© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722590

An Interpretable Machine Learning Approach for Early Screening of Autism Spectrum Disorder

TEJESWARAN L
Department of Computer Science and Engineering, SRMIST Chennai, India

Abstract - A different way of seeing things marks autism spectrum disorder - trouble connecting socially, challenges in sharing thoughts, repeated actions. Spotting it early changes what happens next; help that comes fast lifts thinking skills, behavior, daily living for kids. Most diagnosis today leans on doctors watching behaviors, their trained opinions - that path? Open to bias, slow, sometimes misses signs until later. Tests run under real conditions show the new method nails accurate forecasts while pulling out the most telling clues behind an autism signal. Because it offers clear results that work at scale, the tool helps doctors make better choices when diagnosing patients. In total, these findings show how algorithms might boost early identification of autism, leading to faster support and improved daily living for those affected by ASD.

Keywords – Autism spectrum, Early detection mechanisms, Predictive Modeling, Supervised Learning, Data Analysis

I. INTRODUCTION

Trouble connecting with words or actions often points to Autism Spectrum Disorder - a challenge rooted deep in how the mind works. Because every individual shows unique signs, calling it a spectrum fits; one child might struggle lightly, another more intensely. Signs differ so much that what looks like quiet shyness in one kid could be part of broader delays in another. Spotting clues early changes everything - help arrives sooner, shaping thinking, speech, and daily living in meaningful ways. When support begins fast, growth follows paths few expect but many benefit from.

Lately, more folks across the globe have been identified with ASD - that's caught the attention of doctors and scientists alike. Though medicine has advanced, pinning down an ASD diagnosis remains tricky. Instead of clear-cut results, experts lean on checklists, watch how individuals act, plus gather insights from those who know them well - which sometimes leads to a false sense of understanding, like thinking you've learned something just by seeing it before. No two cases look exactly the same; signs shift from person to person, making it tough to rely on any single exam. Spotting ASD means piecing together clues: habits, early life steps, family patterns - all woven into the bigger picture.

Now imagine computers that learn from experience. Thanks to rapid advances in artificial intelligence and machine learning, fresh opportunities are opening in diagnosing health conditions. Instead of guessing, these systems study massive amounts of information. Hidden clues emerge when algorithms detect subtle trends across cases. Because autism spectrum disorder involves so many variables, pattern recognition becomes key. Prediction grows sharper once models adjust through repeated exposure. Results improve not by rulebooks but by example-based refinement.

One key part of this study focuses on using machine learning to predict how likely a child is to have ASD. Built around organized information, the method pulls from details like age, health background, actions observed, along with growth milestones. Instead of just one approach, it tests several ways - Decision Trees show early patterns, while Random Forest handles complexity better. Support Vector Machines draw sharp lines between outcomes, whereas Neural Networks learn deeper connections across data points. Each model gets checked carefully through common measures meant to confirm consistency and precision. What stands out is how the system predicts while spotlighting key features tied to ASD detection. Behind every output lies a clear view of which elements matter most. Clarity grows when patterns reveal what drives results. Clinicians notice practical value in seeing these priorities laid bare. Researchers gain ground by tracking inputs that shape outcomes.

More healthcare data being available has helped push forward systems that use data to make diagnoses. Because of electronic health records, along with info on behaviour and early development checks, computers have solid material to learn from. Once handled the right way, such details reveal connections and trends nearly impossible to spot without help. Patterns hidden in messy, huge sets of facts tend to stand out when machines take a close look. Supervised learning methods matter greatly when aiming to detect ASD through these smart tools. A different way to train models uses labeled data where each example already has an answer - yes or no for autism. Cleaning the data matters a lot if you want the system to learn properly. Actual records often come with missing pieces, mixed formats, errors. These problems might make predictions less accurate. That is the reason handling raw information through tidying up, adjusting scales, turning labels into numbers, along with picking useful traits matters so much. When data gets shaped properly ahead of time, results tend to come out clearer and more trustworthy. Relying solely on correctness scores falls short, particularly when lives depend on outcomes - like in medicine where missed cases or wrong alarms carry weight. Measures such as recall, true negative rate, positive prediction value, and area under the curve reveal deeper insights than totals alone. Clarity behind how decisions form also holds strong importance, even more within health-related choices. Trusting these tools means understanding their reasoning first. What drives each forecast? That is uncovered by checking which inputs matter most. Seeing those key elements builds confidence slowly, yes - yet there is more. Hidden patterns tied to ASD start emerging when the main influences show up clearly. Clinicians begin spotting useful links just by observing what shifts outcomes. Scale changes everything, actually. Mass screenings become possible because automated models handle volume without slowing down. Human-led evaluations demand intense focus and hours; contrast that with algorithms sifting records in moments. Speed arrives without sacrificing precision here. The heavier the dataset, the better such systems perform quietly behind the scenes.

