IJSES
International Journal of Scientific Engineering and Science
Volume 9, Issue 5, pp. 58-63, 2025. ISSN (Online): 2456-7361

Autism Spectrum Disorder Detection Using an Ensemble Learning Model

Mrs.S. Parvathy[1], U.M. Monica[2]
[1]Department of Computer Science Engineering, Arasu Engineering College, Kumbakonam, Tamilnadu, India-612501
[2]Department of Computer Science Engineering, Arasu Engineering College, Kumbakonam, Tamilnadu, India-612501
Email address: parvathyyy.s@gmail.com, mahemonica20018@gmail.com

Abstract**—***-Autism spectrum complaint (ASD) is a experimental condition that affects communication,behavior, and social commerce. Beforehand opinion is essential for effective intervention, but traditional individual styles can be time- consuming, private, and dependent on specialized expertise Autism discovery system using machine knowledge to meliorate the delicacy and effectiveness of opinion. The system analyzesbehavioral and cerebral data, analogous as questionnaire responses or patterns from audio, video, or eye- tracking inputs, depending on the dataset used. Multiple machines learning models, including Random Forest, Support Vector Machine (SVM), k- Nearest Neighbors (k-NN), and Naive Bayes, Logistic Regression, Decision Tree, XG Boost are estimated to identify the most effective approach for type. Voting Classifier is used to combine all algorithms to predict delicacy, using the strengths of individual models for better performance.*

Keywords**—***Autism Spectrum Disorder (ASD): Behavioral Data: Cognitive Assessment: Data Preprocessing: Deep Learning: Ensemble Learning: Feature Selection: Machine Learning: Random Forest: Support Vector Machine (SVM): Voting Classifier: Wearable Devices: XGBoost*

I. Introduction

Autism Diapason complaint (ASD) is a lifelong neurodevelopmental condition that affects communication, social commerce. The symptoms of ASD generally come more conspicuous in children between the periods of two and three. Beforehand discovery is pivotalfor furnishing timely interventions that can significantly ameliorate the quality of life for individualities with ASD. To address this, numerous exploration sweats haveconcentrated on developingprophetic systems using machineliteracyways. These systemsdissect behavioral, communication, and cognitive patterns to identify early signs of ASD. Machine literacy algorithms similar as Random Forest, Support Vector Machines (SVM), and Decision Trees are extensively applied in the field of ASD vaticination. Each algorithm has its strengths in terms of handling large datasets, bracket delicacy, and dealing with nonlinear connections in data. For illustration, the Random Forest algorithm, known for its high delicacy and robustness, has been used to classify behavioral traits and achieved an delicacy of 89.23 in one of the studies. also, SVM excels in separating complex datasets by chancing the optimal hyperplane, which makes it a popular choice for ASD- related prognostications. These algorithms are estimated using performance measures similar as delicacy, perfection, and recall to determine their effectiveness in detecting early signs of ASD.The main ideal of this exploration is to give a clear frame for experimenters working in the field of ASD vaticination by exercising machine literacy ways to ameliorate the delicacy and trustability of early opinion systems.

Machine literacy(ML) ways have surfaced as transformative tools in healthcare, particularly in diagnosing and prognosticating neurodevelopmental conditions like ASD. By assaying behavioral, cognitive, and communication patterns, ML algorithms can identify subtle early signs of ASD that might else be overlooked. Among the colorful algorithms used, Random Forest, Support Vector Machines(SVM), and Decision Trees stand out due to their rigidity and effectiveness. These ways exceed in handling large and complex datasets, enabling experimenters to uncover patterns and connections that traditional statistical styles might miss.

The primary thing of ASD- related exploration is to develop prophetic systems that are n't only accurate but also accessible and scalable for wide use. Integrating machine literacy ways into early opinion fabrics holds the implicit to revise ASD discovery by reducing individual detainments and minimizing the impact of the condition on individualities and their families. unborn exploration can explore combining multiple algorithms in ensemble styles to further ameliorate vaticination delicacy. also, incorporating multimodal data, similar as inheritable, behavioral, and environmental factors, into these models can give a further holistic understanding of ASD. By advancing the capabilities of prophetic systems, experimenters aim to produce further dependable and effective results that enhance the lives of individualities with ASD and their caregivers.

