation from TD through four actions: spinning, arm flapping, hand movements, and head banging. Eye trackers can record eye movement trajectory features during visual information processing, with eye movement behavior being the most commonly used behavioral indicator for autism classification based on machine learning. Specific indicators include fixation duration and frequency. Wan et al. (2019) used an SMI RED250 portable eye-tracking system to record fixation durations of 4-6-year-old ASD and TD children on different body parts while watching a 10-second video of a female speaker, building a classifier using support vector machine algorithms. They found that differences in children' s fixation durations on the speaker' s mouth and body could effectively distinguish ASD from TD (Acc = 85.1%, Sen = 86.5%, Spe = 83.8%). Additionally, since some autism populations exhibit motor impairments, studies have shown that indicators such as postural control and upper limb movement can also serve as early classification bases, with application accuracy reaching over 90% (Li et al., 2020; Wedyan et al., 2019).

**3.2.4 Multi-Modal Data Classification** "Multi-modal" refers to the diversity of symbolic systems or effective means of representation, while multi-modal data refers to the integration of multiple classification methods and indicators to achieve research objectives. In this study, it specifically refers to combining single-modal data collection methods such as scales, imaging, and behavior, using two or more methods simultaneously for autism classification model construction. Kang et al. (2020) collected rs-fMRI data and eye movement indicators in response to own-race and other-race human faces from 3-6-year-old ASD and TD children, combining support vector machine with minimum redundancy-maximum relevance (MRMR) feature selection methods for participant classification. They found that when EEG and eye movement data served as simultaneous classifier inputs, classification accuracy reached 85.44% with AUC exceeding 93%, significantly higher than the 66.9% accuracy achieved using single behavioral indicators alone. Bosl et al. (2018) longitudinally tracked high- and low-risk ASD infants' EEG data between 3-36 months, measuring participants' autism symptom severity using ADOS at 36 months, and found that EEG data from 3-month-old infants could predict their ADOS scores at age 3. Furthermore, Han Baozhen (2018) found that brain network data fused from multi-modal information achieved the best classification results for ASD and TD.

The application of machine learning in autism early classification has expanded from single-modal to multi-modal data, with increasingly diverse classification indicator types. Collection methods have evolved from explicit behavioral indicators to internal neurological markers, with technology becoming more objective and comprehensive. Rich data collection methods and types can assist model construction at the source, providing a foundation for achieving optimal classification results. However, several issues in current data collection require improvement. First, imaging techniques are relatively monolithic, mainly fMRI and sMRI, with limited application of ERP, fNIRS, and other technologies in this field. Second, traditional diagnostic tools are systematically designed with fixed purposes for each module, yet different machine learning methods apply inconsistent item or module reduction criteria, yielding non-unified results. The basis and standards for tool reduction lack normative discussion. Traditional tool measurement duration ensures more detailed and accurate testing; whether shortening time to over ten minutes with machine learning may cause information omission and result bias remains a significant concern. Finally, classification indicators based on typical behaviors are overly broad, while research based on genetic materials such as DNA and RNA is rare. Different researchers focus on vastly different areas, making it debatable how to select the most representative one or several indicators from numerous options. These issues require extensive discussion and demonstration before applying machine learning to autism classification. Only by truly solving these source problems can technological advantages be genuinely realized for the autism population.

## 4.1 Machine Learning Algorithms

Algorithms are the technical elements of model construction. Machine learning algorithms are broadly classified into supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning based on learning methods, and into traditional machine learning and deep learning based on model structure depth. Current applications in early autism identification primarily use traditional algorithms, including Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbors (KNN), Naive Bayes (NB), Regression, Random Forest (RF), and Gradient Boosted Decision Trees (GBDT). However, with continuous technological development, deep learning methods such as Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Deep Neural Networks (DNN), and Multilayer Perceptrons (MLP) have gradually attracted researchers' attention (Eni et al., 2020; Raj & Masood, 2020).

SVM is one of the representative algorithms of traditional machine learning, particularly suitable for binary classification problems. Liu et al. (2016) used a Radial Basis Function (RBF) kernel to project non-linear data into high-dimensional space for separation, effectively distinguishing ASD from TD with 88.51% accuracy. However, research shows that traditional algorithms like SVM often classify samples directly based on features without performing or only performing one feature transformation and selection, overly relying on upstream-provided features, exhibiting sensitivity to missing data and overfitting problems in multi-classification scenarios, with limited effectiveness when facing the complex etiology and manifestations of autism (Namdeo & Singh, 2021; Yang Jianfeng et al., 2019). Deep learning, without relying on upstream features, can directly input large amounts of raw data into models, detecting and calculating abstract features through multiple data transformations to aid classification, with improved classification performance and generalization ability compared to traditional methods (Hu Yue et al., 2019; Li et al., 2019; Dubreuil et al., 2020). Salari et al. (2019) used rs-fMRI data as classification indicators, employing SVM, KNN, RF, and CNN to build classifiers separately, finding that CNN (Acc = 70.22%) achieved higher classification accuracy than SVM (Acc = 69.35%), KNN (Acc = 62.11%), and RF (Acc = 59.94%). Dong et al. (2021) built a CNN model using resting-state EEG as classification indicators, achieving classification accuracy of 92.63%. Artificial neural networks, as basic components of deep learning, have also demonstrated high classification accuracy. Grossi et al. (2017) built a novel classification system based on ANN, combining Multi-Scale Ranked Organizing Map (MS-ROM) with Implicit Function as Squashing Time (I-FAST) algorithms, using EEG images as discrimination indicators, and found the algorithm achieved 100% accuracy in ASD prediction. However, research also shows that traditional algorithms can self-optimize by increasing the number of classifiers. Random Forest combines bootstrap resampling with decision tree algorithms to build a collection of tree-based classifiers (Cao Zhengfeng, 2014), with significantly higher classification speed, accuracy, and noise resistance than single classifier algorithms, and can reduce overfitting problems to some extent. Wingfield et al. (2020) developed an ASD screening program and found that RF significantly outperformed single NB and DT classifiers, with AUC reaching 98%.

