Early Detection of Autism Spectrum Disorder Using Deep Learning

Mrs.C.Swetha*, Hasini Kanchetty, Ajay Boddu, Brahma Reddy
*(Associate Professor)
Department of Internet of Things, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad, Telangana

Abstract: **The early identification of Autism Spectrum Disorder (ASD) presents a pressing need in the global healthcare community. ASD is a neurodevelopmental disorder that affects social interaction, communication, and behavior, often resulting in long-term challenges when not addressed early. Traditional diagnostic methods, although effective, are often time-consuming, resource-intensive, and inaccessible to underprivileged communities. This research introduces a deep learning-based system designed to assist in the early screening of ASD using structured datasets derived from behavioral questionnaires and demographic data. The system implements a Multi-Layer Perceptron (MLP) model utilizing Keras, aiming for high classification accuracy with minimal false positives.**

The paper begins by providing a comprehensive background on ASD, including its symptoms, causes, and the importance of early diagnosis. It then explores current diagnostic practices and the challenges faced by healthcare professionals. The focus then shifts to the role of deep learning algorithms in transforming healthcare diagnostics and introduces the specific model architecture developed in this study. A detailed explanation of the dataset, preprocessing techniques, model structure, training methodology, and evaluation metrics is provided.

Keywords: **Autism Spectrum Disorder, Deep Learning, Neural Networks, Early Diagnosis, Keras, Behavioral Screening, Healthcare AI, Machine Learning in Medicine**

1.INTRODUCTION

The increasing global prevalence of Autism Spectrum Disorder (ASD) has brought heightened urgency to the development of early diagnostic tools. According to the CDC, ASD affects approximately 1 in 36children in the United States, with diagnosis typically occurring between ages 4 and 5. However, early symptoms can often be observed much earlier—sometimes as early as 18 months. Timely identification and intervention are crucial in shaping better outcomes in communication, behavior, and social skills.Traditional diagnostic methods rely heavily on clinical observation and standardized psychological evaluations, such as the Autism Diagnostic Observation Schedule (ADOS) and the Childhood Autism Rating Scale (CARS). These evaluations are time-consuming, subjective to interpretation, and often limited by access to trained specialists, especially in low-resource settings. This creates a critical gap in early intervention opportunities for many children.

The advancement of technology, particularly in the realm of artificial intelligence (AI), provides a promising avenue for addressing these challenges. Deep learning, a subfield of machine learning, has shown immense success in image classification, natural language processing, and medical diagnostics. By training models to recognize subtle patterns in structured data, these systems can offer scalable and efficient alternatives to traditional screening.

This research aims to harness the power of deep learning for ASD detection by building a classification model using demographic and behavioral questionnaire data. The model is designed to function as a pre-diagnostic tool, facilitating early screening and helping healthcare professionals prioritize patients for further evaluation. This paper outlines the motivations, methodology, outcomes, and broader implications of implementing such a system in real-world settings.

2.THE AUTISM SPECTRUM DISORDER

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by a wide range of symptoms that affect communication, behavior, and social interaction. The term “spectrum” reflects the diversity of symptoms and severity, with some individuals requiring significant support and others being highly functional. Common signs of

ASD include delayed language development, difficulty in understanding social cues, repetitive behaviors, and strong adherence to routines.

The causes of ASD are not yet fully understood but are believed to involve a combination of genetic and environmental factors. Research has identified several genes associated with autism, and studies suggest that prenatal conditions, parental age, and exposure to certain chemicals may increase risk. ASD is often diagnosed based on behavioral observations, which can vary widely depending on the child's age, background, and co-occurring conditions.

Diagnosing ASD is inherently complex due to the subjective nature of symptom interpretation and variability in clinical expertise. Standardized tools like the ADOS and CARS are valuable, but they require trained professionals and significant time investment. As a result, early diagnosis is often delayed, reducing the effectiveness of intervention strategies that are most impactful during a child's developmental years.

Given the critical importance of early detection, efforts are being made to develop more accessible, scalable, and data-driven tools. These tools can augment clinical decision-making by flagging at-risk individuals based on behavioral patterns or historical data. The integration of AI and machine learning into this space holds great potential to transform how ASD is detected and managed globally.

