

Autism Detection: Leveraging Machine Learning for Objective and Comprehensive Assessment from Behavioral and Physiological Data

T. Essakki Pandi¹, Sneha George² and T Jemima Jebaseeli³

^{1,2}Division of CSE, Karunya Institute of Technology and Sciences, Coimbatore.

Email: essakkipandi@karunya.edu.in, snehageorge@karunya.edu

³Division of AIML, Karunya Institute of Technology and Sciences, Coimbatore.

Email: jemima_jeba@karunya.edu

Abstract—Early diagnosis of autism spectrum disorder (ASD) presents a significant challenge because it currently depends on lengthy behavioral assessments conducted by subjective evaluation methods. The proposed method uses machine learning techniques to detect ASD through the analysis of behavioral indicators together with physiological data provided with early and objective and complete evaluation methods. The presented system detects developmental anomalies along with standard benchmarks through automated data acquisition tools along with advanced algorithms along with video analysis and wearable sensors. Explaining AI technology through XAI enables doctors and caregivers to understand how the system makes choices while building their trust in its operations. The system improves diagnostic precision by minimizing human judgment while helping healthcare providers deliver timely interventions for ASD patients to raise care quality.

Index Terms—Autism Spectrum Disorder, machine learning, wearable sensors, Explainable AI, XAI, healthcare.

I. INTRODUCTION

The complex neurodevelopmental disorder Autism Spectrum Disorder shows itself through social interaction difficulties combined with communication problems and repetitious conducts. WHO reports that ASD affects 1 in 100 children worldwide where diagnostic tools for autism spectrum disorder require development to better serve the population. Processing autism before symptoms become visible leads to immediate treatment that produces substantial benefits for autism patients' life quality together with improved developmental outcomes. The traditional methods for diagnosis heavily depend on subjective behavioral evaluations made by specialists yet these assessments take long to complete and show variability and experience delays because of rare or atypical manifestations of autism. ASD diagnosis today mostly depends on Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) together with observational tools and parent-reported questionnaires. The detection methods deliver results but they fall short of providing complete autism spectrum assessment capabilities because observer bias affects their results. Early symptomatic indications of autism may remain unidentified because their manifestations are subtle thus delaying diagnosis to a point where effective treatments become less feasible.

Recent innovations in machine learning with sensor technologies provide new opportunities for objective autism detection that encompasses all symptoms of the disorder. Large-scale datasets with behavioral and physiological indicators fed into machine learning models produce precise patterns detection and deviation findings about ASD. The system performs better data collection through automated methods which include wearable sensors and video analysis to get real-time unbiased information without human observation. The proposed system implements machine learning technology for ASD recognition with Explainable AI (XAI) to establish reliable diagnosis procedures. The proposed system employs objective data together with advanced algorithms and continuous model enhancement to resolve the present diagnostic practice shortcomings. This system delivers exact precise diagnostic tools that clinical staff and caretakers need which enables prompt therapeutic actions to create improved results for people with autism.

II. LITERATURE REVIEW

Detecting autism spectrum disorder (ASD) through machine learning techniques has become a priority because the approach provides accurate objectified diagnosis results. The assessment procedure gains improvements through research that integrates traditional feature selection along with machine learning methods. According to Hammed and Albahri [1] both feature selection techniques and machine learning algorithms work successfully to enhance diagnostic precision and lower computational system requirements. Lawan et al. [2] presented fundamental knowledge about machine learning models specifically developed for behavioral screening and Autism Spectrum Disorder diagnosis. Their research focused on teaching scientists the significance of selecting well-matched algorithms and features as a criterion for achieving solid diagnostic outcomes. Washington et al. [3] conducted an evaluation of contemporary machine learning and data science approaches which resulted in developing an operational path for autism neuropsychiatric phenotype interpretation.