Machine learning algorithms can automatically detect patterns in data without explicit coding, handling complex non-linear data with powerful functional advantages over traditional identification and diagnosis methods. Recent applications in ASD classification research have also shown a trend transitioning from traditional algorithms to deep learning, substantially improving diagnostic efficiency. However, over 75% of current studies still use traditional machine learning algorithms, indicating a need for increased attention to deep learning. Additionally, numerous algorithms are currently used in research, with significant differences among researchers and varying results from the same algorithm applied to different populations, leading to ongoing controversy over optimal algorithms. Future research should continue exploring how to identify the best algorithms to maximize technological value.

## 4.2 Machine Learning Effectiveness Evaluation

Effectiveness evaluation serves as the guiding principle and navigation tower for model construction. After data collection and algorithm selection, model validity must be verified. Machine learning effectiveness evaluation is generally calculated by using professional physician diagnosis as the criterion and computing correlation coefficients or successful prediction proportions. Based on original binary classification data characteristics, classification effects are generally divided into positive and negative classes, where true positives and true negatives represent good predictions in positive and negative classes, respectively, while false negatives and false positives represent poor predictions (see Table 3). Specific model performance evaluation indicators include: (1) **Accuracy (Acc)**, the proportion of correctly classified samples to total samples—generally, higher accuracy indicates better classification effects, calculated as Acc = (TP+TN)/(TP+FN+FP+TN); (2) **Sensitivity (Sen)**, the proportion of correctly classified individuals with autism in the positive class, also called recall or True Positive Rate (TPR), calculated as Sen = TP/(TP+FN); (3) **Specificity (Spe)**, the proportion of correctly classified individuals without autism in the negative class, also called True Negative Rate (TNR), calculated as Spe = TN/(FP+TN); (4) **Receiver Operating Characteristic (ROC) curve**, a graph with false positive rate on the X-axis and true positive rate on the Y-axis, primarily comparing classification performance by calculating the Area Under the ROC Curve (AUC/AUROC)—generally, larger AUC/AUROC indicates better classification effects. In addition to these four commonly used indicators, some studies also employ Positive Predictive Value (PPV), Negative Predictive Value (NPV), and False Discovery Rate (FDR) to examine classifier effectiveness.

**Table 3. Confusion Matrix for ASD Binary Classification Data**

| | ASD | Non-ASD |
|---|---|---|
| **ASD** | True Positive (TP) | False Negative (FN) |
| **Non-ASD** | False Positive (FP) | True Negative (TN) |

## 4.3 Machine Learning Process

A logically clear, complete, and orderly data processing process is the explicit manifestation of effective model implementation. Machine learning application in early ASD identification and diagnosis generally follows four steps: First is the data collection stage, selecting imaging, scale, behavior, and other data indicators required for ASD classification. Second is the raw data processing stage, primarily aimed at feature selection through oversampling, undersampling, and other methods to reduce noise from duplicate records, missing values, and unbalanced data, decreasing computational complexity and improving classification precision to form datasets required for classification. Third is the ASD vs. control group classification stage, typically employing k-fold cross-validation to randomly divide the entire dataset into k portions, with one portion as the test set and k-1 portions for training. The training data can be further randomly divided into training and validation sets, conducting cross-validation k times so each portion serves as the test set once, averaging k results to obtain final outcomes. The specific process can be summarized as: (1) Input the training set, use machine learning algorithms for training to build an ASD classification model; (2) Input the validation set into the model, compare differences between model predictions and "autism" labeled data to verify model classification accuracy; (3) Use the test set to evaluate model classification performance. Finally, obtain the optimal classification model. Epalle et al. (2021) built a DNN model following this procedure, using brain imaging data from ABIDE-I to distinguish ASD from TD. In the data processing stage, they used oversampling for feature selection, increasing features by 19,900 in the CC200 set, 4,005 in the AAL90 set, and 12,720 in the DOS160 set. In the classification stage, they used 10-fold cross-validation, 5-fold stratified cross-validation, and leave-one-site-out methods to build classifiers and evaluate results, setting 70% for training, 20% for validation, and 10% for testing in each data portion. Final model validation showed the classifier could achieve 78.07% accuracy in distinguishing the two groups.

Machine learning application in ASD child classification automatically allocates training and test sets, selects and extracts data features through filtering and repeated testing, discovers hidden patterns in data, reduces data bias at the source, and increases research internal validity. Model validity verification covers multiple dimensions including Acc, Sen, Spe, AUC, PPV, and NPV, with increasing emphasis on objective model measurement, which also enhances external validation to some extent. However, literature review also reveals that classification indicator combinations vary dramatically across studies, with inconsistent numbers of reported indicators and a lack of consistent, widely recognized classification effectiveness indicator systems. Additionally, Falkmer et al. (2013) suggest that autism classification model accuracy should reach at least 80% to be considered acceptable, yet some current studies report accuracy below 60%, far from meeting requirements (Bussu et al., 2018). How to continue optimizing and standardizing these issues at the technical level requires deeper discussion in future research.

## 5.1 Predicting Autism Risk Factors

As a neurodevelopmental disorder with unclear etiology, autism cannot be prevented through routine prenatal examinations alone. Machine learning technology can effectively predict autism risk by extensively collecting imaging and biological indicator information, fully mining and summarizing information, and calculating predictive data models, providing new directions for autism etiology exploration and comprehensive diagnosis and treatment.

