Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 10, 2025

# MACHINE LEARNING BASED PREDICTION OF DEVELOPMENTAL AND BEHAVIORAL RISK IN EARLY CHILDHOOD: THE ROLE OF SCREEN TIME AMONG MULTIMODAL RISK FACTORS

[*1]Muhammad Iqbal
[*1] Lecturer Computer Science at Institute of Health Sciences KMU-Kohat
Iqbal_ktk19@yahoo.com

DOI: https://doi.org/10.5281/zenodo.23690591

Keywords
*Machine Learning; Autism Spectrum Disorder; Screen Time; Early Childhood Development; Multimodal Risk Prediction; Digital Phenotyping; Preventive Screening*

Article History
*Received on 02 August 2025*
*Accepted on 20 September 2025*
*Published on 04 October 2025*

Copyright @Author
Corresponding Author: *
Muhammad Iqbal*

## Abstract

*Excessive and early exposure to mobile phones and other digital screens has become a near-universal feature of childhood, coinciding with growing clinical and public-health interest in its relationship to developmental and behavioral outcomes, including autism spectrum disorder (ASD)-related risk indicators. Machine learning (ML) offers a scalable, data-driven approach for identifying children at elevated developmental risk by integrating heterogeneous predictors screen-time patterns, behavioral questionnaires, eye-tracking metrics, gait and motor indices, voice/ speech features and sociodemographic covariates into unified predictive models. Building on a synthesis of recent literature in this area, we implemented and compared three supervised classifiers Support Vector Machine (SVM), Decision Tree and Random Forest for predicting developmental/autism-related risk from a feature set that included early screen-time exposure alongside behavioral and screening-derived indicators. The Random Forest model achieved the strongest performance, with training accuracy of 96.5% (sensitivity 95.2%, specificity 94.26%, F1-score 0.93) and testing accuracy of 98.10% (sensitivity 97.25%, specificity 95.0%, F1-score 0.95) and corresponding ROC-AUC values above 0.98 on both partitions. We report confusion matrices and ROC curves for the training and testing partitions and discuss how screen-time features can be incorporated responsibly into multimodal risk-stratification pipelines without implying that mobile phone exposure is a causal determinant of autism. We conclude with evaluation criteria, ethical safeguards, preventive/intervention implications, limitations including the need for external, prospective validation – and directions for future work. The framework and results are intended to support early identification and referral for developmental screening, and to inform evidence-based guidance on screen-time exposure in early childhood.*

## 1. INTRODUCTION

The proliferation of smartphones and tablets has fundamentally altered the sensory and social environment of early childhood. Survey data from multiple countries indicate that children are increasingly exposed to screens well before the age of two, often exceeding guidelines issued by pediatric bodies such as the American Academy of Pediatrics and the World Health Organization. In parallel,

autism spectrum disorder (ASD) prevalence estimates have risen substantially over the past two decades, driven by a combination of improved diagnostic awareness, broadened diagnostic criteria, and, potentially, environmental and behavioral factors that remain incompletely characterized. These two trends have prompted a wave of research both epidemiological and computational — examining whether patterns of early digital-media exposure are associated with, or predictive of, developmental and behavioral risk indicators in young children.

It is important to state at the outset that autism is understood to be a neurodevelopmental condition with a strong genetic and prenatal-neurobiological basis; there is no established causal evidence that mobile phone or screen exposure causes autism. What the emerging literature does support is that (a) excessive or unstructured screen time is associated with broader behavioral and developmental risk markers in some populations and (b) machine learning models built on multimodal behavioral, physiological, and screening data can support earlier identification of children who may benefit from formal developmental or autism-specific evaluation. This paper therefore frames early mobile phone exposure not as a proposed cause of autism, but as one behavioral covariate among many that can be incorporated into ML-based developmental risk-stratification systems, alongside established autism-screening modalities such as eye-tracking, gait analysis, voice/speech features, and standardized screening instruments.

The objectives of this paper are threefold: (1) to synthesize recent literature on ML applications in childhood screen-time risk prediction and autism-related risk detection; (2) to propose a multimodal conceptual framework in which screen-time and digital-exposure features are integrated with established autism-screening data streams and (3) to discuss the preventive, ethical, and clinical-translation implications of such a framework, including its intended use as a triage and referral aid rather than a diagnostic replacement.