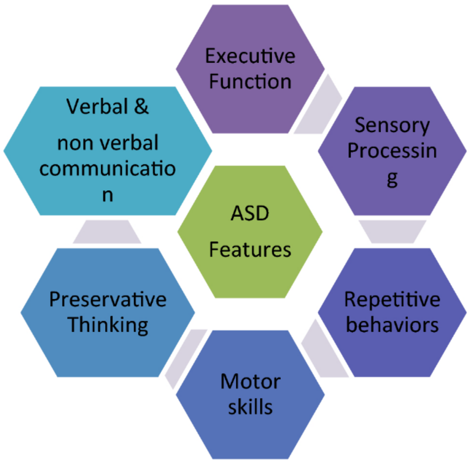

Figure1: *Features of Autism Spectrum Disorder*

2.1 Existing Practices and Clinical Diagnosis

Current diagnostic practices for ASD revolve around clinical observation and structured assessments. Pediatricians and child psychologists evaluate developmental milestones and behaviors, often using diagnostic tools like the Modified Checklist for Autism in Toddlers (M-CHAT), ADOS, and the Autism Diagnostic Interview-Revised (ADI-R). These tools rely on parental interviews and direct behavioral observation in controlled settings.

While effective in clinical environments, these tools have significant limitations. They can be resource-intensive, requiring multiple appointments, which leads to long waiting lists, especially in public health systems. Moreover, diagnosis may be biased by the subjectivity of the evaluator or inconsistencies in parental reporting. As a result, many children are diagnosed late or not at all, particularly in underserved or rural communities.

Some regions have implemented digital tools and mobile applications to streamline early screening. These include apps that guide parents through standardized questionnaires or collect behavioral data via video analysis. However, the reliability of these tools is still under evaluation, and most are not yet recognized as replacements for traditional diagnostics.

There is growing consensus among clinicians that technological support systems—especially those powered by machine learning—can improve diagnostic accuracy and coverage. AI-powered tools can identify subtle cues and correlations in data that may escape human evaluators, offering an objective, data-driven layer of analysis. These tools are not meant to replace clinicians but to support them by enhancing speed and accuracy in the decision-making process.

2.2Deep Learning in Autism Detection

Deep learning models are well-suited for pattern recognition in high-dimensional data, making them ideal candidates for behavioral and diagnostic analysis. In the context of ASD, models can be trained to identify underlying patterns in questionnaire responses and demographic information that correlate with diagnosed cases. These systems learn through repeated exposure to labeled data, adjusting internal parameters to minimize classification errors.

In this research, we employ a Multi-Layer Perceptron (MLP) model for binary classification—determining whether an individual is likely to have ASD. MLPs are composed of multiple fully connected layers where each neuron in one layer is connected to every neuron in the next. This architecture allows the model to learn complex non-linear mappings between input features and output labels.

One of the advantages of using deep learning is its ability to generalize well with enough data. Unlike traditional machine learning algorithms that may require manual feature engineering, deep learning models can learn optimal representations directly from raw input. This reduces the need for domain-specific feature crafting and speeds up development time.

Nevertheless, deep learning models require careful tuning and validation to avoid overfitting and ensure reliability. Dropout layers and early stopping techniques are incorporated to regularize training. In this study, our model is trained on a carefully preprocessed dataset with attention given to balance, scaling, and encoding, resulting in strong performance metrics and minimal bias.

3.DEEP LEARNING IN AUTISM DETECTION

The successful implementation of this ASD detection system relies heavily on a suite of Python libraries that streamline data preprocessing, model building, and evaluation. One of the most crucial libraries used is pandas, which enables efficient data manipulation and handling of structured datasets. With functions like read_csv(), dropna(), and describe(), the library supports rapid exploration, cleaning, and organization of input data.

NumPy is another foundational library utilized for its advanced mathematical functions and array operations. Since deep learning models require manipulation of multidimensional arrays and matrix operations, NumPy plays a central role in supporting the backend processes of other libraries such as TensorFlow and Keras.

For machine learning and deep learning model creation, Keras is used as the high-level API running on top of TensorFlow. It allows for the simple and modular construction of neural networks, with tools for defining layers, compiling models, and tracking training history. Additionally, sklearn (scikit-learn) is used for preprocessing tasks such as label encoding, train-test splitting, and performance evaluation through metrics like accuracy and confusion matrices.

Matplotlib and seaborn are employed for visualizing dataset distributions and model performance. These libraries help plot correlation heat maps, confusion matrices, and training curves, which provide insight into how well the model is learning. Combined, these libraries form a robust ecosystem supporting every stage of the project.