Both perinatal environment and maternal physical and mental health factors can affect embryonic development quality. Prenatal medical examination and monitoring of pregnant women (especially high-risk pregnant women) and fetuses may help identify autism risk factors. Machine learning can infer ASD risk factors from imaging and biomarker information collected through ultrasound examinations, amniocentesis, X-ray examinations, villous cell examinations, and fetoscopy. For example, Caly et al. (2021) predicted newborn autism risk by collecting fetal ultrasound measurements (including femur length, head circumference, abdominal circumference) and biological characteristics (such as IgG cytomegalovirus levels). Using Least Absolute Shrinkage and Selection Operator (LASSO), DT, and eXtreme Gradient Boosting (XGBoost) algorithms to build classification models, they found the model' s positive predictive value could reach 77%. Ten indicators including maternal family history of immune disease, maternal immunity to cytomegalovirus, late-pregnancy fetal femur length, late-pregnancy fetal white blood cell count, newborn heart rate, and gender could serve as ASD predictive markers. Genes are internal factors determining life and health. Genetic testing can extract DNA molecular information from cells to determine newborn gene types and defects. Machine learning technology can help discover genetic factors related to autism by sorting out complex genetic information. Bahado-Singh et al. (2019) used the Infinium HumanMethylation450 BeadChip to detect white blood cell DNA from ASD and TD newborn blood spots, finding significant abnormal methylation at CpG sites in cytosine genes, with prediction results achieving AUC of 100%. Researchers also found epigenetic regulation disorders in autism-related genes including EIF4E, FYN, SHANK1, VIM, LMX1B, GABRB1, SDHAP3, and PACS2. Other scholars have reported that long non-coding RNA (lncRNA) can also predict newborn ASD risk (AUC = 83.9%; Gök, 2018). Qiu et al. (2020) built a prediction model based on SVM showing that multiple metabolites in folate-dependent one carbon metabolism (FOCM) and transsulfuration (TS), as well as maternal pre-pregnancy and pregnancy folic acid intake, could serve as indicators for newborn autism risk prediction, with model prediction accuracy exceeding 96%. Using machine learning to track functional connectivity Magnetic Resonance Imaging (fcMRI) in infants from HR-ASD families advanced ASD risk prediction age to 6 months post-birth (Emerson et al., 2017).

## 5.2 Screening for High-Risk Autism

Early autism screening can identify high-risk children in large populations, which is significant for improving social adaptation in autistic children, preventing secondary developmental disorders, and reducing family and social burdens (Chen Guanghua et al., 2022). If screening can be directly operated by parents or others familiar with the child, potential ASD patients can be identified earlier, avoiding missed intervention opportunities. Thabtah et al. (2018) used machine learning models to simplify traditional scales (CHAT, AQ), developing a mobile application suitable for ASD screening—ASD Tests. This tool can be freely downloaded and used on mobile phones, covering four stages: infancy (1-3 years), childhood (4-11 years), adolescence (12-16 years), and adulthood (17+ years), with only 10 test questions per stage and testing time under 3-5 minutes, easily operable by non-professionals. Research found the program' s screening sensitivity, specificity, and accuracy for HR-ASD across child, adolescent, and adult stages all exceeded 95%, demonstrating good classification performance (Thabtah et al., 2019). Raj and Masood (2020) built an autism screening model based on 21 features including individual gender, ethnicity, and presence of congenital jaundice, finding that CNN could achieve screening accuracies of 98.30%, 96.88%, and 99.53% in child, adolescent, and adult populations, respectively, enabling rapid identification of autistic patients in large populations. Additionally, neuroimaging data and behavioral performance, as important internal and external references for risk identification, have been widely applied in screening or diagnostic model construction, achieving satisfactory results (Rakić et al., 2020; Dickinson et al., 2020; Lai et al., 2020; Qiu et al., 2020).

## 5.3 Assisting Autism Diagnosis

Diagnosis involves more comprehensive and detailed professional testing of children identified as high-risk through screening. Traditional diagnostic methods obtain data through observing children' s external behaviors and parent reports, which are susceptible to evaluator subjectivity and reporter recall bias. Problems such as long waiting periods and misdiagnosis in the diagnostic process also hinder subsequent clinical intervention implementation. Machine learning-assisted professional diagnosis can incorporate large numbers of objective classification indicators, reducing subjective testing drawbacks to some extent, while also reducing time consumption and waiting costs. Social robots are commonly used tools in ASD intelligent diagnosis, designed to be friendly and communicate with test-takers through rich facial and body language, reducing social anxiety in autism. Ramírez-Duque et al. (2019) used RGBD sensors to collect behavioral features of HR-ASD and TD toddlers interacting with the ONO social robot, and used CNN to build a facial recognition model for the two groups of children to assist professionals in diagnosing ASD children. They found the model could achieve automated ASD identification within 40 minutes, with accuracy basically consistent with clinical expert diagnosis. Imaging and traditional scales also provide effective evidence for machine learning-assisted diagnosis (Bone et al., 2016). Tariq et al. (2018) built a model using ADOS and ADI-R as indicators, achieving diagnostic accuracy over 94% and sensitivity exceeding 90%. Abdolzadegan et al. (2020) analyzed EEG data from 3-12-year-old ASD and TD children, establishing diagnostic models based on SVM and KNN, with accuracy reaching 94.68%.