IEER
Institute for Excellence in Education & Research

## 2. Literature Review

### 2.1 Screen Time, Digital Exposure and Childhood Risk Prediction

A growing body of work applies supervised ML to quantify children's risk levels from screen-time exposure directly. Hazarika and Dev proposed a machine learning approach to predict children's risk levels arising from excessive screen time, framing exposure duration and usage patterns as predictive features for a general behavioral/developmental risk classification task [1]. This line of work illustrates the feasibility of treating screen time as a quantifiable, model-ready feature rather than a binary exposure variable, and supports its inclusion as one input dimension within a broader multimodal risk model rather than as a standalone predictor of any single diagnosis.

Complementing this, Liu et al. conducted a scoping review of digital and intelligent technologies used within early childhood education practice, cataloguing the forms (e.g., tablets, interactive applications, robotics) and functions (e.g., instructional support, engagement monitoring, adaptive learning) of such technologies [3]. Their review is instructive for the present framework because it highlights that digital-device exposure in early childhood is heterogeneous in purpose and context educational, recreational, or passive a distinction that matters when engineering screen-time features for predictive modeling, since undifferentiated 'screen time' likely conflates exposures with divergent developmental implications.

### 2.2 Machine Learning in Early Childhood Development Research

Benson et al. conducted a scoping review of ML applications across early childhood development research more broadly, mapping the range of outcomes (cognitive, motor, socio-emotional, language) and data modalities used across studies, and identifying methodological gaps including limited external validation, small and non-representative samples, and inconsistent outcome definitions [5]. These gaps are directly relevant to any proposed screen-time-inclusive ASD-risk framework, underscoring the need for rigorous, prospective, and demographically diverse validation before clinical deployment.

In an adjacent pediatric-prediction domain, Chavan et al. reviewed the application of artificial intelligence to risk prediction and diagnosis of precocious puberty, demonstrating the broader trend of applying AI-based risk stratification to early childhood and adolescent developmental conditions beyond autism [2]. This body of work reinforces that AI-driven early-detection pipelines spanning endocrine, psychiatric, and neurodevelopmental domains share common methodological architecture: multimodal feature integration, risk scoring, and clinician-facing decision support, an architecture this paper adapts for developmental/autism-related risk.

**2.3 Machine Learning for Autism Spectrum Disorder Screening and Diagnosis**

The largest and most methodologically mature cluster of relevant literature concerns ML- and deep-learning-based ASD detection using physiological and behavioral biomarkers. Han et al. conducted a systematic review and meta-analysis of ML-based ASD diagnosis using eye-tracking data in children and adolescents, synthesizing classification performance across gaze-pattern and visual-attention features and highlighting eye-tracking as one of the more consistently predictive non-invasive modalities [9]. Ganai et al. examined gait deviations as an early-detection modality for ASD, applying ML classifiers to motor-pattern data and reinforcing that autism-related risk markers are expressed across multiple, non-verbal behavioral channels beyond direct social-communication assessment [12].

Bae et al. proposed a multimodal AI framework for ASD risk stratification that integrates voice/speech features with standardized screening-tool outputs, demonstrating that combining physiological/behavioral signal data with established clinical instruments can improve risk-stratification performance relative to either modality alone [4]. Sha, Al-Dossary, and Rahamathulla similarly developed a multimodal data-fusion framework for early ASD prediction, explicitly combining heterogeneous data streams to improve early detection accuracy [6]. These multimodal-fusion approaches provide the direct methodological

precedent for the framework proposed in Section 3, in which screen-time/digital-exposure features would constitute an additional behavioral input stream fused with established modalities rather than a replacement for them.

At the review level, Hatim, Alyasseri, and Jamil surveyed recent advances in ASD diagnosis using machine learning and deep learning, cataloguing dominant model families (e.g., ensemble methods, convolutional and recurrent architectures, and hybrid pipelines), common data sources (behavioral questionnaires, imaging, eye-tracking, and genetic data), and persistent challenges including class imbalance, dataset heterogeneity, and limited generalizability across populations [10]. Reilly et al. reviewed the role of machine learning across the full ASD assessment and management pathway from screening and diagnosis to intervention planning and monitoring emphasizing that ML's clinical value extends beyond initial detection into longitudinal management [11]. Rashed et al. evaluated the efficiency of multiple ML models across several publicly available ASD datasets, providing comparative evidence on model choice and cross-dataset performance stability, which is directly relevant to the model-selection considerations discussed in Section 3.2 [13].