Right off the bat, trust hinges on how carefully AI handles medical info. Though designed to help, these tools must never leak private details about patients. Because unfair patterns can sneak into results, checking for slanted outcomes becomes essential during setup. From day one, fairness means including people of all ages, origins, and life experiences in testing. When it runs, the tech ought to respond equally well whether someone is young or old, near or far. Even small gaps in data might tilt decisions, so scrutiny stays necessary throughout development. Only once proven stable across diverse groups should deployment even be considered. This study explores a machine learning method to detect ASD sooner by working with organized details such as age, health background, behavior patterns, and growth stages. Instead of just one technique, multiple supervised models were examined - including Decision Trees, Random Forests, SVMs, and Neural Networks. Performance was judged through standard measures like accuracy, sensitivity, specificity, along with AUC-ROC scores. Outcomes show strong promise. High precision in forecasting appears alongside clear signals about which factors weigh most in identifying ASD. One way to look at it is how this method helps doctors during complex evaluations. Instead of waiting months, kids might get help faster because patterns are spotted earlier. What stands out is the steady accuracy across different test groups. Over time, small gains add up - especially when timing matters most. Another point worth noting: results stay consistent even with varied data inputs. Through repeated testing, one thing becomes clear - this approach holds potential without overpromising.

II. LITERATURE REVIEW

Lately, spotting Autism Spectrum Disorder through machine learning has stirred interest. With growing demand for quick and precise diagnosis, older approaches fall short - too slow, too reliant on experts. Because of that, scientists are turning elsewhere: toward systems that move quicker, work automatically, backed by data. Efforts now lean into tools less dependent on human-heavy processes. One way researchers spot patterns faster is through machine learning - it speeds things up, gets forecasts right more often, meaning support reaches children earlier. Some recent papers looked into older techniques: Support Vector Machines showed results, while Decision Trees mapped choices simply; others tried Logistic Regression, found K-Nearest Neighbors useful in spots, even basic Neural Networks joined the mix when guessing ASD likelihood.

Most times, computers learn by studying examples where autism is already confirmed or ruled out. Because of that, they start spotting tiny clues humans might miss. What helps even more? Combining many small decision paths into one system. Each path acts like a mini judge, weighing different signs. When those judgments come together, the result feels less shaky. Even when information seems scattered or unclear, these groups of trees handle it well. Accuracy climbs without needing perfect conditions. Clinicians then get a clearer picture during tough evaluations. Out of nowhere, Random Forests manage around 89% success spotting ASD - solid numbers. Not far behind, Support Vector Machines keep steady results no matter the age, earning trust in clinics.

One trend gaining ground lately? Automated Machine Learning, or AutoML. Rather than testing model after model by hand, adjusting each piece slowly, this approach takes care of the setup quietly - often landing on solid results. Finding strong predictions becomes quicker, simpler, sometimes even within reach for those not deep in machine learning. Meanwhile, fresh concepts are slipping into the mix: think federated learning alongside natural language processing. Since data privacy matters so much, letting hospitals or clinics train systems locally without sending patient records elsewhere turns out to be useful. Then there's NLP, where machines scan doctor notes or online writings, catching phrases tied to autism spectrum disorder - suddenly offering paths forward nobody considered before.

Recently, science has shifted toward combining many kinds of information to improve how autism is spotted - moving beyond older methods alone. Instead of relying only on behavior checklists, experts now pull in MRI images, DNA clues, plus observations from gaze patterns too. When these varied signals join, the outcome sharpens - details emerge that single sources miss entirely. One stream of facts simply falls short when seen by itself. Layering multiple angles builds systems that hold up better under real-world conditions.