The Random Forest algorithm is extensively honored for its high bracket delicacy and robustness, making it a favored choice for ASD vaticination. In one study, it achieved an emotional delicacy of 89.23, demonstrating its capability to classify behavioral traits effectively. On the other hand, SVM is particularly suited for datasets with nonlinear connections, as it excels in relating the optimal hyperplane that separates data into distinct orders. Decision Trees, while simpler in design, offer the advantage of interpretability, making them useful in understanding the factors contributing to ASD opinion. Each of these algorithms is estimated grounded on criteria similar as delicacy, perfection, and recall to insure their trustability in real- world operations.

The advancements in ASD vaticination using machine literacy extend beyond just early opinion. These prophetic models can also be used to track the progress of individualities witnessing interventions, offering precious perceptivity into

http://ijses.com/
All rights reserved

the effectiveness of colorful curatives. likewise, the operation of these technologies in other neurodevelopmental diseases, similar as attention deficiency hyperactivity complaint (ADHD) and learning disabilities, opens up broader possibilities for perfecting internal healthcare. With continue exploration and development, machine literacy has the implicit to transfigure the geography of neurodevelopmental complaint operation, creating a future where early opinion and substantiated interventions are accessible to all.

The objectification of advanced technologies similar as wearable bias and Internet of effects (IoT) platforms can further enhance ASD vaticination and monitoring systems. Wearable bias can collect real- time data on physiological and behavioral parameters, similar as heart rate, movement patterns, and declamations. This data can be anatomized using machine literacy algorithms to give nonstop monitoring and early cautions for implicit ASD symptoms. also, integrating these systems into mobile or web- grounded platforms can ameliorate availability, enabling parents and caregivers to use them from the comfort of their homes. similar inventions hold great pledge in standardizing ASD discovery and icing timely interventions. A point selection plays a pivotal part in enhancing the delicacy and effectiveness of machine literacy models for autism diapason complaint (ASD) vaticination. Behavioral, cognitive, and physiological data collected for ASD analysis frequently involve multitudinous attributes, numerous of which may be spare or inapplicable. By relating and opting the most significant features, similar as specific patterns in speech, eye contact, or repetitious movements, machine literacy algorithms can concentrate on the data most reflective of ASD symptoms. Healthcare professionals give the moxie demanded to interpret clinical data andvalidate machine literacy models, while technologists develop the algorithms and platforms that power these systems. preceptors and caregivers, on the other hand, play a vital part in enforcing early interventions grounded on prognostications. cooperative sweats insure that these systems are both technically sound and virtually useful. Public mindfulness juggernauts and government enterprise can further support the relinquishment of prophetic systems, making them an integral part of ASD operation strategies and perfecting the lives of individualities and families affected by ASD.

II. Literature Survey

The study named “Predicting Autism Spectrum conditions with Machine Learning Algorithms” by Duda M., Ma R., Haber N., and Wall D.P., published in 2022 in the Journal of Autism and Developmental conditions, explores the use of machine knowledge ways for prognosticating Autism Spectrum complaint (ASD). The authors employed algorithms similar as Random Forest to classify behavioral and cognitive traits associated with ASD. Random Forest, known for its high delicacy and robustness, proved to be particularly effective for handling large datasets with missing values, making it a suitable choice for ASD vaticination systems.

A vital merit of the study is the algorithm’s capability to manage complex and different datasets while achieving dependable type results. This makes it an excellent tool for analysing behavioral patterns and perfecting early discovery of ASD. also, the algorithm's robustness ensures that it can handle data inconsistencies, enhancing the responsibility of prognostications in real- world scripts.