## 5.4 Managing Assessment and Intervention Processes

The ultimate goal of early autism prediction, identification, and diagnosis is to provide effective early intervention for children. Collecting information on assessment and intervention processes for autism populations and tracking intervention implementation effects can provide insights for determining children' s physiological and pathological mechanisms and treatment targets. However, traditional screening, diagnosis, intervention, and re-evaluation are relatively independent and scattered stages, making systematic collection and integration of information difficult, which in turn makes it harder to discover and grasp internal trends and hidden features in data, resulting in data waste. A major advantage of machine learning-based classification is effective data utilization, constructing integrated intelligent assessment and intervention management systems based on existing assessment and intervention data to enhance longitudinal monitoring of diagnosis and intervention for autistic children. Cognoa (Cognoa ASD Screener/Cognoa ASD Diagnosis Aid) is a new ASD assessment program supported by machine learning technology, combining screening, diagnosis, and risk prediction functions. The program primarily serves infants and preschool autistic children (18 months to 5 years), requiring parents to complete a 15-item questionnaire and upload two or more 1-2 minute videos of children' s daily life (such as mealtime, playtime). The system automatically scores and outputs risk values after data upload, with the entire process taking about 15 minutes (Kanne et al., 2018). The program automatically saves results after each assessment. In June 2021, the U.S. Food and Drug Administration (FDA) evaluated the feasibility of applying this program to autism diagnosis, finding that program diagnosis results for 425 children from 14 medical centers were basically completely consistent with expert conclusions, leading to approval for market use. Machine learning technology-based intelligent management systems for autism pictorial assessment can also record data on children' s ability changes from early screening to intervention treatment, obtaining children' s re-evaluation information (Perera et al., 2017). Expert systems supported by artificial neural networks can even intelligently respond to and solve uncertainty problems without pre-programmed settings, providing real-time suggestions for autism diagnosis and intervention, serving as "capable assistants" in clinical assessment (Jin Yuchang et al., 2022; Rahman et al., 2020; Negin et al., 2021).

In summary, machine learning applications in autism classification cover etiological analysis, risk prediction, screening of at-risk populations, assisted diagnosis, and intelligent disease course management. These efforts help increase practitioners' understanding of autism etiology and disease course information, achieve dynamic management of screening, diagnosis, intervention, and re-evaluation, save human and material resources required for long-term follow-up, more quickly identify "commonalities" among numerous individual factors, and improve treatment effectiveness—providing a reference for future clinical diagnosis and treatment.

## 6.1 Advantages

Currently, machine learning technology has been applied to early prediction, screening, and diagnosis of autistic children, with its objectivity, convenience, and effectiveness undergoing extensive validation. Classification performance continues to optimize and improve, with some intelligent tools beginning to be promoted and applied. Compared to empirical diagnosis and traditional measurement, the integration of machine learning demonstrates several advantages of artificial intelligence technology, concentrated in four aspects. First, machine learning has stronger capabilities for acquiring and organizing data indicators, helping to build multi-modal classification models. Traditional classification methods primarily draw conclusions through observation of explicit behavioral indicators with limited reference variables, while machine learning can combine various types of data indicators obtained through different observation methods, considering both internal and external influencing factors, reducing subjective errors, and improving comprehensiveness and objectivity of prediction and identification results. Additionally, machine learning models can longitudinally manage assessment and diagnosis processes, increasing data utilization rates. Second, machine learning reduces time costs for prediction and identification, improves evaluation efficiency, and reduces pressure on diagnostic personnel. Classification model establishment can quickly identify high-risk autistic patients through large-scale screening, alerting parents and relevant personnel earlier, while also compressing diagnosis time by reducing test items and shortening evaluation processes, decreasing waiting time and professional physician pressure. Third, machine learning increases research internal and external validity. The modeling process and computational principles of machine learning differ from traditional methods, automatically dividing raw data into training and validation sets without requiring data generation models and parameter estimation, directly inputting performance indicators. Models can also continuously optimize with sample size accumulation, enhancing research internal validity and external generalizability, and accelerating the speed of model application to practice. Fourth, machine learning can automatically extract key features to help establish risk early warning mechanisms. Machine learning can capture subtle expression or gesture changes unobservable to the naked eye, reducing omission of key indicators, and is suitable for non-linear data, capable of extracting crucial information from complex data, making it more appropriate for analyzing diseases like ASD with unclear etiology and complex structure, reducing the possibility of symptom deterioration. Especially for infants and preschool children, early detection and intervention can maximize benefits, helping these children and their families achieve higher quality of life and reducing various social adaptation problems caused by disabilities.

## 6.2 Limitations and Future Directions

Although machine learning has certain methodological advantages in ASD child identification and diagnosis, and related research achievements are gradually increasing, there remains some distance from large-scale application. Specific problems faced and possible solutions can be summarized in three aspects.