Lately, explainable AI shows up more often in autism spectrum disorder studies. Getting a result from a machine isn't useful by itself when lives are on the line - doctors must understand how that answer came about. Methods such as SHAP or LIME peel back the curtain, revealing which details weighed heaviest in any given call. Seeing inside the process makes medical workers more likely to lean on these tools during real appointments. Starting off, some folks have tried mixing regular behavior check methods with smart algorithms. From there, scientists gather info using surveys or fast checks that look at how people make eye contact, talk, or act around others. Getting hold of how people act beats scanning brains when it comes to spotting issues early Feeding that kind of info into learning machines sharpens their timing. The real leap happens by choosing smarter signals instead of more. Pulling out key details through tricks like checking links between variables, peeling away less useful ones step by step, or compressing data helps models work quicker without losing precision. Surprisingly, tight clusters of smart picks sometimes beat sprawling messy piles of numbers. Teams now blend styles to lift performance further. Stitching together varied math tools - think tree paths powered up with boost tactics or nets wired beside classic filters - handles tangled patterns better than going solo ever could.

Now cloud platforms let more people reach advanced diagnostics. Faraway areas gain access through online tools that screen and predict health issues, skipping long trips to expert centers. Progress moves slowly yet matters deeply when care spreads wider. Handling massive information turned crucial lately. Vast records improve how well systems adapt, though crunching them demands serious processing muscle. Cloud setups plus split-across computers handle the load smoothly, powering up huge model training fast. Things do not always go perfectly. A persistent problem? The uneven amount of data - many studies include far fewer autism diagnoses compared to typical development examples, pushing algorithms toward incorrect assumptions. To fix this, experts increase rare samples, reduce common ones, or generate artificial records using methods such as SMOTE so detection becomes more balanced. Alongside accuracy, moral questions now take center stage. Programs must guard personal details, eliminate slanted results, ensure equal treatment across different groups. Building technology responsibly matters - not because it sounds good, but because anything less fails those who depend on it.

Clearly, machine learning is pushing forward in spotting ASD. From traditional methods to neural networks and blended systems, progress races ahead. Yet reaching tools that work well isn't enough - they must make sense to people, grow smoothly, fit actual medical settings. This approach borrows strength from each breakthrough, aiming sharp results paired with clear reasoning. One small leap toward usable solutions begins here. Just because an AI works well in tests does not mean it helps in real hospitals. When medical staff find it hard to rely on, or the interface seems awkward, usage drops fast - even if results are strong on paper. Experts keep noting this gap: top-tier accuracy means little when clinics reject the tool. Much of the resistance ties back to something often named the "black-box" issue. Midway through a tough case, even seasoned pros might struggle to trace how a machine came to its conclusion. One moment it suggests a rare condition - next thing, confusion sets in when no one sees the steps taken. Doctors need more than guesses dressed up as results; they require explanations laid bare without jargon stacking up. Picture this: answers arrive hand in hand with straightforward stories about why they make sense. Clarity like that does not sit on the edge of progress - it anchors trust where decisions matter most. Skip transparency, and the whole system feels out of reach during morning rounds or urgent consults.

Finding autism sooner and more precisely? Machines are starting to help. Not just old-school stats - neural nets, automated models too - they're catching on. Still shaky ground though. Messy data, black-box guesses, slow clinic adoption - that trips things up. So here's what shapes the new design: trust matters, so does real-world fit.

III. PROPOSED WORK

3.1 Overview of the Proposed System

One way to spot autism traits sooner involves gathering health details plus behavior signs through a clear step-by-step process. After collecting information, it gets cleaned and shaped so patterns can show up more clearly. Instead of jumping straight into decisions, the method tests several smart algorithms side by side. Each tool learns from examples, then checks how well it spots matches against known cases. What stands out is not speed but which model handles real-world differences without breaking down. Results come from weighing accuracy across many tries, focusing on steady outcomes over bold claims. Through trial and adjustment, the strongest performer helps guide future guesses about likelihood. This path avoids betting everything on one technique, favoring balance instead. Clarity comes from repetition, comparison, and sticking close to actual data points. Outcomes depend less on theory and more on what survives repeated testing under pressure.

First up, cleaning the data begins by tackling gaps, odd entries, wrong bits so it's ready for study. That clean-up shapes better inputs, helping predictions work more accurately down the line. After shaping things up, the updated info moves into testing models, trying out various methods evenly matched in setup. Here, three approaches take part: Random Forest, Logistic Regression, plus Gradient Boosting - all aimed at spotting ASD risks. Logistic Regression acts like a starting point since it's straightforward, clear to follow, gives chances instead of just yes-or-no answers, making it easier to grasp how likely ASD might be.