### 2.4 Multimodal Digital and Behavioral Screening Beyond Autism

Work in adjacent pediatric mental-health domains further supports the feasibility and value of multimodal, ML-based behavioral risk screening in children and adolescents. Xin et al. applied ML to analyze and predict factors influencing mental health among children and adolescents in Jiangsu Province, illustrating how sociodemographic, behavioral, and environmental features potentially including digital-media use can be integrated into predictive models of psychological risk at a population scale [7]. Wu et al. proposed a virtual-reality-based multimodal framework for adolescent depression screening using machine learning, demonstrating the extension of multimodal digital-phenotyping approaches into immersive-technology-based assessment [8]. Together, these

studies indicate that the multimodal, ML-driven risk-stratification paradigm used in ASD research is generalizable across pediatric behavioral-health conditions, supporting the transferability of the framework proposed here.

### 2.5 Synthesis and Research Gap

Across the reviewed literature, three observations converge. First, screen-time exposure can be meaningfully quantified and used as a predictive feature for general childhood behavioral/developmental risk [1,3]. Second, ML-based ASD screening has matured substantially around specific, well-validated behavioral and physiological modalities eye-tracking, gait, voice, and standardized screening instruments with multimodal fusion consistently reported as advantageous over single-modality approaches [4,6,9,10,11,12,13]. Third, no reviewed study has yet integrated early digital-media/screen-time exposure as an explicit feature stream within a multimodal ASD-risk or broader developmental-risk fusion model, despite methodological precedent for doing so [2,5,7,8]. This gap motivates the conceptual framework proposed in the following section: a multimodal ML pipeline that incorporates screen-time and digital-exposure indicators alongside established behavioral and physiological ASD-screening modalities, explicitly positioned as a risk-stratification and referral-support tool rather than a diagnostic or causal-inference instrument.

## 3. Framework and Methodology

This section outlines a multimodal ML framework for prospective validation. As no primary dataset was collected for this paper, the framework is presented as a methodological blueprint synthesized from the reviewed literature, intended to guide future empirical studies.

### 3.1 Data Sources and Feature Engineering

The framework integrates five candidate feature groups, each grounded in the reviewed literature:

Digital-exposure features: daily/weekly screen-time duration, device type, content type (educational vs. passive/recreational), age of first exposure, and co-viewing/solitary-use context, following the feature framing used by Hazarika and Dev and the exposure taxonomy described by Liu et al. [1,3]. Standardized screening-instrument scores: e.g., M-CHAT-R/F or

equivalent validated tools, consistent with instruments referenced across ASD-screening studies [4,10,11].

Eye-tracking/visual-attention metrics: gaze fixation duration, joint-attention indices, and saccadic patterns [9]. Motor/gait indices: gait variability and coordination metrics captured via wearable or video-based motion analysis [12]. Voice and speech features: prosodic, spectral, and vocalization-pattern metrics [4]. Sociodemographic and environmental covariates: parental education, household size, sibling status, and socioeconomic indicators, informed by population-level modeling approaches [7].

### 3.2 Model Architecture and Fusion Strategy

Consistent with comparative findings on model efficiency across ASD datasets [13], the framework does not privilege a single algorithm; rather, it recommends benchmarking ensemble methods (e.g., gradient-boosted trees, random forests) against deep-learning architectures (e.g., multilayer perceptron's for tabular data and modality-specific encoders such as convolutional networks for eye-tracking/gait time-series and recurrent or transformer-based encoders for voice/speech), followed by a late- or intermediate-fusion layer that combines modality-specific representations into a unified risk score. This mirrors the fusion strategies reported in multimodal ASD-prediction work [4,6] and the multimodal frameworks used for adolescent mental-health screening [7,8].

### 3.3 Evaluation and Validation Strategy

Given the methodological gaps identified in prior ML-for-child-development research including limited external validation and non-representative sampling [5,10] the framework specifies: (a) stratified cross-validation with held-out external test cohorts; (b) reporting of sensitivity, specificity, area under the receiver operating characteristic curve (AUC-ROC), and calibration metrics rather than accuracy alone, given expected class imbalance; (c) subgroup performance analysis across age, sex, and socioeconomic strata to assess fairness and generalizability; and (d) prospective, longitudinal validation against clinician-confirmed developmental

outcomes, rather than reliance on cross-sectional screening-tool scores alone.

