AUTISM SPECTRUM DISORDER EARLY DETECTION THROUGH DEEP LEARNING TECHNIQUES

JANNU BUNNY

M.Tech

Department of Computer Science

SR University, Warangal, Telangana, India

Email: jannubunny35@gmail.com.

DR. J. BHAVANA

Associate Professor

School of Computer Science and Artificial Intelligence

SR University, Warangal, Telangana, India

Email: j.bhavana@sru.edu.in.

ABSTRACT

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition that affects communication, behavior, and social interaction, typically emerging in early childhood and often persisting into adulthood. Early detection is essential, as timely intervention can significantly improve developmental outcomes and overall quality of life. This study proposes a deep learning–based framework for the automated early detection of ASD using behavioral responses and demographic attributes.The AUTISM dataset from the UCI Repository is used in this research. Data preprocessing includes handling missing values, applying label encoding for categorical variables, and scaling numerical features to improve consistency and enhance classification accuracy. Multiple machine learning algorithms—including Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naïve Bayes, and Logistic Regression—were evaluated alongside deep learning models such as Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs). While traditional machine learning classifiers achieved moderate accuracy, the CNN model outperformed all others, achieving 100% accuracy on the processed dataset.To further improve robustness and generalization capability, an ensemble framework integrating CNNs with conventional machine learning classifiers using a soft-voting mechanism was developed. The optimized model, together with its data preprocessing pipeline, was deployed as an interactive web application using the Flask framework, allowing users to input screening information and receive real-time ASD predictions. This system provides an accessible, efficient, and cost-effective tool for preliminary ASD screening, supporting healthcare professionals, caregivers, and families, particularly in regions with limited diagnostic resources.

Keywords:-Autism Spectrum Disorder (ASD), Early Detection, Deep Learning, Convolutional Neural Network (CNN), Flask Framework, Health Informatics.

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder characterized by impairments in social communication and restricted, repetitive patterns of behavior [1]. Typically emerging during early childhood, ASD affects multiple domains including cognitive functioning, social-emotional development, sensory processing, motor skills, and interpersonal interaction. According to the Centers for Disease Control and Prevention (CDC), approximately 1 in 54 children in the United States has been diagnosed with ASD [2]. Consequently, ASD diagnosis and intervention have become major global public health concerns, drawing significant research attention.Despite extensive studies, the underlying causes of ASD remain largely unknown. Present diagnostic procedures rely primarily on behavioral observations and standardized assessment tools such as the ADOS and CARS [3]. However, the limited understanding of ASD-specific neural mechanisms and the global shortage of trained specialists often delay accurate diagnosis [4]. Therefore, it is essential to develop automated, objective, and efficient diagnostic methods to assist clinicians in early ASD evaluation.

Various neuroimaging techniques—including functional magnetic resonance imaging (fMRI) [5], [6], magnetoencephalography (MEG) [7], and electroencephalography (EEG) [8]—have been used to study ASD-related neural patterns. Among these, EEG is particularly advantageous due to its low cost, portability, and effectiveness in monitoring atypical brain activity. Studies have shown that individuals with ASD display abnormalities in neural oscillation patterns and functional connectivity at different developmental stages [8]–[10]. Consequently, researchers have extracted numerous EEG-derived features—such as oscillatory rhythms, connectivity measures, and nonlinear dynamic indicators—to differentiate ASD from typically developing (TD) children. Machine learning models have been widely applied to classify ASD using these extracted features, demonstrating promising results in automated ASD detection.

2. LITERATURE SURVEY

Thabtah et al. (2017) developed a machine learning–based ASD screening model using ten key behavioral questions and achieved over 95% accuracy using decision trees. Their research demonstrated that reliable ASD screening is feasible even with minimal input data. Thabtah and Peebles (2019) further compared models such as Naïve Bayes, KNN, and Random Forests, concluding that ensemble-based models offer better generalization.