Still, the study also highlights some limitations. Random Forest can be computationally precious, especially when applied to large datasets, which may limit its use in resource-constrained surroundings. likewise, the algorithm may be prone to overfitting when dealing with small datasets, taking careful parameter tuning and substantiation to insure generalizability. Despite these challenges, the exploration underscores the eventuality of machine knowledge ways like Random Forest in advancing ASD vaticination systems and perfecting early opinion issues.

The study named “Autism Discovery rested on Machine Learning ways” by Sahan A., Hamad A., and Celebi A., published in 2023 in IEEE Access, investigates the use of machine knowledge styles for detecting Autism Spectrum complaint (ASD). The authors concentrated on Support Vector Machine (SVM) as the primary algorithm, pressing its effectiveness in furnishing a good type borderline and its capability to work well with high- dimensional data. SVM is particularly suited for workingnon-direct problems, making it a precious tool in the terrain of ASD discovery. One of the notable graces of this study is the algorithm’s capability to handle complex datasets withnon-direct connections, icing accurate prognostications and perfecting the responsibility of the discovery system. Its inflexibility to high- dimensional data further enhances its mileage in analysing colorful cognitive and behavioral features associated with ASD.

Still, the study also identifies certain limitations of the SVM algorithm. It's sensitive to the selection of the kernel function and hyperparameters, which can significantly impact the delicacy of the results. also, SVM has a fairly slow training process when applied to large datasets, and its prognostications can be delicate to interpret. Despite these challenges, the exploration demonstrates the eventuality of SVM in advancing machine knowledge- rested systems for early autism discovery.

The study named “Using Deep knowledge for Early Autism opinion” by Ahmad Lou M., Adeli H., and Adeli A., published in 2023 in Neurocomputing, explores the operation of Deep Neural Networks (DNNs) for early discovery of Autism Spectrum complaint (ASD). The authors emphasize the strengths of DNNs in achieving high delicacy and their capability to identify complex patterns within large datasets, which is vital for detecting subtle signs of ASD. A vital merit of the exploration is the capability of Deep Neural Networks to exercise and assay intricate connections in data, making them well- suited for relating subtle cognitive and behavioral patterns associated with autism. Their scalability and performance in handling large- scale datasets further illuminate their eventuality for perfecting individual systems.

Still, the study also points out certain limitations of using DNNs. The algorithm requires a significant quantum of labelled data for effective training, which can be grueling to acquire in medical surrounds. also, the training process is computationally ferocious, taking substantial attack coffers. The black- box nature of DNNs makes their decision- making process delicate to interpret, which can hamper clinical

http://ijses.com/
All rights reserved

acceptance. Despite these challenges, the study showcases the transformative eventuality of deep knowledge in advancing early autism opinion.

The study named “Classifying Autism diapason complaint Using Machine Learning Approaches” by Zheng Z., Chen Y., and Yu J., published in 2022 in Frontiers in Psychiatry, investigates the use of machine knowledge ways for autism type, with a focus on the K- Nearest Neighbors (KNN) algorithm. The authors illuminate KNN as a simple and intuitive algorithm that performs effectively on small datasets and does n't bear a training phase, making it accessible and straightforward to apply. A notable merit of this exploration is the algorithm's capability to deliver dependable results in scripts with limited data, which is constantly the case in autism- related studies. Its simplicity ensures that it's easy to understand and apply, particularly in clinical surroundings with constrained computational coffers.

Still, the study identifies several limitations associated with KNN. The algorithm can be computationally precious when applied to large datasets, as it requires storing and comparing all data points during type. also, its performance is largely sensitive to inapplicable features in the data and the choice of distance metric, which can lead to sour results. Despite these challenges, the study underscores the eventuality of KNN in autism type and its connection in specific surrounds with manageable datasets and well- curated features.

III. Problem Statement

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition that impacts communication, behavior, and social interaction, often appearing in early childhood. Timely and accurate diagnosis is crucial for effective intervention, yet conventional diagnostic methods are time-intensive, subjective, and rely heavily on clinical expertise. With the increasing availability of behavioral, cognitive, and genetic data, there is a growing need for intelligent systems that can assist in early and reliable ASD detection. This study aims to address the limitations of traditional diagnostic approaches by developing an automated, accurate, and scalable prediction system using machine learning and ensemble techniques. By integrating diverse datasets and leveraging multiple classification algorithms, the proposed system seeks to improve diagnostic precision, support clinical decision-making, and enable early intervention.