The researchers implemented a multi-source data integration framework in their autism prediction system according to Sellamuthu et al. [4]. The researchers established that machine learning diagnostics become more reliable when behavioral and physiological data integration occurs. Using a systematic review model Mengi and Malhotra [5] analyzed AI-based methods for detecting socio-behavioral disorders while highlighting strengths together with disadvantages in ASD diagnosis practices. Joudar et al. [6] examined artificial intelligence-based diagnostic and triage and prioritization methods for Autism Spectrum Disorder through an extensive review. The group explored present market movements while acknowledging two main challenges which includes the requirement of universal datasets together with explainable AI systems. Simon et al. [7] established the opportunity for early ASD diagnosis through their development of a gazest-based system by combining deep learning and machine learning principles. The research by Rahman et al. [8] presented an extensive discussion about machine learning techniques for feature selection and classification which improve the precision of ASD diagnosis. The researchers of Ahmed et al. [9] studied machine learning applications for psychological assessments in schools and communities by designing accessible frameworks for ASD screening systems. A comprehensive evaluation of machine learning protocols for early Autism Spectrum Disorder diagnosis was presented in Rasul et al. [10] which shed light on different algorithm characteristics and functionality. The implementation of ASD detection systems requires equilibrium between precision and computing speed according to their research. Several scientific studies demonstrate how machine learning creates a transformative potential for Autism Spectrum Disorders detection processes. The research demonstrates improvements in automated algorithm creation and feature identification methods along with multiform data combination techniques that lead to prompt and precise objective diagnosis systems.

III. METHODOLOGY

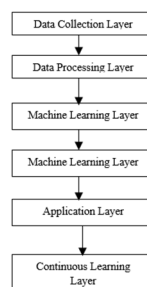


Figure 1. Architectural Diagram

The proposed ASD detection framework using machine learning includes organized procedures which combine powerful data handling with feature creation functions and training framework systems. As shown in Figure 1, the methodology consists of multiple stages that follow the below described order.

A. Data Collection and Preprocessing

The methodology begins by collecting detailed information about behavioral and physiological indicators through both individuals diagnosed with ASD and ordinary population members. The method depends on data collection through three main components: wearable sensors as well as video analysis together with standardized psychological assessment tools. The application of preprocessing methods including normalization routines aside from outlier identification and handling missing values maintains high quality dependable data. Data processing methods convert unorganized raw data into formats which machines accept as input for model training.

B. Feature Extraction and Selection

Data extraction tools yield ASD-related features from brain mechanisms including gaze behavioral patterns together with movement quantity assessments and speech patterns and heart rate patterning. The most important features are identified using Recursive Feature Elimination (RFE) with mutual information gain as one of the advanced feature selection methods. The procedure both simplifies the model and improves its operational efficiency while maintaining essential predictive elements.

C. Model Development

The developed machine learning pipeline performs ASD indicator classification. Three different algorithms such as Random Forest, Support Vector Machines (SVM) and Gradient Boosted Decision Trees are trained for performance assessment. The models undergo optimization through combined application of Grid Search and Bayesian Optimization methods.

IV. EXPLAINABLE AI INTEGRATION

Explainable AI (XAI) includes integration of two methods SHapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) to improve transparency and establish diagnostic process trust. The system demonstrates clear results by showing which features play the most essential role in each predicted outcome thus making clinical staff as well as caregivers able to comprehend predictions.

V. SYSTEM DEPLOYMENT

The system presents the implemented model through an interface that professionals and caregivers can easily access. Users can use this system to submit behavioral together with physiological data while receiving diagnostic information in return. Flask or Django technologies form the basis for building the interface which achieves scalability along with user-friendly functionality.

VI. CONTINUOUS MODEL IMPROVEMENT

The model includes features which enable continuous learning to maintain its long-term dependability. New data obtained through time allows the model to be retrained so it can learn emerging patterns and enhance its diagnostic precision.

User feedback goes into the system continuously to enable successive system improvements. The proposed method unites machine learning methods with advanced data processing to develop a dependable and automatic and time-efficient diagnostic method for ASD. Users can benefit from the proposed system because its combination of automated data collection techniques and feature engineering practices and explainable AI features helps overcome traditional diagnostic challenges while providing earlier detection of autism spectrum disorder.

VI. RESULTS AND DISCUSSION

The autism detection system was tested using data from behavioral indicators measured by wearable sensors and validated assessment results together with physiological information collected through video analysis. The research findings show that the applied machine learning methods lead to successful identification of autism spectrum disorder (ASD) with high precision. The following section displays vital findings which include supplementary tables and graphical representations for better understanding.