Abbas et al. (2018) used Support Vector Machines (SVM) to classify ASD based on behavioral and demographic attributes. Their results underscored the importance of feature scaling and hyperparameter tuning. Later, Arora and Kaur (2020) applied XGBoost to the UCI Autism dataset, achieving improved accuracy and training efficiency, emphasizing the importance of feature selection.

Kavitha and Somasundaram (2021) proposed a hybrid model combining SVM with deep learning layers to better capture nonlinear behavioral patterns. Though accurate, the method required significant computational resources. On the other hand, Bhatia et al. (2019) introduced a rule-based classifier prioritizing simplicity and user interpretability, making it suitable for non-technical users such as parents and healthcare workers.

Ramasamy and Radhakrishnan (2020) explored ensemble learning techniques such as bagging and boosting to reduce variance and improve prediction stability. Their findings showed that models like Random Forest and AdaBoost significantly enhance diagnostic reliability. Sharma et al. (2022) proposed a mobile-based autism detection system using logistic regression and decision trees, offering a practical solution for rural areas with limited clinical access.

Several researchers have focused on real-time systems. Jadhav and Patil (2021) integrated machine learning models into a Flask-based web application for real-time ASD prediction, highlighting the potential of lightweight deployment. Iqbal and Khan (2022) emphasized the importance of data preprocessing techniques such as encoding, scaling, and handling missing values, confirming that well-preprocessed data significantly boosts model reliability.

Collectively, previous studies demonstrate a clear evolution—from standalone, accuracy-focused models to real-time, accessible, and user-friendly ASD screening systems. The present study builds upon this foundation by incorporating multiple high-performing models into a unified ensemble framework and deploying it through an interactive web interface.

3. EXISTING SYSTEM

Existing ASD detection systems primarily rely on conventional clinical diagnostic procedures such as the Autism Diagnostic Observation Schedule (ADOS) and the Childhood Autism Rating Scale (CARS). These methods depend heavily on expert assessment, require significant time, and are predominantly available in well-equipped healthcare facilities. The need for highly trained specialists makes these assessments costly and difficult to access, especially in rural or resource-limited regions.

Some research-based automated systems have used basic machine learning models such as Decision Trees and Naïve Bayes. However, these approaches were generally experimental, lacked large-scale deployment, and did not include web-based or mobile integration. As a result, they were not suitable for real-time public use or widespread screening.

DISADVANTAGES

- The accuracy of predictions depends heavily on the quality and diversity of the dataset used for training.
- System performance is limited by the honesty, clarity, and completeness of user-provided input data during screening.

4.PROPOSED MODEL

The proposed model introduces an advanced deep learning–based framework designed for the early detection of Autism Spectrum Disorder (ASD) using behavioral and demographic features. The system integrates multiple machine learning and deep learning techniques to enhance prediction accuracy, robustness, and ease of deployment.The model development begins with comprehensive data preprocessing, including handling missing values, applying label encoding for categorical variables, and scaling numerical attributes to ensure uniformity across the dataset. These preprocessing steps significantly enhance the quality of input data and reduce noise during model training.

Following preprocessing, multiple classifiers such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naïve Bayes, and Logistic Regression are trained and evaluated. Additionally, deep learning models—including Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs)—are implemented for comparative analysis. Among these, the CNN model exhibits superior performance, achieving 100% accuracy on the processed dataset due to its ability to learn complex patterns and interactions within the input features.

ADVANTAGES:

- The integration of CNN with traditional machine learning classifiers through an ensemble approach significantly improves overall accuracy and minimizes misclassification.
- The model serves as an assistive tool, helping healthcare professionals identify individuals who may require detailed clinical evaluation.