First, there is a lack of research tracking autism pathogenic sources in pregnant women and newborns. Genetic research results indicate that common genetic variations such as gene mutations, translocations, inversions, and copy number variations may be correlated with ASD to some degree, with copy number variation frequency significantly higher in newborns with ASD than in TD children (Sanders et al., 2011). Other research reports suggest ASD may result from a sequence of pathogenic events affecting cell proliferation, migration, and many other fundamental processes (Caly et al., 2021). Autopsy reports of autistic children' s brains also show abnormal excess neurons in the prefrontal cortex, along with brain overgrowth and macrocephaly issues (Tyzio et al., 2014). Rahman et al. (2020) also found that parental age and family medical history increase autism risk in newborns. This evidence indicates autism is influenced by genetic factors, and these conditions may already appear during embryonic formation and development (Caly et al., 2021). In addition to genetic factors, maternal exposure to hazardous environments is also a susceptibility factor for autism. Choi et al. (2016) found that increased ASD incidence is associated with maternal viral or microbial infection, fever, and immune system diseases. Medication use during pregnancy, particularly sodium valproate, and vitamin deficiency may cause autism in fetuses. Cesarean section, premature birth, and neonatal complications also increase ASD incidence (Cloarec et al., 2019). This suggests that internal and external environments in early life are causes of autism. Clarifying pathological, etiological, and disease course factors in early individual development may provide new opportunities for risk prediction and prevention, but unfortunately, current literature rarely includes autism identification research using perinatal data. Among the only two studies in current literature, Caly et al. (2021) collected routine biomarkers and ultrasound measurements from pregnant women from early pregnancy to one day after birth to predict autism, comparing predictions with diagnosis results when children were 4-5 years old, achieving 77% prediction accuracy. Bahado-Singh et al. (2019) collected 249 types of neonatal (24-79 hours after birth) white blood cell epigenome markers, successfully predicting children later diagnosed with autism (Sen = 97.5%, Spe = 100%). This suggests that if we can extensively collect health information on pregnant women and their family members during embryonic formation and early neonatal development, focusing on possible teratogenic factors in genetics and environment, especially relatively stable biological markers, and use the massive data sorting function of machine learning models to extract common factors, we may gradually move from a "passive waiting" to an "active defense" position. Therefore, future research could, following ethical principles, attempt to establish children' s physiological and pathological information databases through pregnancy screening and neonatal physical examination links, collecting process information on children (especially those from high-risk families) from embryonic period to birth, and could also 借鉴 the American Academy of Pediatrics' approach to conduct long-term follow-up studies at fixed time points (such as at birth, 8 months, 18 months, etc.) with participant informed consent (Wan et al., 2019), obtaining original data on autism pathogenic sources in early individual development to summarize possible risk factors and prevent this disorder early.

Second, there is a lack of standardized model classification data systems. While machine learning models provide possibilities for processing massive information, technology' s inclusiveness of data may become a double-edged sword. Although current machine learning classification indicators are diverse in form, their content is complex and lacks clear classification basis and systematic, rule-based application effect analysis, with low verifiability of evaluation results. Taking single-modal classification indicators based on behavior as an example, researchers have used specific indicators including dozens of behavioral manifestations such as repetitive stereotyped behaviors, resistant behavior, self-injurious behavior, fine/gross motor skills, social interaction behavior, social smiling, joint attention, imitation, eye movement changes, expressive and receptive language. Different behavioral manifestations can be further subdivided into more micro-level classification indicators—eye movement changes alone can differentiate into eye scanning, eye tracking, and eye touching/contacting paradigms. However, even with the same behavioral manifestation, conclusions are not completely consistent (Wan et al., 2019; Shic et al., 2014). Scale-based classification also faces this issue, with research literature containing multiple classification tools with inconsistent information and test objects. Even with the same assessment tool, item screening standards and conclusions are not completely unified (Emerson et al., 2017). Multi-modal classification models mix more information with low relevance to evaluation during multi-source information fusion, causing some interference with testing accuracy. The combination of multiple modalities and different classification indicators within the same modality provides researchers with more classification standard choices while also introducing more measurement errors, 不利于 research result promotion and generalization.

Therefore, model construction should emphasize internal orderliness of classification indicators, sorting out standards and basis for indicator selection, gradually identifying a subset of indicators with strong stability, wide applicability, and high accuracy to construct more scientific and stable classification indicator systems. Some researchers propose that age could serve as a basis for classification indicator division. Liu et al. (2016) showed that ASD identification in infants and children is more suitable for non-verbal behavioral features such as eye contact, gestures, and facial expressions, while adolescent ASD should focus more on language communication and emotional interaction with peers. With age increase, social communication and emotional connection become more important classification indicators. Additionally, brain changes often appear before clinical manifestations and have high stability (Wolff et al., 2012; Wolff et al., 2015; Elsabbagh et al., 2012), so objective imaging indicators should be particularly emphasized in classification processes (Hazlett et al., 2017; Emerson et al., 2016). Verbal IQ, autism severity, and other verified and easily collected classification standards can also be directly incorporated into classification systems (Katuwal et al., 2016). However, these criteria remain relatively broad and cannot fully meet the 精细化 requirements of autism specificity for indicator selection. Future research could, based on existing results, continue to refine the applicable scope of behavior, scale, imaging, and other indicators, rank specific indicators according to classification performance (accuracy, stability, etc.), and explore how different modalities can be organically combined and complement each other' s strengths, considering both internal neural mechanisms and external behavioral manifestations of autistic children to better unlock the hidden value of multi-modal data.