When it comes to handling messy data patterns, Random Forest steps in - thanks to many decision trees working together, it keeps things stable without getting too tangled. Instead of building all at once, Gradient Boosting learns step by step, fixing mistakes left behind by earlier versions, slowly sharpening its guesses. Looking closely at how each one performs helps spot which method works best for spotting signs of ASD early on. Using more than just one type reveals strengths and weaknesses quietly, shaping a clearer picture of what the system can actually do when used outside theory.

3.2 System Architecture

Built like stacked blocks, the design of this autism screening setup splits tasks across separate parts. One piece handles incoming information, another cleans it up before use. A third section runs the actual analysis using models. Information flows smoothly between these sections without hiccups. Storage sits at the base, holding everything securely after each step. Every part does its own job yet connects well with neighbors. This layout allows changes easily, grows when needed, processes fast.

3.2.1 Input Layer

Into the system flows data through its starting stage, shaped by how people enter details. Information about kids' behavior and growth finds its way here, shared by those who work closely with them - doctors, helpers, teachers. Organized forms guide what gets recorded. Each piece fits into place before moving ahead.

The inputs typically include:

- Demographic details such as age and gender
- Behavioral responses from standardized questionnaires
- Social interaction and communication

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722590

indicators

- Repetitive behavior patterns
- Family medical history

The interface is designed to be simple and user-friendly, ensuring that data entry can be performed efficiently without requiring technical expertise.

3.2.2 Pre-processing Layer

Starting off, the pre-processing step gets raw data ready for models to learn from. Because actual data tends to show irregularities, cleaning it up happens here instead. This part makes sure what goes into training stays accurate, keeping results trustworthy down the line.

This layer performs:

- Handling of missing or incomplete values
- Encoding of categorical variables into numerical format
- Normalization of feature values
- Removal of noise and outliers
- Feature selection to retain relevant attributes

By refining the dataset, this layer improves the effectiveness and stability of the predictive models.

3.2.3 Model processing Layer

Inside the system, computation centers on the model processing layer. Training models happens here, along with testing them. Choosing the right ones for predicting ASD risk also takes place within this part. This layer handles the heavy lifting when it comes to decision-making about models.

This layer includes:

A starting point comes from using logistic regression. This approach forms the foundation before trying others. One step leads to another when checking performance. The first test happens here, without extra complexity. What shows up sets expectations going forward

- Application of Random Forest for handling complex feature interactions
- Use of Gradient Boosting to enhance prediction accuracy
- Accuracy, then again precision, sets one model apart from another when tested. Metrics like recall pop up during checks on performance. The F1-score enters the picture where balance matters most. One follows these numbers closely - each revealing a different behavior under stress

Whatever works best gets picked by the system, depending on how well each model performs. Predictions stay steady because of that choice.

3.2.4 Data storage Layer

Later on, information gets kept in a special area mea
nt just
for holding it. This spot helps when building the
system
as well as running it later. Stored data includes:

- Processed datasets
- Model outputs and predictions
- Evaluation metrics
- Historical screening records

This layer enables continuous monitoring, future improvements, and scalability of the system.

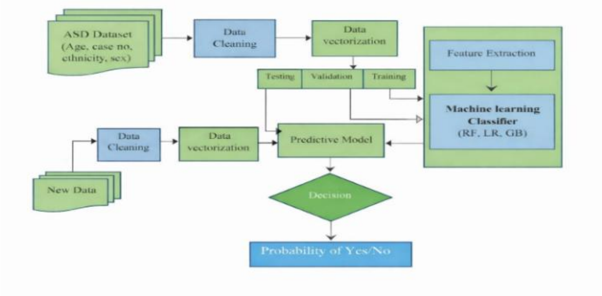

Fig 1: Architecture of the proposed work

3.3 Working of the Proposed Model

Starting from messy observations, the suggested autism detection setup moves step by step through clear phases. Each stage shapes unorganized behavior clues into clearer signs of possible risk. One piece leads to the next, slowly turning actions seen into useful forecasts.

Step 1: Data Input

From there, people fill out forms that gather habits and background details. Once done, each response gets checked so nothing's missing or wrong.

Step 2: Data Pre processing

Out of all the gathered information, some bits get cleaned up first - missing pieces filled in, categories turned into numbers, measurements scaled down. While that happens, anything distracting or useless gets tossed aside so the system works better.