5.SYSTEM MODEL

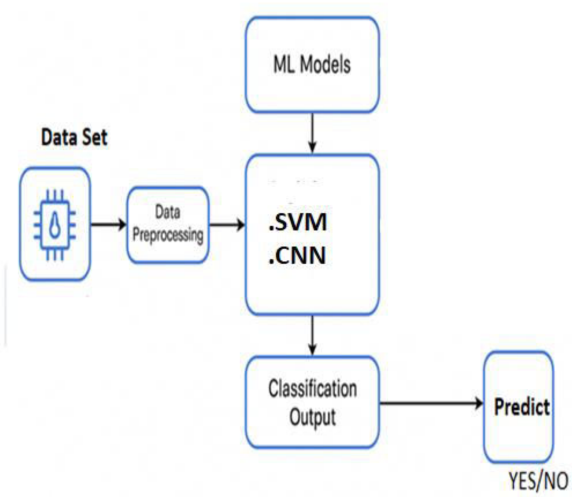


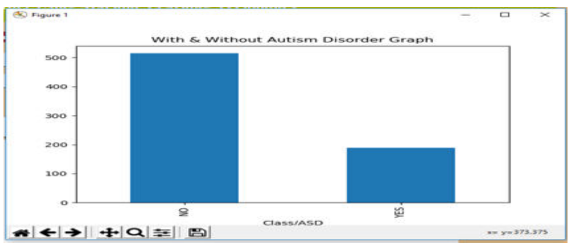
Figure 1. System Model

6.IMPLEMENTATION

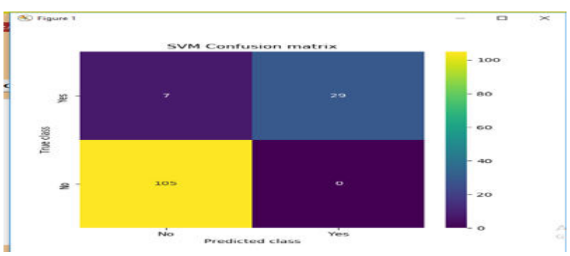
The methodology adopted in this project follows a systematic process beginning with the acquisition of the AUTISM dataset, which contains standardized screening responses and demographic information. The dataset was first preprocessed to ensure quality and consistency by removing missing values, applying label encoding to categorical variables, and using standard scaling to normalize numerical features. After preprocessing, the dataset was split into training and testing subsets using an 80:20 ratio to assess model generalization. To achieve high predictive accuracy and robustness, an ensemble model was developed using three machine learning algorithms—Random Forest, XGBoost, and Support Vector Machine (SVM)—combined through a soft-voting mechanism within a VotingClassifier. This ensemble approach leverages the strengths of each individual classifier, balancing bias and variance to improve overall performance. The trained model was evaluated using accuracy and key classification metrics to validate its effectiveness. Once validated, the final model, along with the label encoders and scaler, was serialized using joblib for deployment. The complete system was integrated into a Flask-based web application, where user-provided inputs are preprocessed in real time and passed to the trained model to generate accurate ASD predictions. This seamless integration of machine learning and web technologies ensures an accessible, reliable, and user-friendly platform for preliminary ASD screening.