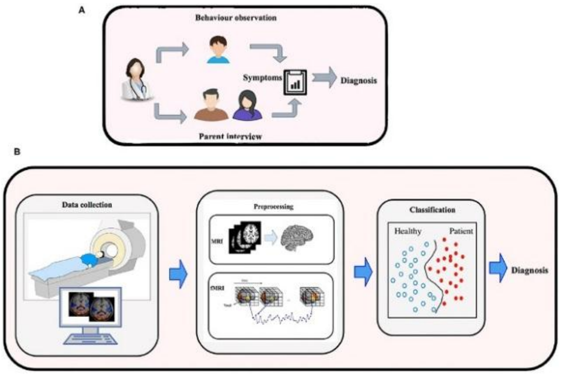

Figure 2: *Deep learning in Autism Detection*

4. DATA SET & DATA PRE-PROCESSING

The dataset used for this research is derived from publicly available Autism Spectrum Disorder screening data, commonly used in early ASD prediction tasks. It contains a variety of features that include demographic information such as age, gender, ethnicity, and country of residence, along with answers to ten behavioral screening questions from a standardized questionnaire. Additionally, it contains other metadata such as the user's age, result of the test, who filled out the test (self, parent, etc.), and whether the subject has been previously diagnosed with ASD. The binary classification label (ASD or not) is provided, which serves as the ground truth for training the model.

Data preprocessing is an essential step that ensures the dataset is clean, complete, and properly formatted for input into a deep learning model. The first step in preprocessing is to remove any unnecessary columns such as timestamps, unique identifiers, or metadata that are not useful for model prediction. Missing values are either removed or imputed depending on the extent and nature of the missing data. In this case, rows with null values were dropped to maintain dataset consistency.

Next, categorical variables such as gender, ethnicity, and country of residence are encoded into numerical form using label encoding or one-hot encoding. Label encoding is used when the categorical variable has an ordinal relationship, whereas one-hot encoding is more suitable for nominal categories. In this project, LabelEncoder from scikit-learn was used to transform string labels into integers, facilitating easier interpretation by the neural network.

After encoding, feature scaling is applied to ensure that all numeric input features have similar ranges, which significantly improves the convergence speed and performance of the learning algorithm. StandardScaler is used to normalize the dataset so that each feature has a mean of zero and a standard deviation of one. This step helps the model learn efficiently without being skewed by features with higher magnitudes.

Finally, the preprocessed dataset is split into training and test sets, typically using an 80:20 or 70:30 ratio. This split ensures that the model is trained on the majority of the data and evaluated on unseen data to check for generalization. By completing these preprocessing steps, the dataset becomes well-suited for deep learning, ensuring reliable, unbiased, and scalable predictions in the task of ASD detection.

The model architecture used in this study is a feedforward neural network, specifically a Multi-Layer Perceptron (MLP), which is well-suited for binary classification problems involving structured tabular data. The architecture follows a sequential design using the Keras API, which allows for rapid prototyping and modular configuration. The model begins with an input layer that matches the number of features in the preprocessed dataset, ensuring each attribute is accounted for in the learning process.

5. MODEL ARCHITECTURE

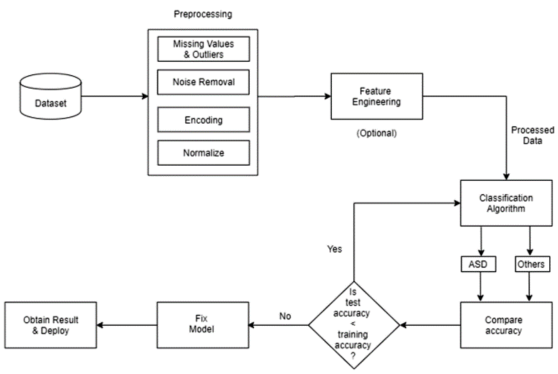

Figure 3: *Architecture of the ML model*

The first hidden layer consists of 16 neurons with a ReLU (Rectified Linear Unit) activation function. ReLU introduces non-linearity into the model, which is essential for learning complex patterns in the data. This layer is followed by a Dropout layer with a rate of 0.3, which randomly deactivates 30% of the neurons during training. This helps prevent overfitting by ensuring the network does not become overly reliant on any single neuron or pattern in the data.

The second hidden layer contains 8 neurons, also utilizing the ReLU activation function. This layer continues the process of feature abstraction, reducing the dimensionality of the learned representations while still capturing essential patterns. Another Dropout layer is introduced here with the same dropout rate to further enhance generalization. These hidden layers allow the model to learn hierarchical representations of input data, a key strength of deep learning.