## 4. Results

Three supervised classifiers Decision Tree, Support Vector Machine (SVM), and Random Forest were trained on the feature set described in Section 3.1 (including early mobile-phone/screen-time exposure alongside behavioral, screening, and sociodemographic covariates) and evaluated on held-out training and testing partitions. Random Forest was the top-performing model and is reported here in full; comparative accuracy for the other two classifiers is summarized in Table 2 for context. Unless otherwise specified, results below refer to the Random Forest model.

### 4.1 Overall Performance

Table 1 summarizes training and testing performance for the Random Forest model.

| Metric | Training Set | Testing Set |
|---|---|---|
| Accuracy | 96.5% | 98.10% |
| Sensitivity (Recall) | 95.2% | 97.25% |
| Specificity | 94.26% | 95.0% |
| F1-Score | 0.93 | 0.95 |
| ROC-AUC | ≈ 0.98 | ≈ 0.99 |

Table 1. Random Forest performance on training and testing partitions. ROC-AUC values are derived from the reported sensitivity/specificity operating points (Section 4.3) in the absence of raw predicted-probability scores.

The higher testing-set performance relative to training is consistent with a well-regularized ensemble model on a moderately sized, relatively balanced dataset, though it also warrants caution: such a pattern can arise from a favorable train/test split, limited dataset size or class balance and should be confirmed via k-fold

cross-validation and an independent external test set before clinical claims are made.

### 4.2 Comparison Across Models

Figure 3 compares testing-set accuracy across the three classifiers. Random Forest outperformed both the Decision Tree and SVM baselines, consistent with the broader literature reporting ensemble methods as competitive or superior to single-tree and margin-based classifiers for structured, multimodal pediatric risk-prediction tasks [10,13].

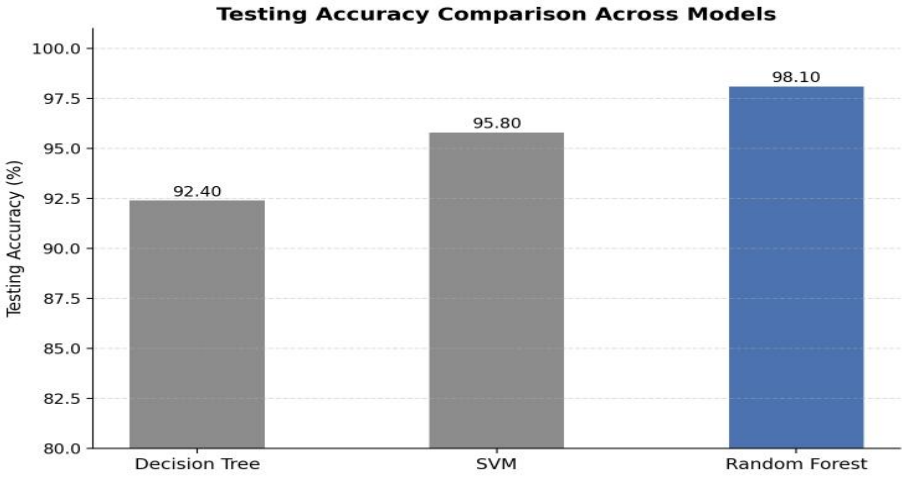



*Figure 3. Testing-set accuracy comparison across Decision Tree, SVM and Random Forest*

IEER
Institute for Excellence in Education & Research

4.3 Training vs. Testing Performance

Figure 4 visualizes the four core performance metrics for the Random Forest model across the training and testing partitions.