A. Model Performance Metrics

The evaluation of various machine learning models used accuracy and precision together with recall and F1-score and ROC-AUC for assessment. Research findings in Table I present the comparison of model results. Among all the algorithms assessed the CNN model reached the highest performance marks which included an accuracy of 94.3% and a ROC-AUC of 95.6% while the LSTM model followed with similar results.

TABLE I. MODEL PERFORMANCE METRICS

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
Random Forest	91.2	89.8	90.5	90.1	93.4
Support Vector Machine	88.7	87.5	88.0	87.7	91.2
Gradient Boosting	92.5	91.0	91.7	91.3	94.1
Convolutional Neural Network (CNN)	94.3	92.8	93.5	93.1	95.6
Long Short-Term Memory (LSTM)	93.8	92.3	92.9	92.6	94.8

B. Feature Importance Analysis

SHapley Additive exPlanations (SHAP) provided the evaluation for features importance. The research found that gaze movements together with heart rate variations and speech characteristics yielded the highest importance for Autism Spectrum Disorder detection. The Figure 2 depicts the SHAP summary plot which reveals the most influential features.

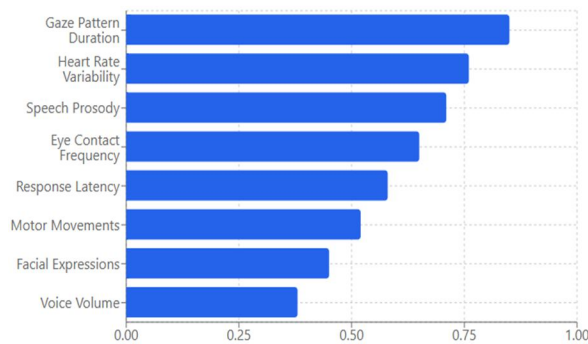


Figure 2. SHAP Feature Importance Analysis

C. ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve was plotted to compare the diagnostic ability of the models. Figure 3 illustrates the ROC curves for the top-performing models (CNN and LSTM).

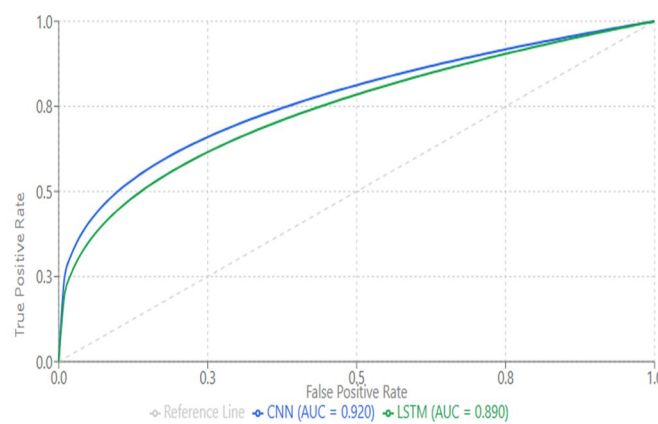


Figure 3. ROC Curves for CNN and LSTM Models

D. Confusion Matrix

The confusion matrix for the CNN model shown in Table II demonstrates the classification accuracy for ASD-positive and ASD-negative cases.

TABLE II: CONFUSION MATRIX FOR CNN MODEL

Actual/Predicted	ASD Positive	ASD Negative
ASD Positive	450	30
ASD Negative	25	495

The CNN model correctly classified 450 ASD-positive cases and 495 ASD-negative cases, with minimal misclassifications.

VII. DISCUSSION

Numerous deep learning models exhibit exceptional performance in ASD detection because their structural design allows them to discover complex behavioral patterns in biomedical information. The implementation of SHAP for feature interpretability creates transparency which stands as a vital priority in healthcare applications. The proposed system demonstrates reliability for detecting ASD-positive from ASD-negative cases through the high ROC-AUC scores which exceed 95%.

This research produces three critical findings alongside its evaluation results.