IV. Proposed Methodology

The proposed methodology is designed to enhance the early vaticination and opinion of Autism Spectrum complaint (ASD) through a robust integration of multi-source data and advanced machine learning fabrics. The foundation of the system rests on comprehensive data collection involving behavioral assessments, cognitive testing, and inheritable profiling. All data inputs are acquired through well-established and formalized protocols to insure uniformity, trustability, and interoperability across different clinical and exploration surroundings. Once collected, the raw data suffer preprocessing, which includes cleaning, normalization, missing value insinuation, and metamorphosis into machine-readable formats. point selection ways similar as Recursive point Elimination (RFE), star element Analysis (PCA), and collective Information areapplied to identify and retain only the most applicable features. In addition to lowering dimensionality, this also improves the model's performance and interpretability.

A mongrel machine literacy approach is employed, which combines traditional bracket models similar as Logistic Retrogression, Random Forest, Support Vector Machine (SVM), Naive Bayes, Decision Trees, and K- Nearest Neighbors (KNN) with advanced deep literacy infrastructures including Convolutional Neural Networks (CNNs) and intermittent Neural Networks (RNNs). These models are trained and estimated on structured and unshaped data using stratified k-fold cross-validation. Ensemble strategies, particularly a Voting Classifier, are enforced to aggregate prognostications, thereby maximizing conception and reducing bias.

To enable real- time analysis, wearable detectors are integrated into the system. These bias continuously capture behavioral data similar as stir, eye movement, communication patterns, and physiological criteria like heart rate variability. The data is streamed to a pall structure where it's anatomized in near real- time using thepre-trained models, allowing for early discovery of atypical geste.The system offers a stoner-friendly interface for clinicians and caregivers.