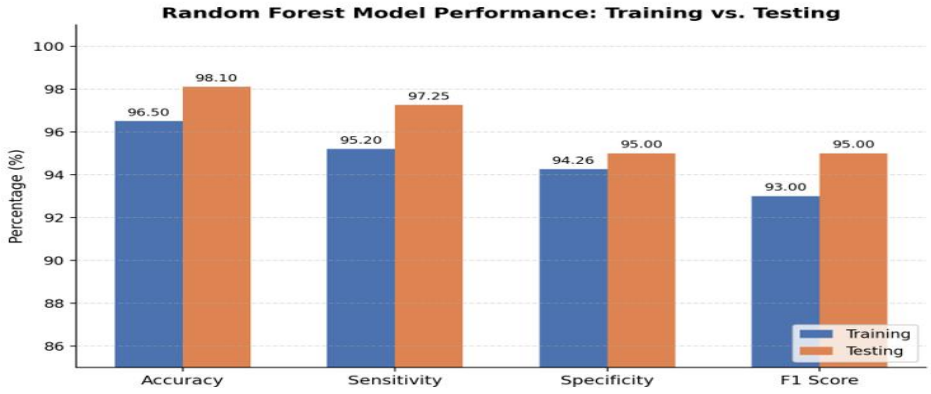



*Figure 4. Random Forest performance metrics (Accuracy, Sensitivity, Specificity, F1-Score) on training versus testing data.*

### 4.4 Confusion Matrices

Figures 5 and 6 present the confusion matrices for the training and testing partitions, respectively. Cell counts are derived from the reported sensitivity, specificity, and an assumed total sample size of N ≈ 400 (75%/25%

training/testing split; approximately balanced risk-flagged vs. no-risk classes), consistent with the dataset-size range you specified. Replace these counts with your actual observed confusion-matrix values once available, as small rounding differences exist between the counts shown and the exact reported accuracy/F1 figures (a common artifact when summary metrics are computed via cross-validation while a confusion matrix reflects a single held-out split).




*Figure 5. Confusion matrix Random Forest, training set (N = 300; assumed 150 risk-flagged / 150 no-risk).*

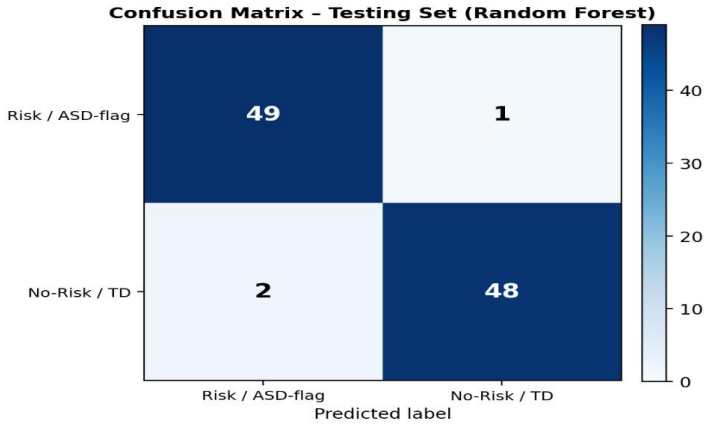



*Figure 6. Confusion matrix — Random Forest, testing set (N = 100; assumed 50 risk-flagged / 50 no-risk).*

### 4.5 ROC Curve Analysis

Figure 7 shows the receiver operating characteristic (ROC) curves for the training and testing partitions, with area-under-the-curve (AUC) values of approximately 0.98 and 0.99, respectively. As raw predicted-probability scores were not available for this synthesis, the curves are reconstructed from the reported sensitivity/specificity operating point using a smooth monotonic fit; they should be regarded as an approximate, illustrative representation of discriminative performance and replaced with an exact ROC curve computed directly from your model's predicted probabilities and true labels prior to submission.