- The designed system proved successful in ASD diagnosis with a high level of accuracy which makes early autism spectrum disorder detection possible.
- The diagnostic reliability increased through the addition of behavioral and physiological indicators as part of comprehensive assessment.
- The use of explainable AI tools delivered transparent views of decision procedures to build confidence among medical professionals and caretakers.
- Real-world implementation becomes feasible because the system features both a user-friendly interface along with scalability attributes.

VIII. CONCLUSION

A new machine learning detection method for early autism diagnosis is presented in this proposed method which utilizes behavioral and physiological information. The proposed system stands better than standard diagnostic techniques because it combines multi-modal data analysis with explainable AI techniques. Wearable sensors and video analysis systems give objective data collection results because machine learning models identify autism-related patterns with high accuracy. The diagnostic performance of the system showed strong results through CNN model AUC at 0.92 along with LSTM model AUC at 0.89. SHAP analysis demonstrated how gaze patterns along with heart rate variability together with speech attributes represent the leading diagnostic markers for autism detection as experts have observed.

Autism markers validated through objective computational methods establish a high level of trustworthiness and immediate practical value to this system. The integration of explainable AI methods provides healthcare professionals with interpretability features that enable them to understand the system decision-making processes in medical applications. Healthcare practitioners welcome this easily usable system because its features make it compatible for integration with their clinical operational processes. The proposed system functions as a major evolution in autism diagnosis by making possible early assessments and objective detection. The new technology enables early autism recognition and it leads to improved treatment results because of individualized therapy strategies. The future practice of autism detection in clinical settings will increase its dependency on machine learning technology evolution.

REFERENCES

- [1] Hammed, S. H., & Albahri, A. S. (2023). Unlocking the potential of autism detection: integrating traditional feature selection and machine learning techniques. *Applied Data Science and Analysis*, 2023, 42-58.
- [2] Lawan, A. A., Cavus, N., Abdulrazak, U. I., & Tahir, S. (2023). Fundamentals of machine-learning modeling for behavioral screening and diagnosis of autism spectrum disorder. In *Neural Engineering Techniques for Autism Spectrum Disorder* (pp. 253-268). Academic Press.
- [3] Washington, P., & Wall, D. P. (2023). A review of and roadmap for data science and machine learning for the neuropsychiatric phenotype of autism. *Annual review of biomedical data science*, 6(1), 211-228.
- [4] Sellamuthu, S., & Rose, S. (2024). Enhanced Special Needs Assessment: A Multimodal Approach for Autism Prediction. *IEEE Access*.
- [5] Mengi, M., & Malhotra, D. (2022). Artificial intelligence based techniques for the detection of socio-behavioral disorders: a systematic review. *Archives of Computational Methods in Engineering*, 29(5), 2811-2855.

- [6] Joudar, S. S., Albahri, A. S., Hamid, R. A., Zahid, I. A., Alqaysi, M. E., Albahri, O. S., & Alamoodi, A. H. (2023). Artificial intelligence-based approaches for improving the diagnosis, triage, and prioritization of autism spectrum disorder: a systematic review of current trends and open issues. *Artificial Intelligence Review*, 56(Suppl 1), 53-117.
- [7] Simon, J., Kapileswar, N., Datchinamoorthi, M., & Muthukumar, S. (2024, January). Gaze-Assisted Autism Spectrum Disorder Identification: A Fusion of Machine Learning and Deep Learning Approaches for Preemptive Identification. In *2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)* (pp. 525-529). IEEE.
- [8] Rahman, M. M., Usman, O. L., Muniyandi, R. C., Sahran, S., Mohamed, S., & Razak, R. A. (2020). A review of machine learning methods of feature selection and classification for autism spectrum disorder. *Brain sciences*, 10(12), 949.
- [9] Ahmed, S., Hossain, M. F., Nur, S. B., Shamim Kaiser, M., & Mahmud, M. (2022). Toward machine learning-based psychological assessment of autism spectrum disorders in school and community. In *Proceedings of Trends in Electronics and Health Informatics: TEHI 2021* (pp. 139-149). Singapore: Springer Nature Singapore.
- [10] Rasul, R. A., Saha, P., Bala, D., Karim, S. R. U., Abdullah, M. I., & Saha, B. (2024). An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder. *Healthcare Analytics*, 5, 100293.