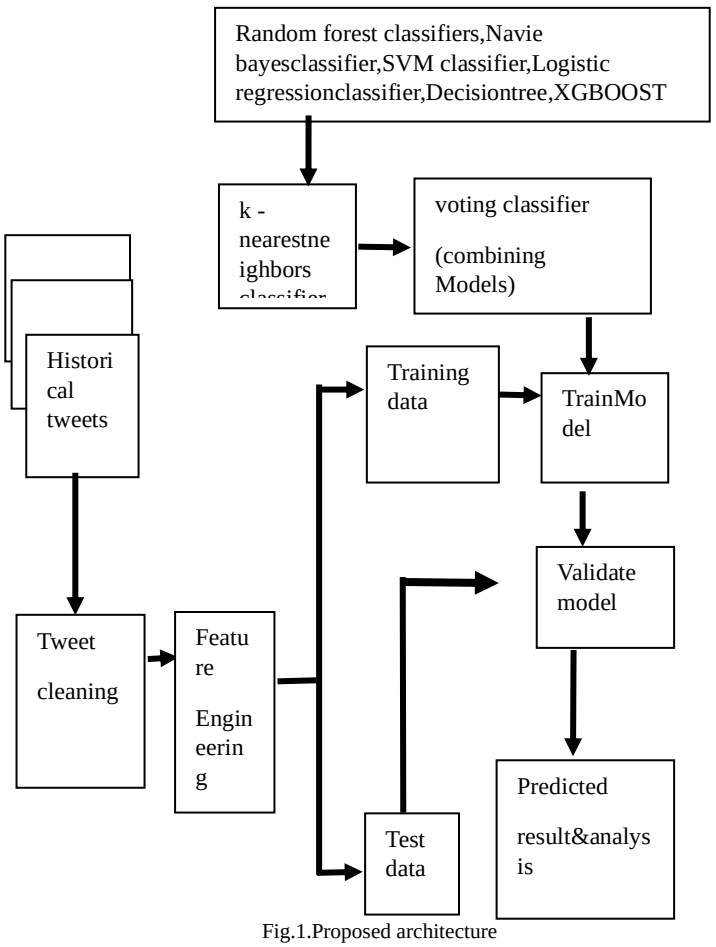


Fig.1.Proposed architecture

http://ijses.com/
All rights reserved

With the help of solvable AI tools like SHAP values, this interface provides visual dashboards, alerts, and explanatory perception derived from model work. The interface also supports feedback mechanisms that contribute to nonstop literacy and adaption of the prophetic models grounded on new patient data. This integrative and adaptive methodology ensures that the ASD vaticination system remains accurate, scalable, interpretable, and clinically feasible, making it a precious tool in the sphere of early neurodevelopmental complaint discovery.

V. Implementation Of Core Components

1.Data Collection

This module collects behavioral and cerebral data from colorful sources similar as questionnaires, checks, and multimedia inputs(audio, videotape, or eye- shadowing). It gathers both structured data(similar as check responses) and unshaped data(similar as patterns from videotape or audio), enabling a comprehensive analysis of ASD- related actions. The data collection process is designed to benon-invasive and can be fluently integrated into being clinical settings or exploration surroundings.

2.Data Preprocessing

Raw data collected from different sources frequently requires preprocessing to insure that it's clean, standardized, and ready for analysis. This module handles tasks similar as removing noise, normalizing data, handling missing values, and transubstantiating the data into a suitable format for machine literacy models. Preprocessing also includes point birth and selection, where the most applicable attributes are linked and used for model training.

3.Feature Selection

This module identifies the most applicable attributes from the dataset to ameliorate model performance. By applying ways similar as Recursive point Elimination and collective Information, this module reduces dimensionality, mitigates overfitting, and enhances the system's capability to concentrate on the crucial factors impacting autism. discovery.

4.Machine Learning Models

This module tools and trains multiple algorithms, including Random Forest, Support Vector Machine (SVM), k-Nearest Neighbors (k- NN), Naive Bayes, Logistic Regression, and XGBoost. This module also focuses on optimizing hyperparameters to insure each model performs at its stylish for autism bracket.

5.Voting Classifier

The Module combines prognostications from all trained machine literacy models to enhance overall delicacy. By using either weighted or hard voting mechanisms, this module leverages the strengths of individual models, performing in further dependable and robust bracket issues.

6.Evaluation

The Module measures the performance of the system using criteria similar as delicacy, perfection, recall, F1 score, and ROC- AUC. This module compares the performance of individual models with the voting classifier, furnishing perceptivity into their effectiveness and imaging results through confusion matrices and bracket reports

7.Deployment

This Module prepares the system for real- world operation by creating a secure and stoner-friendly interface, similar as a web or mobile operation, for healthcare professionals. This module supports batch and real- time prognostications while icing secure storehouse and reclamation of patient data, making the system accessible and practical for clinical surroundings

8.Future Integration

This Module focuses on extending the system's capabilities by integrating multimodal data sources similar as speech analysis and imaging data. It aims to enhance the system's rigidity to different workflows in clinical surroundings, icing scalability and bettered effectiveness for early autism discovery.

VI. Algorithm Implementation

1. Random Forest classifier

Random Forest is an ensemble literacy system that constructs multiple decision trees during training and merges their prognostications to ameliorate delicacy and help overfitting.

2. Support Vector Machine (SVM)

SVM is a supervised literacy model that finds the optimal boundary(hyperplane) to separate data into distinct classes, especially effective for high- dimensional ornon-linear datasets.

3. k- Nearest Neighbors (k- NN)

k- NN is a simple algorithm that classifies data points grounded on the maturity class among their 'k' closest neighbors in the point space.

4. Naive Bayes classifier

Naive Bayes is a probabilistic classifier grounded on Bayes' Theorem, assuming point independence, which makes it fast and effective for textbook and bracket problems.

5. Logistic Retrogression

Logistic Retrogression is a direct model used for double bracket, prognosticating the probability of a class by applying a sigmoid function to a direct combination of input features.