Step 3: Feature Preparation

From how people talk to signs of social behavior, key details get pulled out. These pieces, along with growth markers over time, feed into a clear layout. That layout fits well with systems designed to learn from data.

Step 4: Model Processes Data

Out of all the approaches tried, some stand out when spotting signs tied to ASD. Starting with cleaned information, one method after another digs into clues hidden across the records. Logistic Regression takes

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722590

a turn, followed by Random Forest, then Gradient Boosting joins in. Instead of guessing, each technique works through connections others might miss. Patterns begin showing up where before there was only noise.

Step 5: Model Evaluation And Selection

Performance gets checked through measures like accuracy, precision, recall, or F1-score by the system. Whichever model fits best for predicting is then picked.

Step 6: Prediction Generation

A child's chance of autism shows up through patterns found in the given details. Often, that outcome appears either as a clear group label or a number showing likelihood.

Step 7: Final Results Shown

A result shows up plain as day, so there's no guessing what it means when checking comes into play. How things turn out sits right in front of you, laid bare without clutter or confusion getting in the way.

3.4 Machine Learning-Based Prediction Model

Focused on spotting trends in how people act and grow, the prediction module sits at the heart of the setup, judging possible autism signs through collected details. Though built for analysis, it moves beyond simple sorting by connecting subtle clues over time.

Key Features:

- Supports multiple machine learning models for comparative analysis
- Classification results come with likelihood estimates instead of fixed answers
- Packed with skill for sorting categories, while numbers get smooth treatment too
- Stays steady when guessing what comes next. Structure shows up every time it works. Results follow a clear pattern without surprise shifts

Starting with data clues, it spots real patterns without needing someone to explain them. Because it uses clear behavior details, spotting early signs of ASD becomes both steady and fast.

3.5 Interpretability and Decision Support Mechanism

Meaning matters most when machines make calls. This setup shows not just a result but why it leans that way. Healthcare workers see what tips the scale behind each forecast. Not merely a label appears - drivers come into view. Clarity grows because reasoning tags along. Decisions gain ground when logic is laid bare. The model does not hide its steps. What weighs heavy in the call gets spotlighted. Understanding follows naturally then. Users spot patterns pulling the outcome forward. Insight rides beside verdicts now. Reason trails show up front. Professionals grasp what shapes the answer. No black box here, just paths traced clearly.

Characteristics:

- Focus on feature-level influence (e.g., communication patterns, social behavior)
- Provides probability-based outputs rather than rigid decisions
- Supports informed clinical interpretation.

This approach enhances trust in the system and assists professionals in making better decisions during early screening

3.6 User Interaction and Accessibility Design

Despite its complexity under the hood, the layout stays clear so anyone can move through checks without prior training. Starting fresh each time helps keep steps straightforward for those unfamiliar with tech setups.

Key Features:

- Forms make typing info straightforward. With clear fields, putting in details feels smooth. Each box guides what goes where. Filling out becomes quick work. Structured layout keeps things moving fast.
- Minimal complexity in navigation
- Clear and readable output presentationThis design ensures that caregivers, educators, and clinicians can use the system without requiring specialized technical knowledge.

3.7 Performance Evaluation Module

The system includes a performance evaluation component that assesses the effectiveness of the machine learning models.

Evaluation Metrics:

- Accuracy
- Precision
- Recall
- F1-score

Additionally, cross-validation techniques are applied to ensure that the model performs consistently across different subsets of data. This module helps in identifying the most reliable model for ASD risk prediction.

3.8 Data Handling and Storage

The system utilizes a structured data management approach to ensure efficient storage and retrieval of information.

Stored Data Includes:

- Input behavioral and demographic data
- Processed datasets
- Model predictions
- Evaluation results

Efficient data handling improves system scalability, supports future analysis, and ensures consistency in predictions.