Finally, the transformation from theoretical models to practical application is not yet fully mature. Machine learning can help parents or other practitioners understand disease risk early and formulate rehabilitation plans promptly. However, current research remains at the theoretical stage of model construction, and several issues require careful consideration before model application to practice. First is the representativeness of research samples to the population. Participants in existing research models are typically selected, mostly from high-risk autism families, with autism proportions far exceeding normal values in typical development populations (20%-33% vs. 1%). Whether evaluation results can be directly applied to general populations requires compatibility testing (Thabtah, 2019; Rahman et al., 2020). Second is model applicability and generalizability. Limited by research feasibility factors, over one-third of current studies have sample sizes under 300, potentially creating sampling bias, and few studies report validation of results across different cultural backgrounds and geographical regions, leaving theoretical model generalizability uncertain (Mazumdar et al., 2021; Liaqat et al., 2021). Third is algorithm accuracy and stability needing improvement. Current models still mostly use traditional machine learning algorithms, with some classification accuracy below 60% and stability insufficient for large-scale testing requirements. Fourth is relatively single model function. Current research achievements can only make simple binary judgments on whether individuals have autism, with 极少 direct application to diagnosis (Millar et al., 2019). The application of computational models to prediction and evaluation process management is also superficial and insufficiently in-depth. Fifth is the need to improve application audience recognition of models. While user-friendly intelligent identification software like ASD Tests and Cognoa can be easily downloaded on mobile phones, with quick screening results obtained by uploading videos or completing test questions, the vast majority of parents, teachers, and even relevant professionals are unaware of these tools. Even if parents and relevant personnel know about these testing tools, their recognition and acceptance of tools related to identification and especially "diagnosis" results may fall short of researcher expectations. Compared to traditional methods where professional physicians conduct long-term observations and parent interviews, combining highly structured scales for careful consideration (with ADI-R alone taking 2.5 hours and containing numerous natural real-environment interpersonal interaction clues) (Abbas et al., 2020), whether machine learning models can obtain genuine and effective data while shortening testing time, whether pre-set training data can fully cover test objects, how to interpret evaluation results, and how to apply results to subsequent intervention and rehabilitation are all concerns for parents and other application audiences. As machine learning technology continues to mature, deeper thinking and demonstration of these issues will help accelerate the transformation from theoretical models to practice. Future research could, based on enriching research subjects and optimizing machine learning algorithms, fully investigate actual needs and concerns of different audiences, solve problems targeted, increase promotion of theoretical research results, and continuously deepen theoretical model penetration and promotion in practice.

## References

*Abbas, H., Garberson, F., Glover, E., & Wall, D. P. (2018). Machine learning approach for early detection of autism by combining questionnaire and home video screening.* Journal of the American Medical Informatics Association*, 25(8), 1000–1007.

*Abbas, H., Garberson, F., Liu-Mayo, S. Glover, E., & Wall, D. P. (2020). Multi-modular AI approach to streamline autism diagnosis in young children.* Scientific Reports*, 10(1), 1-8.

*Abraham, A., Milham, M., Martino, A. D., Craddock, R. C., Samaras, D., Thirion, B., & Varoquaux, G. (2016). Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example.* NeuroImage*, 147, 736–745.

*Abdolzadegan, D., Moattar, M. H., & Ghoshuni, M. (2020). A robust method for early diagnosis of autism spectrum disorder from EEG signals based on feature selection and DBSCAN method.* Biocybernetics and Biomedical Engineering*, 40(1), 482–493.

Autism Brain Imaging Exchange. (2017). Retrieved 2021, http://fcon_1000.projects.nitrc.org/indi/abide/

*Bahado-Singh, R. O., Vishweswaraiah, S., Aydas, B., Mishra, N. K., Yilmaz, A., Guda, C., & Radhakrishna, U. (2019). Artificial intelligence analysis of newborn leucocyte epigenomic markers for the prediction of autism.* Brain Research*, 1724, 146457.

*Bone, D., Bishop, S., Black, M. P., Goodwin, M. S., Lord, C., & Narayanan, S. S. (2016). Use of machine learning to improve autism screening and diagnostic instruments: Effectiveness, efficiency, and multi-instrument fusion.* Journal of Child Psychology and Psychiatry*, 57(8), 927–937.

*Bosl, W. J., Tager-Flusberg, H., & Nelson, C. A. (2018). EEG analytics for early detection of autism spectrum disorder: A data-driven approach.* Scientific Reports*, 8(1), 6828.

*Bussu, G., Jones, E., Charman, T., Johnson, M. H., & Buitelaar, J. K. (2018). Prediction of autism at 3 years from behavioural and developmental measures in high-risk infants: A longitudinal cross-domain classifier analysis.* Journal of Autism and Developmental Disorders*, 48(7), 2418–2433.

*Caly, H., Rabiei, H., Coste-Mazeau, P., Hantz, S., Alain, S., Eyraud, J. L., ⋯ Ben-Ari, Y. (2021). Machine learning analysis of pregnancy data enables early* *identification of a subpopulation of newborns with ASD.* Scientific Reports*, 11(1), 1-14.

Centers for Disease Control and Prevention. (2021). Prevalence and characteristics of autism spectrum disorder among children aged 8 years —autism and developmental disabilities monitoring network, 11 sites, United States, 2018. Retrieved January 5, 2022, from https://www.cdc.gov/mmwr/volumes/70/ss/ss7011a1.htm

Choi, G. B., Yim, Y. S., Wong, H., Kim, S., Kim, H., Kim, S. V., ···Huh, J. R. (2016). The maternal interleukin-17a pathway in mice promotes autism-like phenotypes in offspring. *Science*, 351(6276), 933–939.

Cloarec, R., Riffault, B., Dufour, A., Rabiei, H., Gouty-Colomer, L. A., Dumon, C., ···Ben-Ari, Y. (2019). Pyramidal neuron growth and increased hippocampal volume during labor and birth in autism. *Science Advances*, 5(1), eaav0394.

*Dickinson, A., Daniel, M., Marin, A., Gaonkar, B., Dapretto, M., McDonald, N. M., & Jeste, S. (2020). Multivariate neural connectivity patterns in early infancy predict later autism symptoms.* Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 6(1), 59–69.

*Dong, H. Y., Chen, D., Zhang, L., Ke, H. J., & Li, X. L. (2021). Subject sensitive EEG discrimination with fast reconstructable CNN driven by reinforcement learning: A case study of ASD evaluation.* Neurocomputing*, 449, 180–191.

Dubreuil-Vall, L., Ruffini, G., & Camprodon, J. A. (2020). Deep learning convolutional neural networks discriminate adult ADHD from healthy individuals on the basis of event-related spectral EEG. *Frontiers in Neuroscience*, 14, 251.