6. Decision Tree

A Decision Tree is a flowchart- suchlike model that splits data into branches grounded on point values, helping make opinions or groups in a hierarchical structure.

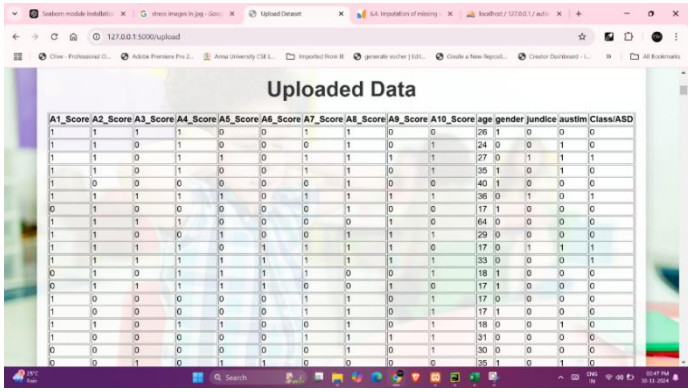
7. XGBoost (Extreme Gradient Boosting)

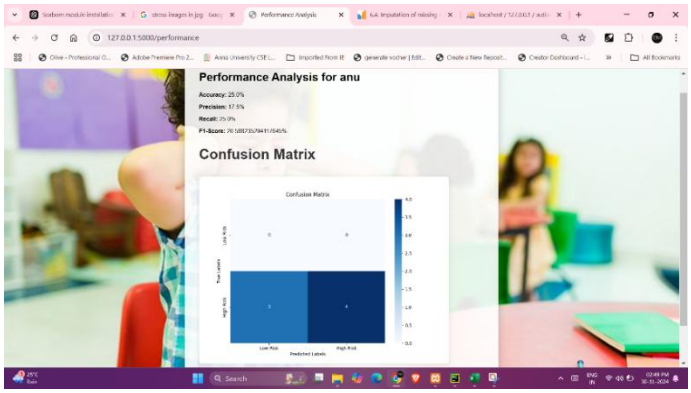
XGBoost is an advanced boosting fashion that builds decision trees successionally, where each new tree corrects the crimes made by former bones, optimizing performance and speed.

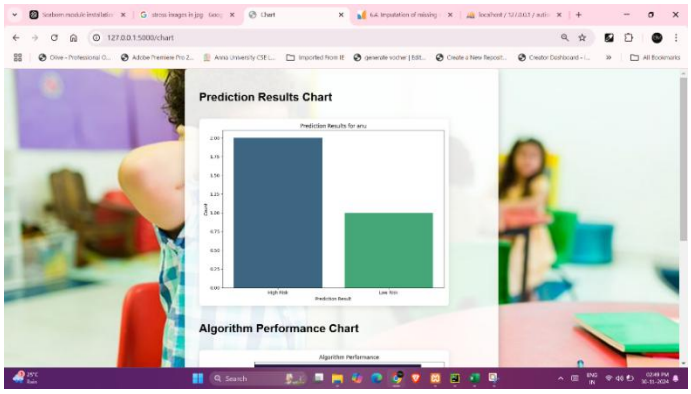
http://ijses.com/
All rights reserved

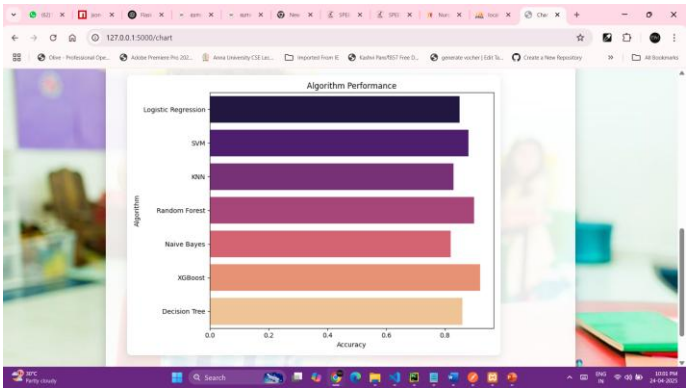
VII. Result

The system is estimated using criteria analogous as delicacy, perfection, recall, F1- score, and ROC- AUC. Individual model performances are compared with the ensemble Voting Classifier. Experimental results demonstrate that the crossbred ensemble model outperforms individual classifiers in both prophecy delicacy and robustness. The addition of real- time monitoring data further enhances predictive capabilities by landing behaviour patterns in natural settings. Visualizations analogous as confusion matrices and ROC angles are used to anatomize performance.