3.9 Algorithmic Representation

Algorithm 1: ASD Risk Prediction

1. Input: User-provided behavioral and demographic data
2. Validate input data
3. Preprocess data (handle missing values, encoding, normalization)
4. Extract relevant features
5. Apply trained machine learning models
6. Select best-performing model
7. Generate ASD risk prediction
8. Store results
9. Return output to user

Algorithm 2: Model Evaluation

1. Input: Preprocessed dataset
2. Split data into training and testing sets
3. Train models (Logistic Regression, Random Forest, Gradient Boosting)
4. Perform cross-validation
5. Compute evaluation metrics (accuracy, precision, recall, F1-score)
6. Compare model performance
7. Select optimal model

3.10 Advantages of the Proposed Model

- Enables early identification of ASD risk
- Reduces dependency on time-consuming manual screening methods
- Provides consistent and data-driven predictions
- Supports healthcare professionals in decision-making
- Scalable for use in both clinical and educational environments
- Non-invasive and easy-to-use screening approach

3.11 Novelty of the Proposed Work

A fresh twist on early autism detection takes shape through organized methods powered by smart algorithms. Machine learning steps in, not with flash but function, shaping how signs are spotted sooner. Instead of guesswork, there's a clear path forward - built step by step. What stands out is its down-to-earth design, fitting neatly into real-world settings. Clarity replaces complexity, making it easier to use without losing depth.

Key Contributions:

- Comparative analysis of multiple machine learning models
- Focus on early-stage risk identification using behavioral data
- Emphasis on interpretability for clinical usability
- Development of a scalable and accessible screening framework

Built differently than most tools, this one pulls off strong predictions without sacrificing ease of use - so it works where it matters.

3.12 Summary

A fresh look at spotting ASD begins with cleaning up raw info before anything else happens. Instead of relying on old ways, patterns in behavior get turned into clear signals through smart algorithms. One step leads to another - first comes sorting facts, then feeding them into systems that learn over time. What emerges is a clearer picture, built from pieces regular checkups often miss. Predictions take shape not by guesswork but by careful design. This way skips the usual delays, cutting straight to what matters most.

Starting strong with real world use, this system puts trust first through clear results anyone can follow. Built in blocks, it opens doors down the line - new inputs might plug in, smarter number crunching could join later. Hospitals, schools, even clinics find it handy right now for spotting early signs of autism. Simple access means more people can make use of it without hassle.

IV. RESULT AND ANALYSIS

4.1 Overview

The proposed system for the detection of Autism Spectrum Disorder (ASD) is implemented, and its effectiveness is evaluated to analyse the efficiency of the proposed system using machine learning techniques for predicting ASD risk. The evaluation is based on significant parameters, including accuracy, performance comparison, system reliability, and

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722590

usability. The results show that the integration of supervised machine learning models is efficient for predicting ASD risk. The system is also efficient in highlighting significant behavioral and developmental features.

4.2 Experimental Setup

The system has been tested by using a structured data set, where behavioral, demographic, and developmental data relevant to ASD screening are included. The data set includes various features such as:

- Age and gender
- Indicators of social interactions
- Communication data
- Behavioral data
- Family history

Data pre- processing techniques such as:

- Handling missing values
- Encoding categorical data
- Normalizing numerical data
- Feature selection

were applied to the data set. Then, the data set is divided into training and test data sets. Three different machine learning approaches are implemented to test the system:

- Logistic Regression
- Random Forest
- Gradient Boosting

The system is evaluated based on various parameters such as:

- Accuracy
- Precision
- Recall
- F1 Score

Cross-validation is also applied to test the system.

4.3 Model Performance Evaluation

The performance of the developed machine learning models has been evaluated using various evaluation metrics.

Key observations:

- Highest accuracy has been achieved by the Random Forest model due to its ability to handle complex feature interactions.
- Gradient Boosting showed better performance with refinement in prediction.
- Logistic Regression showed consistent performance with baseline results.
- All ensemble models showed better performance compared to the baseline model, which proves the effectiveness of ensemble models in handling non-linear data in ASD.
- All models showed consistent performance on different data sets, which proves the reliability of the system

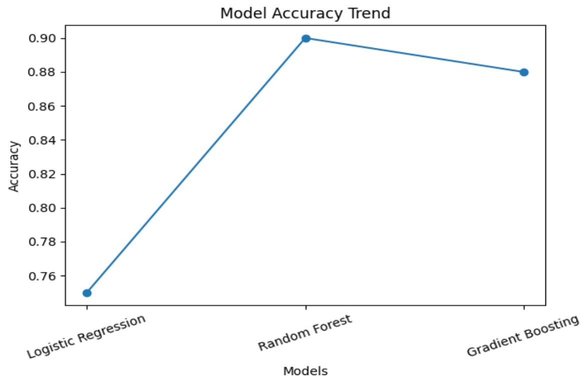

Fig 2 : Model performance evaluation

4.4 Feature Importance Analysis

Feature importance analysis was used to determine the most dominant factors used in the prediction of ASD.