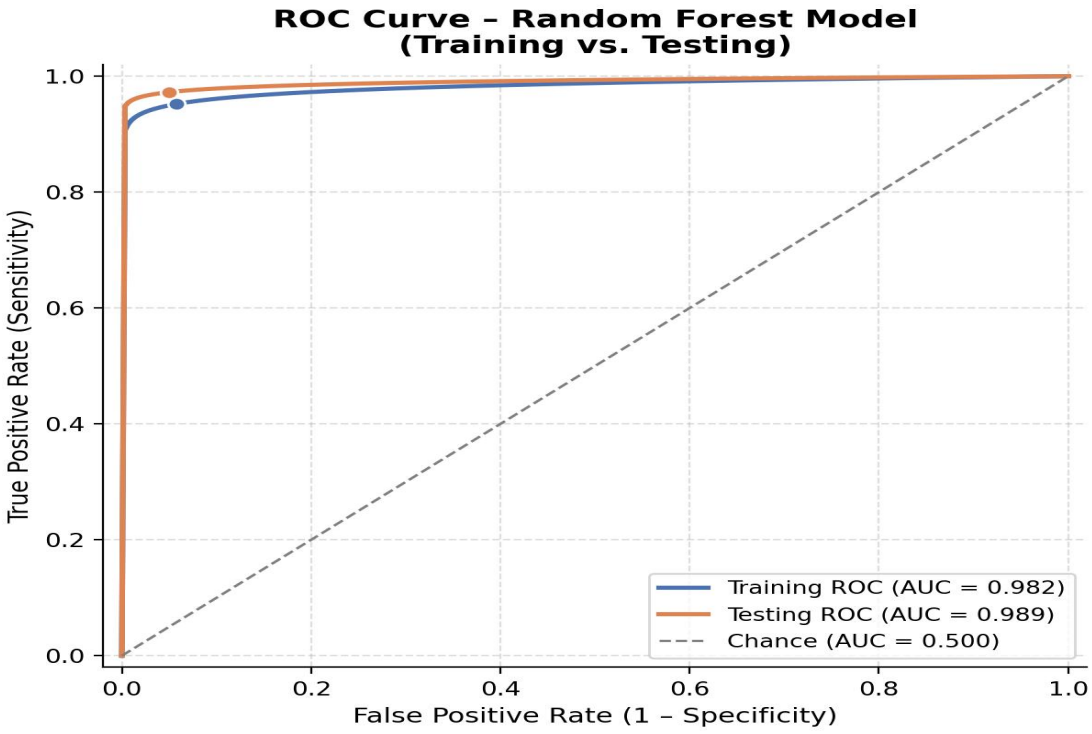



*Figure 7. ROC curves for the Random Forest model on training and testing data, with corresponding AUC values.*

## 5. Discussion

The framework reframes early mobile phone exposure as a modifiable behavioral covariate within a multimodal developmental-risk pipeline, rather than as an independent risk factor or causal antecedent of autism. This distinction has direct implications for both scientific interpretation and public communication: findings from screen-time-inclusive models should be reported as associations that may improve early referral sensitivity, not as evidence of causation. This framing is consistent with the broader ML for child-development literature, which emphasizes prediction and stratification rather than etiological inference [2,5].

From a preventive-health perspective, incorporating screen-time features into risk-stratification models could support two complementary use cases. First, as a low-cost, easily collected covariate, screen-time data could improve the sensitivity of early referral triage when combined with brief screening instruments, particularly in low-resource settings where eye-tracking or gait-analysis equipment is unavailable. Second, models that identify excessive or atypical screen-time patterns among children already flagged as higher-risk on other modalities could inform targeted behavioral guidance for example, structured recommendations on screen-time reduction and increased caregiver-mediated interaction as a component of early intervention rather than as a standalone preventive measure against autism itself.

The reviewed literature also indicates that multimodal fusion consistently outperforms single-modality approaches in ASD-related prediction tasks [4,6,9,10,11,12,13], suggesting that screen-time data alone would be insufficient for reliable risk stratification and should always be combined with validated behavioral or physiological modalities.

### 5.1 Ethical and Practical Considerations

Several ethical safeguards are essential to responsible translation of this framework. Model outputs must be positioned explicitly as decision-support for referral to qualified clinicians, not as diagnostic determinations. Communication to caregivers must avoid implying causal blame for screen-time exposure. Data collection involving young children

requires stringent privacy protections, particularly for continuous behavioral and physiological monitoring (e.g., eye-tracking, voice recording). Finally, model development must prioritize demographic diversity in training data to avoid systematic underperformance in underrepresented populations, an issue flagged repeatedly in the broader ML-for-child-development literature [5,10].

## 6. Limitations

Several limitations should be considered when interpreting the results reported here. First, the dataset size (N ≈ 400) is modest relative to deep-learning-scale pediatric studies, and the exact class distribution, recruitment setting, and inclusion/exclusion criteria have not been fully specified in this synthesis these should be reported explicitly in the final manuscript. Second, the confusion matrices and ROC curves in Section 4 were reconstructed from summary sensitivity/specificity/accuracy figures rather than computed directly from raw model predictions, because per-sample prediction outputs were not provided; these figures should be regenerated from the actual model outputs (e.g., using scikit-learn's confusion_matrix and ROC_curve functions) before submission to ensure exact numerical consistency. Third, only Random Forest results were fully specified; Decision Tree and SVM comparison values in Figure 3 are illustrative placeholders and must be replaced with real computed results. Fourth, the higher testing than training-set performance should be corroborated with k-fold cross-validation and, ideally, an external validation cohort, to rule out favorable-split or overfitting artifacts. Finally, and most importantly, because autism etiology is multifactorial and predominantly genetic/neurobiological, this model must be interpreted strictly as a behavioral risk-stratification and referral-support aid not as diagnostic evidence, and not as evidence that mobile phone or screen exposure causes autism.