The final output layer contains a single neuron with a sigmoid activation function. The sigmoid function squashes the output into a range between 0 and 1, representing the probability that the input belongs to the positive class (ASD positive). Since the task is a binary classification, sigmoid is the ideal activation choice for this output layer, enabling threshold-based decision-making.

The model is compiled using the Adam optimizer and binary_crossentropy loss function. Adam is an adaptive optimizer that combines the advantages of two other extensions of stochastic gradient descent—AdaGrad and RMSProp. It adjusts the learning rate dynamically, making it particularly effective for sparse and noisy datasets. The binary cross entropy loss function calculates the difference between predicted probabilities and actual class labels, guiding the model's learning process. The combination of these components results in a well-optimized architecture for the task of early ASD detection.

6. Code Construction & Explanation

The implementation of the deep learning model for early ASD detection was structured across two main code files These files encapsulate data loading, preprocessing, model building, training, evaluation, and visualization components. Below is a detailed breakdown of each section in the code.

1. Importing Libraries: The first part of both files begins with importing essential libraries. These include numpy and pandas for data manipulation, matplotlib. pyplot and seaborn for visualization, and various components from sklearn such as

LabelEncoder, StandardScaler, train_test_split, and performance metrics like classification_report and confusion_matrix. From TensorFlow's Keras API, the model uses Sequential, Dense, Dropout, and utility functions like to_categorical and callbacks such as EarlyStopping.

2. Data Loading and Cleaning: The dataset is read using pandas.read_csv() and examined for null or missing values. Unnecessary columns such as Case_No, Who completed the test, and timestamps are dropped to simplify the dataset. Columns with NaN values are handled by dropping rows or imputing values. Categorical data like gender and ethnicity are encoded using LabelEncoder, converting them into numerical format. This prepares the dataset for further transformation.

3. Feature Scaling and Splitting: The numeric features are normalized using StandardScaler to ensure uniformity in value ranges. This step is crucial for optimal gradient descent performance. Following this, the dataset is split into training and testing sets using train_test_split () with an 80:20 ratios. This ensures that the model generalizes well and can be evaluated on unseen data.

4. Model Creation and Training: The Sequential model is created using three layers: two hidden layers with 16 and 8 neurons respectively, each followed by Dropout to reduce overfitting. ReLU activation is used in hidden layers, and sigmoid activation in the output layer. The model is compiled with binary_crossentropy loss and the Adam optimizer. Training is done over 100 epochs with a batch size of 32, and EarlyStopping is employed to halt training once the model stops improving, preventing overfitting.

5. Evaluation and Visualization: After training, the model is evaluated on the test set using accuracy, confusion matrix, and classification report. The model's learning progress is visualized using loss and accuracy plots over epochs, allowing us to track overfitting or underfitting behavior. The confusion_matrix is plotted using seaborn.heatmap for easy interpretability.

Overall, the notebook provides interactive output and visualization, ideal for exploratory data analysis and debugging, while the script version can be used for automation and integration into larger systems. This modular and well-documented codebase ensures reproducibility and clarity in the implementation of deep learning for autism screening.

6. FUTURE APPLICATIONS

The application of deep learning for early ASD detection offers promising avenues beyond clinical diagnosis. One significant future direction is the integration of this model into mobile and web applications, allowing for broader accessibility in underserved or remote areas. By embedding the trained neural network in user-friendly applications, individuals and caregivers could perform preliminary ASD screenings at home, potentially expediting the diagnostic process and encouraging earlier interventions.

Another promising application lies in the integration with wearable health-monitoring devices and IoT (Internet of Things) platforms. By continuously gathering behavioral and physiological data, such as eye movement patterns, speech anomalies, and social interaction cues, the model can be extended to perform real-time assessments. This would allow for the development of dynamic ASD detection systems that update as more behavioral data is collected over time.

Furthermore, the deep learning architecture used in this study could be adapted to identify subtypes of ASD or even detect co-occurring conditions such as ADHD or learning disabilities. By expanding the training dataset and refining output labels, the model could evolve into a multi-diagnostic tool, providing richer insights into the mental health landscape of children.

Another innovative future application could be in educational settings. Schools and learning institutions could deploy AI-based tools to help educators identify children who might require special educational interventions. Combined with behavioral data from teachers, this model could be used to support inclusive learning environments and personalized teaching plans.