*Duda, M., Ma, R., Haber, N., & Wall, D. P. (2016a). Use of machine learning for behavioral distinction of autism and ADHD.* Translational Psychiatry*, 6(2), 1–5.

*Duda, M., Daniels, J., & Wall, D. P. (2016b). Clinical evaluation of a novel and mobile autism risk assessment.* Journal of Autism and Developmental Disorders*, 46(6), 1953–1961.

*Duda, M., Haber, N., Daniels, J., & Wall, D. P. (2017). Crowdsourced validation of a machine-learning classification system for autism and ADHD.* Translational Psychiatry*, 7(5), e1133.

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., ···the BASIS Team. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology*, 22(4), 338–342.

*Emerson, R. W., Adams, C., Nishino, T., Hazlett, H. C., Wolff, J. J., Zwaigenbaum, L., ···Piven, J. (2017). Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age.* Science Translational Medicine*, 9(393). https://doi.org/10.1126/scitranslmed.aag2882

*Eni, M., Dinstein, I., Ilan, M., Menashe, I., Meiri, G., & Zigel, Y. (2020). Estimating autism severity in young children from speech signals using a deep* *neural network.* IEEE Access*, 8, 139489–139500.

*Epalle, T. M., Song, Y., Liu, Z., & Lu, H. (2021). Multi-atlas classification of autism spectrum disorder with hinge loss trained deep architectures: ABIDE I results.* Applied Soft Computing*, 107(3), 107375.

*Eugenia, C., Alessandra, R., Letizia, P., Spera, G., Bosco, P., Biagi, L., ···Calderoni, S. (2020). Autism spectrum disorder and childhood apraxia of speech: Early language-related hallmarks across structural MRI study.* Journal of Personalized Medicine*, 10(4), 275.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: A systematic literature review. *European Child and Adolescent Psychiatry*, 22(6), 329–340.

Food and Drug Administration. (2021). FDA authorizes marketing of diagnostic aid for autism spectrum disorder. Retrieved October 21, 2021, from https://www.fda.gov/news-events/press-announcements/fda-authorizes-marketing-diagnostic-aid-autism-spectrum-disorder

*Gök, M. (2018). A novel machine learning model to predict autism spectrum disorder risk gene.* Neural Computing and Applications*, 31(4), 6711–6717.

*Grossi, E., Olivieria, C., & Buscema, M. (2017). Diagnosis of autism through EEG processed by advanced computational algorithms: A pilot study.* Computer Methods and Programs in Biomedicine*, 142, 73–79.

Guinchat, V., Chamak, B., Bonniau, B., Bodeau, N., Perisse, D., Cohen, D., & Danion, A. (2012). Very early signs of autism reported by parents include many concerns not specific to autism criteria. *Research in Autism Spectrum Disorders*, 6(2), 589–601.

*Guo, X. Y., Dominick, K. C., Minai, A. A., Li, H. L., Erickson, C. A., & Lu, L. J. (2017). Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method.* Frontiers in Neuroscience*, 11, 460.

Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., ··· The IBIS Network. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature*, 542(7641), 348–351.

*Kang, J. N., Han, X. Y., Song, J. J., Niu, Z. K., & Li, X. L. (2020). The identification of children with autism spectrum disorder by SVM approach on EEG and eye-tracking data.* Computers in Biology and Medicine*, 120(24), 103722.

*Kanne, S. M., Carpenter, L. A., & Warren, Z. (2018). Screening in toddlers and preschoolers at risk for autism spectrum disorder: Evaluating a novel mobile-health screening tool.* Autism Research*, 11(7), 1038–1049.

*Katuwal, G. J., Baum, S. A., Cahill, N. D., & Michael, A. M. (2016). Divide and conquer: Sub-grouping of ASD improves ASD detection based on brain* *morphometry.* Plos One*, 11(4), 1-24.

Klin, A., Lang, J., Cicchetti, D., & Volkmar, F. R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders*, 30(2), 163-167.

*Lai, M., Lee, J., Chiu, S., Charm, J., So, W. Y., Yuen, F. P., ···Zee, B. (2020). A machine learning approach for retinal images analysis as an objective screening method for children with autism spectrum disorder.* EClinicalMedicine*, 28, 100588.

*Levy, S., Duda, M., Haber, N., & Wall, D. P. (2017). Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism.* Molecular Autism*, 8(1), 65.

*Li, X., La, R., Wang, Y., Niu, J., Zeng, S., Sun, S., & Zhu, J. (2019). EEG-based mild depression recognition using convolutional neural network.* Medical and Biological Engineering and Computing*, 57(6), 1341-1352.

*Li, Y., Mache, M. A., & Todd, T. A. (2020). Automated identification of postural control for children with autism spectrum disorder using a machine learning approach.* Journal of Biomechanics*, 113, 110073.

*Liaqat, S., Wu, C., Duggirala, P. R., Cheung, S. S., Chuah, C. N., Ozonoff, S., & Young, G. (2021). Predicting ASD diagnosis in children with synthetic and image-based eye gaze data.* Signal Processing Image Communication*, 94(3), 116198.

*Liu, W., Li, M., & Li, Y. (2016). Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework.* Autism Research*, 9(8), 888-898.

*Mazumdar, P., Arru, G., & Battisti, F. (2021). Early detection of children with autism spectrum disorder based on visual exploration of images.* Signal Processing Image Communication*, 94(3), 116184.

Millar, L., Mcconnachie, A., Minnis, H., Wilson, P., Thompson, L., Anzulewicz, A., ···Delafield-Butt, J. (2019). Phase 3 diagnostic evaluation of a smart tablet serious game to identify autism in 760 children 3-5 years old in Sweden and the United Kingdom. *BMJ Open*, 9(7), e026226.