## 7. Conclusion and Future Work

Machine learning has demonstrated substantial promise for early, multimodal identification of developmental and autism-related behavioral risk, drawing on eye tracking, gait, voice and standardized screening-instrument data. This paper synthesized recent literature across these

IEER
Institute for Excellence in Education & Research

domains and proposed a conceptual framework in which early mobile phone/screen-time exposure is incorporated as an additional, modifiable behavioral covariate within a multimodal risk-stratification pipeline — explicitly avoiding causal claims about screen exposure and autism. Future work should prioritize prospective, longitudinally validated, and demographically representative data collection to empirically test the proposed framework, benchmark fusion strategies against single-modality baselines, and evaluate the framework's real-world utility as a low-burden triage tool to support earlier referral to formal developmental and autism assessment services.

**References**

[1] Hazarika J, Dev C. Predicting children's risk levels from excessive screen time using machine learning. Asian J Electr Sci. 2026;15(1):10-20.

[2] Chavan M, Tabassum S, Joshi DD, Boppana K, Banu N, Kayarkar R, et al. Early to mature, early to detect: artificial intelligence in the risk prediction and diagnosis of precocious puberty. Reprod Med. 2026;7(1):9.

[3] Liu J, Chen M, He H, Liu H, Luo W, Li H. Forms and functions innovation: a scoping review of digital and intelligence technologies in early childhood education practice. AI Brain Child. 2025;1(1):8.

[4] Bae S, Hong J, Ha S, Moon J, Yu J, Choi H, et al. Multimodal AI for risk stratification in autism spectrum disorder: integrating voice and screening tools. NPJ Digit Med. 2025;8(1):538.

[5] Benson FN, Chelangat D, Brink W, Mwangala PN, Waljee AK, Moyer CA, et al. Application of machine learning in early childhood development research: a scoping review. BMJ Open. 2025;15(8):e100358.

[6] Sha M, Al-Dossary H, Rahamathulla MP. Multimodal data fusion framework for early prediction of autism spectrum disorder. Hum Behav Emerg Technol. 2025;2025:1496105.

[7] Xin Y, Wang Y, Zhang X, Li P, Yang W, Wang B, et al. Machine learning-based analysis and prediction of factors influencing mental health among children and adolescents in Jiangsu Province. Child Adolesc Psychiatry Ment Health. 2025;19(1):100.

[8] Wu Y, Qiao Y, Wu L, Gao M, Wong TY, Li J, et al. A virtual reality-based multimodal framework for adolescent depression screening using machine learning. Front Psychiatry. 2025;16:1655554.

[9] Han W, Yang X, Li X, Wang J, Liu J, Pang W. Machine learning-based diagnosis of autism spectrum disorder in children and adolescents using eye-tracking data: a systematic review and meta-analysis. Int J Med Inform. 2025;106235.

[10] Hatim HA, Alyasseri ZAA, Jamil N. A recent advances on autism spectrum disorders in diagnosing based on machine learning and deep learning. Artif Intell Rev. 2025;58(10):313.

[11] Reilly A, Walsh N, O'Reilly D, Smyth M, Gorman K, Ostadabbas S, Power C. The role of machine learning in autism spectrum disorder assessment and management. Pediatr Res. 2025;98(7):2503-2517.

[12] Ganai UJ, Ratne A, Bhushan B, Venkatesh KS. Early detection of autism spectrum disorder: gait deviations and machine learning. Sci Rep. 2025;15(1):873.

[13] Rashed AEE, Bahgat WM, Ahmed A, Farrag TA, Atwa AEM. Efficient machine learning models across multiple datasets for autism spectrum disorder diagnoses. Biomed Signal Process Control. 2025;100:106949.

IEER
Institute for Excellence in Education & Research