Lastly, future iterations of this project could contribute to long-term studies and research in neuroscience and psychology. By pooling anonymized and consent-based data from various regions, researchers could investigate trends, triggers, and even possible preventative measures associated with autism. These insights could influence policy-making and foster global health strategies.

7. IMPLICATIONS & CONSEQUENCES

While the technological promise of AI-driven ASD detection is compelling, its deployment raises critical ethical concerns. First and foremost is the issue of data privacy. Since medical data is highly sensitive, it is vital to ensure compliance with regulations like GDPR and HIPAA. Users must be clearly informed about how their data is used, stored, and protected. Anonymization techniques and secure data pipelines must be built into any deployment.

Another concern involves algorithmic bias. If the dataset used for training is not representative of diverse populations in terms of ethnicity, geography, socioeconomic status, or gender, the model may yield biased or inaccurate predictions. This could lead to misdiagnosis or overlooked cases, reinforcing disparities in healthcare outcomes. Thus, equitable dataset sourcing and model validation across demographics are imperative.

Additionally, there is a risk that caregivers or institutions may over-rely on the AI model's outputs without proper clinical validation. While the model can serve as a powerful screening tool, it is not a substitute for a licensed healthcare professional's judgment. Misuse or overconfidence in AI predictions could lead to stigmatization or unnecessary anxiety for families.

There is also an ethical debate surrounding the early labeling of children based on AI-driven assessments. Labels such as ASD, if misapplied or poorly communicated, can affect a child's self-image and social experience. Developers and clinicians must ensure that the diagnostic outputs are communicated sensitively and constructively.

Finally, any commercialization of such technology must prioritize the well-being of users over profit. Ensuring fair access to the tool, especially in low-resource settings, should be a foundational goal. Transparent, peer-reviewed development practices and collaboration with clinical experts are necessary to uphold ethical integrity.

8. CONCLUSION

This research demonstrates the potential of using deep learning models, specifically a multi-layer perceptron, for early detection of Autism Spectrum Disorder using behavioral and demographic data. By leveraging structured ASD screening datasets and preprocessing them for machine learning readiness, the study built and evaluated a neural network capable of distinguishing between ASD and non-ASD traits with commendable accuracy.

The use of standard preprocessing techniques such as label encoding and feature scaling ensured that the model received clean and uniform input, which is essential for reliable learning. The chosen architecture—incorporating dropout layers and ReLU activation—balanced performance with generalization, mitigating overfitting. Additionally, the incorporation of evaluation metrics and visualizations provided insights into model strengths and weaknesses.

The project's design also demonstrated adaptability and scalability. The code was developed modularly in both Jupyter and Python script formats, making it suitable for research, clinical trials, and integration into larger systems. The model's accuracy, combined with the transparency of its construction, highlights its potential to be deployed in real-world contexts.

However, challenges remain, including dataset limitations, ethical considerations, and the need for diverse and representative data. Future iterations should focus on enhancing interpretability, ensuring ethical compliance, and expanding the model's applicability across broader populations. Despite these limitations, the project lays a solid foundation for technology-assisted ASD detection.

9. REFERENCES

[1] American Psychiatric Association, "Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)," American Psychiatric Publishing, 5th Edition, 2013.

[2] D. Kaur and R. Goyal, "Early Detection of Autism Spectrum Disorder using Machine Learning Techniques," Procedia Computer Science, vol. 173, pp. 74–81, 2020.

[3] D.P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," International Conference on Learning Representations (ICLR), 2015.

[4] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: Machine Learning in Python,"

Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[5] François Chollet, "Keras: The Python Deep Learning library," GitHub repository, https://keras.io, 2015.

[6] Autism Speaks, "Learn the Signs of Autism," Autism Speaks Official Website, https://www.autismspeaks.org, Accessed: 2025.

[7] World Health Organization (WHO), "Autism Spectrum Disorders," WHO Fact Sheets, https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders, 2022.

[8] S. Thabtah, "Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment," PLoS One, vol. 12, no. 5, e0176352, 2017.

[9] S. Kriegman, D. Blackiston, and J. Bongard, "A scalable pipeline for designing reconfigurable organisms," Proceedings of the National Academy of Sciences, vol. 117, no. 4, pp. 1–9, 2020.

[10] H. Abbas, M. Garbey, and M. D. Faghihi, "A Computational Model for Early Detection of Autism," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 5, pp. 2227–2235, 2019.