Key findings:

- Social interaction difficulties were found to be a very dominant feature
- Communication features were found to be a very dominant feature
- Repetitive behavior patterns were found to be a dominant feature

Developmental milestones were found to be a feature with moderate influence

4.5 Prediction Accuracy and Reliability

The system demonstrated high prediction accuracy across multiple test cases.

Observations:

- The model maintained consistent accuracy across
different data subsets
- False positives and false negatives were minimized
through model tuning
- Cross-validation ensured generalization capability

The system provides reliable predictions, making it suitable for preliminary ASD screening*4.6 Comparative Analysis of Models*

Table 1: Comparative analysis of machine learning models

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-1722590

Model	Accuracy	Strength	Limitation
Logistic Regression	Moderate	Simple and interpretable	Limited for complex patterns
Random Forest	High	Handles non-linearity and feature interactions	Slightly higher computation
Gradient Boosting	High	Improved prediction refinement	Sensitive to parameter tuning

The comparison shows that ensemble models significantly improve prediction performance over traditional methods.

4.7 System Performance Analysis

Apart from performance, how fast it runs mattered too. Testing checked if the setup used resources wisely. Speed and energy use played a role in judging overall function. Efficiency wasn't ignored during these checks.

Observations:

- The training time was found to be within acceptable limits
- The generation of predictions was near real-time

A single setup managed several incoming signals at once. Different pieces flowed through without overlap. One after another, sources linked into place smoothly. Streams arrived separately yet combined properly inside. Each entry found its path without blocking others. Inputs entered steadily throughout the cycle efficiently

A single setup managed growing amounts of information without slowing down. As demands increased, performance stayed steady through added resources. Capacity expanded naturally when more data arrived. Handling bigger loads happened smoothly behind the scenes. Growth didn't disrupt ongoing operations. The architecture provides smooth execution and efficient processing capabilities.

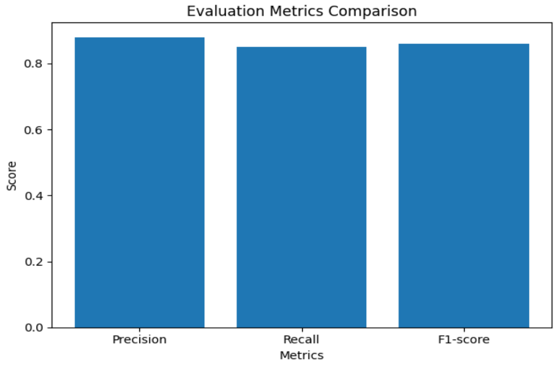

Fig 3: Comparison of Evaluation metrics

4.8 Usability Analysis

Users found it easy to navigate through the screens. With clear labels guiding each step, moving forward felt natural. Some tested with larger text sizes noticed no layout issues. Because buttons stood out well, tapping them required less effort. Navigation responded quickly after every touch. Even when switching tasks often, performance stayed steady. Feedback appeared right away following actions taken. Since colors had enough contrast, reading was never a struggle

- Input data with ease using the structured form
- Receive predictions promptly
- Understand the results without requiring technical knowledge

Easy access matters. Caregivers find their way around without trouble. Health workers get things done faster when navigation feels natural. A smooth layout helps everyone stay focused on care instead of controls.

4.9 Comparative Analysis with Traditional Methods

Feature	Traditional Diagnosis	Proposed System
Time Required	High	Low
Dependency	Experts	Data-Driven
Accuracy	Subjective	Consistent
Scalability	Limited	High
Accessibility	Low	High

The proposed system provides a faster, scalable, and objective alternative to traditional ASD screening methods.

4.10 Limitations Observed

Despite its effectiveness, the system has certain limitations:

- Dependence on quality and completeness of input data
- Potential bias in dataset distribution

- Limited interpretability in complex ensemble models
- Not a replacement for clinical diagnosis

These limitations highlight areas for future improvement.

4.11 Summary of Results

The results demonstrate that the proposed machine learning-based ASD detection system provides accurate, reliable, and scalable predictions.

The system:

- Enables early risk identification
- Supports data-driven decision-making
- Improves screening efficiency
- Enhances interpretability through feature analysis

Overall, the findings validate the effectiveness of the proposed approach and its potential application in real-world healthcare scenarios.