VIII. Conclusion

This paper presents a comprehensive and intelligent system for the early discovery of Autism Spectrum Disorder. By integrating multi-source data, applying mongrel machine literacy models, and exercising wearable detectors for real-time behavioral monitoring, the system offers a scalable and practical result for clinicians. The addition of an interpretable and stoner-friendly interface ensures clinical relinquishment, while interdisciplinary collaboration promotes continual system improvement. unborn work includes expanding the dataset, integrating further detector modalities, and planting the system in real- world clinical settings.

IX. Future Scope

The proposed system lays the groundwork for enhancing early autism diagnosis through intelligent technologies. In the future, this model can be extended by incorporating deep learning frameworks such as CNNs and RNNs to analyze complex, high-dimensional data like speech, facial expressions, or EEG signals. Integration of real-time data from wearable devices and IoT platforms can further enhance continuous monitoring and early warning systems. Additionally, combining multimodal data sources—such as genetic, behavioural, and environmental information—can provide a more comprehensive understanding of ASD. The deployment of this system as a cloud-based or mobile application will also make it more accessible to clinicians and caregivers, especially in remote or resource-limited areas. Future research can focus on improving interpretability, scalability, and personalization of predictions to support individualized therapy and care strategies.

References

[1] Bhat, S. K., et al.(2023). Predicting Autism Spectrum Disorder Using Random Forest Algorithm and Feature Selection. This paper discusses the use of arbitrary timber algorithms in confluence with point selection ways to ameliorate vaticination delicacy for ASD. Journal of Autism and Developmental diseases.

[2] Chakraborty, R., et al.(2023). Machine Learning Approaches for Autism Spectrum Disorder opinion A Methodical Review. This review summarizes colorful machine literacy models and their effectiveness in diagnosing ASD grounded on behavioral and clinical data. Neurocomputing.

[3] Dutta, R., et al.(2023). A Novel Deep Learning Approach for Autism Spectrum Disorder opinion Using EEG Signals. The paper presents a deepliteracy- grounded frame for ASD discovery using electroencephalography (EEG) data. IEEE Deals on Neural Systems and Rehabilitation Engineering.

All rights reserved

[4] Gomez, M. E., et al.(2023). Prophetic Modelling of Autism Spectrum Disorder UsingMulti- Modal Data:A Machine Learning Approach. The study explores prophetic models grounded onmulti-modal datasets, including genetics, clinical data, and behavioral features. Artificial Intelligence in Medicine.
[5] Hernandez, A., et al.(2023). A Comparison of Machine Learning Algorithms for Autism opinion Using inheritable and Behavioral Data. This exploration compares multiple machine learning algorithms for ASD opinion grounded on inheritable and behavioral data. Bioinformatics.
[6] IEEE Xplore Conference Paper(2023). Autism Diapason complaint vaticination Using Machine literacy.This publication discusses the operation of colorful machine learning algorithms for ASD vaticination,fastening on their delicacy and scalability in real- world surroundings.
[7] Kumar, S., et al.(2024). Combining Multimodal Data for Enhanced ASD vaticination Using Deep literacy. This study integrates inheritable, behavioral, and environmental factors, demonstrating bettered vaticination issues.
[8] Lopez, M., and Chen, Y.(2023). Support Vector Machines and Ensemble Learning in AutismDetection. A Case Study. This exploration compares SVM and ensemble styles, showcasing their efficacity in handling complex datasets for ASD opinion.

http://ijses.com/
All rights reserved