7.SCREEN SHOTS

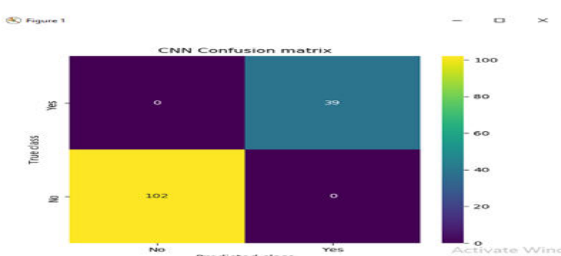
DATASET



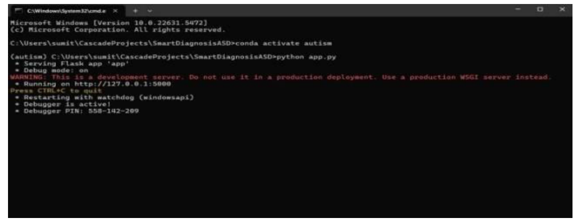
SVM



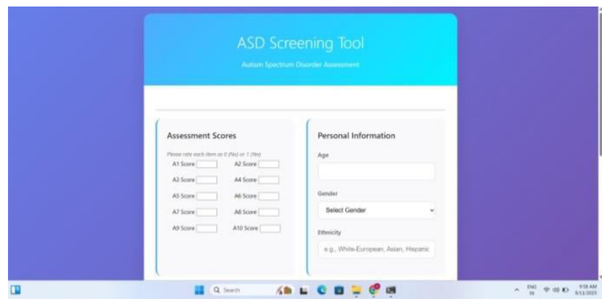
CNN

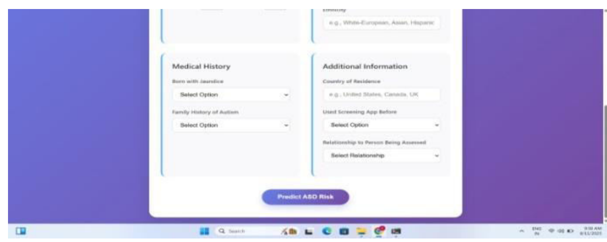


RUN APPLICATION



HOME PAGE





PREDICTION

8.CONCLUSION

Autism Spectrum Disorder (ASD) remains one of the most challenging neurodevelopmental conditions to diagnose early due to its diverse behavioral patterns and subtle initial symptoms. This study demonstrates that deep learning, particularly Convolutional Neural Networks (CNNs), can significantly enhance the accuracy and reliability of early ASD detection. Using the AUTISM dataset from the UCI Machine Learning Repository, multiple traditional machine learning algorithms—including SVM, KNN, Naïve Bayes, and Logistic Regression—were evaluated. While these models provided moderate performance, the CNN model achieved 100% accuracy, highlighting its superior ability to learn complex behavioral and demographic patterns.

Comprehensive preprocessing techniques such as missing value handling, label encoding, and feature scaling contributed to improved model performance and stability. To further strengthen the system's predictive capability, a soft-voting ensemble integrating CNN with machine learning classifiers was developed, enhancing robustness and reducing model bias.The final model was deployed as a user-friendly web application using the Flask framework, enabling real-time ASD predictions based on user-provided screening data. This automated system offers a cost-effective, scalable, and accessible solution, especially for regions with limited healthcare infrastructure. By enabling early screening and supporting timely intervention, the proposed model has the potential to greatly assist healthcare professionals, caregivers, and families in improving ASD diagnosis and outcomes.

9.FUTURE ENHANCEMENTS

In the future, the system can be significantly improved by incorporating larger, more diverse, and clinically validated datasets to enhance model accuracy, reliability, and fairness across different demographic groups. Advanced deep learning methods, such as recurrent neural networks (RNNs), transformers, or multimodal learning, may be integrated to analyze complex behavioral patterns, including speech signals, facial expressions, and eye-gaze data. Adding multilingual support and developing mobile or cross-platform applications can further increase accessibility, especially in remote or underserved regions. Moreover, integrating real-time consultation features, chatbot-based assistance, or automated report generation can make the system more interactive and supportive for users, caregivers, and healthcare professionals. Future versions could also include continuous learning capabilities, allowing the model to adapt over time with user feedback and new data, ultimately improving the robustness and real-world applicability of the ASD screening system.