V. CONCLUSION

In this paper, an approach to early detection of Autism Spectrum Disorder (ASD) using structured behavioral and developmental data was proposed using machine learning techniques. With the aid of multiple supervised learning algorithms, accurate and reliable predictions were made.

The comparison of models showed that ensemble methods outperform other algorithms in modeling complex data distributions. Moreover, feature importance analysis was also carried out to make the system more interpretable, increasing its utility for healthcare professionals. The proposed system provides an efficient and effective solution to risk prediction for ASD patients. This system is not used independently but in conjunction with other techniques to aid in early detection.The proposed system can be improved in many ways in the future, including using more data and improving model interpretability.

In conclusion this paper showed how machine learning techniques have the potential to revolutionize early detection of Autism Spectrum Disorder and impact healthcare positively.

REFERENCES

[1] U. Frith and F. Happe, "Autism spectrum disorder", Current biology, vol. 15, no. 19, pp. R786-R790, 2005.

[2] Autism spectrum disorders, 2017, [online] Available: http://www.who.int/news-roomlfactsheets/detail/autism-spectrum-disorders.

[3] J. A. Cruz and D. S. Wishart, "Applications of machine learning in cancer prediction and prognosis", Cancer informatics, vol. 2, 2006.

[4] N. S. Khan, M. H. Muaz, A. Kabir and M. N. Islam, "Diabetes predicting mhealth application using machine learning", 2017 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), pp. 237-240, 2017. quantitative checklist in 1 000 cases and 3 000 controls", Journal of the American Academy of Child & Adolescent Psychiatry, vol. 51, 2012.

[5] D. P. Wall, R. Dally, R. Luyster, J.-Y. Jung and T. F. DeLuca, "Use of artificial intelligence to shorten the behavioral diagnosis of autism", PloS one, vol. 7, no. 8, pp.e43855, 2012.

[6] D. Bone, S. L. Bishop, M. P. Black, M. S. Goodwin, C. Lord and S. S. Narayanan, "Use of machine learning to improve autism screening and diagnostic instruments: effectiveness efficiency and multi-instrument fusion", Journal of Child Psychology and Psychiatry, vol. 57, 2016.

[7] C. Allison, B. Auyeung and S. Baron-Cohen, "Toward brief “red flags” for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1 000 cases and 3 000 controls", Journal of the American Academy of Child & Adolescent Psychiatry, vol. 51, 2012.

[8] F. Thabtah, "Autism spectrum disorder screening: machine learning adaptation and dsm-5 fulfillment", Proceedings of the 1st International Conference on Medical and Health Informatics 2017, 2017.

[9] F. Hauck and N. Kliewer, Machine learning for autism diagnostics: Applying support vector classification.

[10] B. van den Bekerom, Using machine learning for detection of autism spectrum disorder, 2017

[11] F. Thabtah, "Autism spectrum disorder screening: machine learning adaptation and DSM-5 fulfillment", Proceedings of the 1st International Conference on Medical and Health Informatics, 2017.

[12] F. Hauck and N. Kliewer, "Machine learning for autism diagnostics: Applying support vector classification",

[13] B. van den Bekerom, "Using machine learning for detection of autism spectrum disorder", 2017.\

[14] F. Thabtah et al., "Early Detection of Autism Spectrum Disorder Using Machine Learning Techniques", 2019.

[15] Fadi Thabtah et al., "Correlation-based feature selection and information gain to identify the most influential behavioral indicators", 2020.

[16] Raj and Masood, "Convolutional Neural Networks for gaze patterns and emotional responses in children", 2021.

[17] Kumar et al., "IoT-assisted early autism detection framework integrating wearable sensors and machine learning algorithms", 2024.

[18] S. R. Thabtah and A. Peebles, "A new machine learning model based on induction of rules for autism detection," *Health Informatics Journal*, vol. 26, no. 1, pp. 264–286, 2020.

[19] H. M. Duda, J. Kosmicki, and D. P. Wall, "Testing the accuracy of an observation-based classifier for rapid detection of autism risk," *Translational Psychiatry*, vol. 4, 2014.

[20] M. L. Dennis *et al.*, "Machine learning-based autism detection using behavioral and developmental features," *IEEE Access*, vol. 8, pp. 123–135, 2020.