Namdeo, A., & Singh, D. (2021). Challenges in evolutionary algorithm to find optimal parameters of SVM: A review. *Materials Today: Proceedings*, 3, 288.

*Perera, H., Jeewandara, K. C., Seneviratne, S., & Guruge, C. (2017). Culturally adapted pictorial screening tool for autism spectrum disorder: A new approach.* World Journal of Clinical Pediatrics*, 6(1), 45-51.

*Qiu, N., Tang, C., Zhai, M., Huang, W., Weng, J., Li, C., ···Ke, X. (2020). Application of the still-face paradigm in early screening for high-risk autism spectrum disorder in infants and toddlers.* Frontiers in Pediatrics*, 8, 290.

*Rahman, R., Kodesh, A., Levine, S. Z., Sandin, S., Reichenberg, A., & Schlessinger, A. (2020). Identification of newborns at risk for autism using electronic medical records and machine learning.* European Psychiatry*, 63(1), e22, 1-7.

*Raj, S., & Masood, S. (2020). Analysis and detection of autism spectrum disorder using machine learning techniques.* Procedia Computer Science*, 167, 994-1004.

*Rakić, M., Cabezas, M., Kushibar, K., Oliver, A., & Lladó, X. (2020). Improving the detection of autism spectrum disorder by combining structural and functional MRI information.* NeuroImage: Clinical*, 25, 102181.

Ramírez-Duque, A. A., Frizera-Neto, A., & Bastos, T. F. (2019). Robot-assisted autism spectrum disorder diagnostic based on artificial reasoning. *Journal of Intelligent and Robotic Systems*, 96(2), 267-281.

*Salari, V., Sherkaghanad, Z., Akhondzadeh, M., Salari, S., Zomorodimoghadam, M., Abdar, M., & Khosrowabadi, R. (2019). Automated detection of autism spectrum disorder using convolutional neural network.* Frontiers in Neuroscience*, 13, 1325.

Sanders, S. J., Ercan-Sencicek, A. G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., ···State, M. W. (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*, 70(5), 863-885.

*Shahamiri, S. R., & Thabtah, F. (2020). Autism AI: A new autism screening system based on artificial intelligence.* Cognitive Computation*, 12, 766-777.

*Sen, B., Borle, N. C., Greiner, R., Brown, M. R. G., & Bernhardt, B. (2018). A general prediction model for the detection of ADHD and autism using structural and functional MRI.* Plos One*, 13(4), 1-28.

Sheldrick, R. C., Maye, M. P., & Carter, A. S. (2017). Age at first identification of autism spectrum disorder: An analysis of two US surveys. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(4), 313-320.

Shic, F., Macari, S., & Chawarska, K. (2014). Speech disturbs face scanning in 6-month-old infants who develop autism spectrum disorder. *Biological Psychiatry*, 75(3), 231-237.

*Tariq, Q., Daniels, J., Schwartz, J. N., Washington, P., Kalantarian, H., & Wall, D. P. (2018). Mobile detection of autism through machine learning on home video: A development and prospective validation study.* Plos Medicine*, 15(11), e1002705.

*Thabtah, F., Kamalov, F., & Rajab, K. (2018). A new computational intelligence approach to detect autistic features for autism screening.* International Journal of Medical Informatics*, 117(SEP.), 112-124.

*Thabtah, F. (2019). An accessible and efficient autism screening method for behavioural data and predictive analyses.* Health Informatics Journal*, 25(2), 1-17.

*Thabtah, F., Abdelhamid, N., & Peebles, D. (2019). A machine learning autism classification based on logistic regression analysis.* Health Information Science and Systems*, 7(1), 12.

Tyzio, R., Nardou, R., Ferrari, D. C., Tsintsadze, T., Shahrokhi, A., Eftekhari, S., ···Ben-Ari, Y. (2014). Oxytocin-mediated GABA inhibition during delivery attenuates autism pathogenesis in rodent offspring. *Science*, 343(6171), 675-679.

*Wan, G., Kong, X., Sun, B., Yu, S., Tu, Y., Park, J., ···Kong, J. (2019). Applying eye tracking to identify autism spectrum disorder in children.* Journal of Autism and Developmental Disorders*, 49(1), 209-215.

Wolff, J. J., Gu, H., Gerig, G., Elison, J. T., Styner, M., Gouttard, S., ···the IBIS Network. (2012). Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *American Journal of Psychiatry*, 169(6), 589-600.

Wolff, J. J., Gerig, G., Lewis, J. D., Soda, T., Styner, M. A., Vachet, C., ···the IBIS Network. (2015). Altered corpus callosum morphology associated with autism over the first 2 years of life. *Brain*, 138(7), 2046-2058.

*Wedyan, M., Al-Jumaily, A., & Crippa, A. (2019). Using machine learning to perform early diagnosis of autism spectrum disorder based on simple upper limb movements.* International Journal of Hybrid Intelligent Systems*, 15(5), 1-12.

*Wingfield, B., Miller, S., Yogarajah, P., Kerr, D., Gardiner, B., Seneviratne, S., ···Coleman, S. (2020). A predictive model for paediatric autism screening.* Health Informatics Journal*, 26(4), 2538-2553.

## Appendix 1. Summary of Machine Learning Algorithms and Classification Model Performance

[Note: The table contains detailed information about studies including sample sizes, ages, countries, data indicators, algorithms/methods, and performance metrics. Due to its extensive length and specialized content, it is preserved in its original format as a reference table for researchers.]

*Note: Figure translations are in progress. See original paper for figures.*