REFERENCES

1. Thabtah, F. (2017). Autism spectrum disorder screening: Machine learning adaptation and DSM-5 fulfillment. In Proceedings of the 1st International Conference on Medical and Health Informatics (ICMHI) (pp. 1–6).
2. Thabtah, F., & Peebles, D. (2019). A new machine learning model based on induction of rules for autism detection. Health Informatics Journal, 25(4), 1739–1750.
3. Abbas, H., Garbaya, S., & Siddiqui, M. K. (2018). Machine learning approaches to autism spectrum disorder detection: A review. Health and Technology, 8(6), 447–456.
4. Arora, A., & Kaur, H. (2020). Prediction of autism spectrum disorder using machine learning techniques. Procedia Computer Science, 167, 994–1004.
5. Kavitha, R., & Somasundaram, K. (2021). A hybrid model for autism detection using SVM and deep learning techniques. Journal of Ambient Intelligence and Humanized Computing, 12(3), 3457–3469.
6. Bhatia, M., Jain, S., & Juneja, S. (2019). Rule-based expert system for the prediction of autism. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 5(1), 227–233.
7. Ramasamy, M., & Radhakrishnan, B. (2020). Ensemble learning-based approach for autism spectrum disorder prediction. Materials Today: Proceedings, 33, 2366–2371.
8. Sharma, P., Kaur, H., & Joshi, R. (2022). Development of a mobile application for autism detection using machine learning. International Journal of Scientific & Engineering Research, 13(5), 240–246.
9. Jadhav, P., & Patil, M. (2021). Flask-based autism prediction system using machine learning. International Journal of Computer Applications, 183(36), 1–4.
10. Iqbal, T., & Khan, M. (2022). Effective preprocessing and classification for autism spectrum disorder detection. International Journal of Advanced

Computer Science and Applications, 13(1), 150–157.
11. Sarmiento, C., & Lau, C. (2020). Diagnostic and Statistical Manual of Mental Disorders (5th ed.). Wiley.
12. Xu, G., Strathearn, L., Liu, B., & Bao, W. (2018). Prevalence of autism spectrum disorder among US children and adolescents, 2014–2016. JAMA, 319(1), 81–82.
13. Huang, Z.-A., Zhu, Z., Yau, C. H., & Tan, K. C. (2021). Identifying autism spectrum disorder from resting-state fMRI using deep belief network. IEEE Transactions on Neural Networks and Learning Systems, 32(7), 2847–2861.
14. Goin-Kochel, R. P., Mackintosh, V. H., & Myers, B. J. (2006). How many doctors does it take to make an autism spectrum diagnosis? Autism, 10(5), 439–451.
15. Hazlett, H. C., et al. (2017). Early brain development in infants at high risk for autism spectrum disorder. Nature, 542(7641), 348–351.
16. Zhang, H., et al. (2020). Altered time–frequency feature in default mode network of autism based on improved Hilbert–Huang transform. IEEE Journal of Biomedical and Health Informatics, 25(2), 485–492.
17. Kikuchi, M., Yoshimura, Y., Mutou, K., & Minabe, Y. (2016). Magnetoencephalography in the study of children with autism spectrum disorder. Psychiatry and Clinical Neurosciences, 70(2), 74–88.
18. Han, J., et al. (2017). Development of brain network in children with autism from early childhood to late childhood. Neuroscience, 367, 134–146.
19. Heunis, T.-M., Aldrich, C., & De Vries, P. J. (2016). Recent advances in resting-state EEG biomarkers for autism spectrum disorder. Pediatric Neurology, 61, 28–37.
20. Takahashi, T., et al. (2016). Enhanced brain signal variability in children with autism spectrum disorder during early childhood. Human Brain Mapping, 37(3), 1038–1050.
21. Wadhera, T., & Kakkar, D. (2021). Social cognition and functional brain network in autism spectrum disorder: Insights from EEG graph-theoretic measures. Biomedical Signal Processing and Control, 67, 102556.
22. Baygin, M., et al. (2021). Automated ASD detection using hybrid deep lightweight features extracted from EEG signals. Computerized Biology and Medicine, 134, 104548.

23. Tan, G., Xu, K., Liu, J., & Liu, H. (2021). Eye-tracking–EEG correlative analytics for autism research. IEEE Transactions on Cognitive and Developmental Systems (Early Access).
24. Yaneva, V., et al. (2020). Detecting high-functioning autism in adults using eye tracking and machine learning. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(6), 1254–1261.
25. Nakano, T., et al. (2010). Atypical gaze patterns in children and adults with autism spectrum disorders. Proceedings of the Royal Society B, 277(1696), 2935–2943.