Autism Spectrum Disorder Diagnosis: A Comprehensive Review of Machine Learning Approaches

Deepti Prasad[1], Suman Bhatia[2]

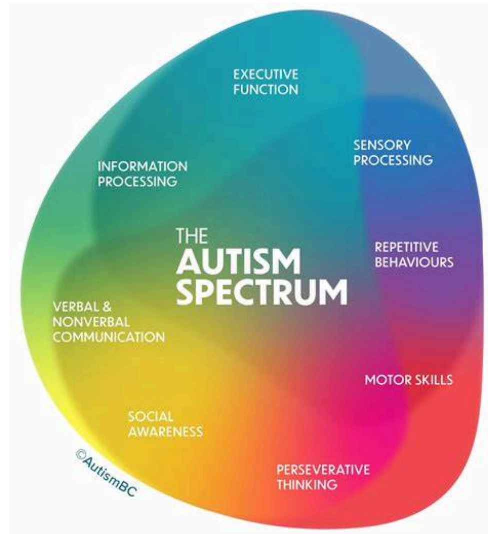
1. **deeptiprasad2015@gmail.com**

2. **ersuman80@gmail.com**

Abstract: Understanding autism spectrum disorder (ASD) and offering effective, individualised therapies present special challenges. This thorough review analyses various facts about ASD, from its early diagnosis and medical treatments to its cutting-edge therapeutic strategies and research endeavours. The study develops through thoroughly examining cutting-edge machine learning (ML) and deep learning (DL) methods used in ASD research. Deep learning, Conditional Inference Forest (CF), Random Forest (RF), and Support Vector Machines (SVM) emerge as key methods for understanding the complicated genomic and neuroimaging landscapes of ASD.

Children with autism benefit from play therapy and game-based interventions because they provide opportunities for social interaction and skill development. The study also looks into the significance of awareness and early intervention. It offers a forward-thinking road map and highlights the critical need for public education, easily available screening services, and specialised educational programmes. The technology will be based on predictions of disease risk, deficits, and behavioural patterns. curating a dataset centred on Indian children and adults and integrating emotional and behavioural nuances. Collaborations with healthcare facilities and governmental organisations can be planned to close the accessibility gap and ensure that ASD treatments are available to people from every socioeconomic background, not just the wealthy few.

Keywords *Autism Spectrum Disorder (ASD), Machine Learning Algorithm, Hemisphere Asymmetry, Magnetic resonance imaging -MRI (fMRI,sMRI), Confusion Matrix, Mean Squared Error (MSE) and Mean Absolute Error (MAE), Conditional Inference Forest (CF)*

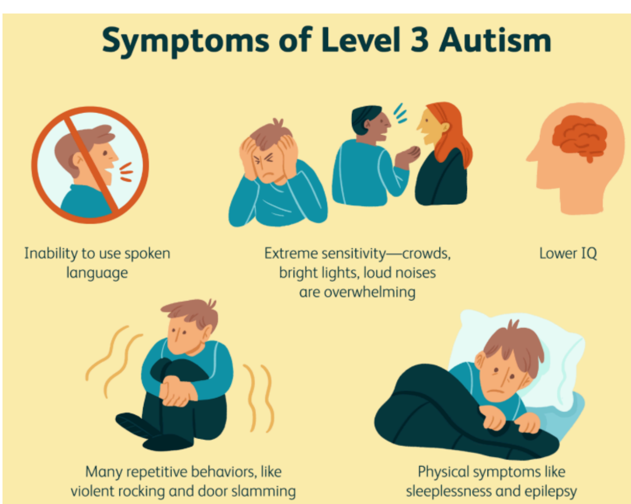


1. Introduction

ASD research is being revolutionised by supervised machine learning in combination with big data, allowing for better insights and customised interventions for people with ASD and related diseases [5][12]. Understanding ASD in the context of brain networks is essential because it highlights the need for thorough diagnostic frameworks that can handle a range of severity levels and support needs [9]—these aspects of ASD when combined with better predictions using machine learning algorithms on available datasets. [ImageSource: https://tinyurl.com/muk8sn68]

1.1 Autism and its diagnosis

ASD, which affects one in 59 children in the USA [5], is characterised by social communication and interaction difficulties. The main approaches for diagnosing ASD at this time are clinical standardised tests like the Autism Diagnostic Interview-Revised (ADI-R) and Autism Diagnostic Observation Schedule Revised (ADOS-R), however, they are time-consuming and expensive [14]. Children with ASD can be identified using screening measures including the Autism Spectrum Quotient (AQ), Childhood Autism Rating Scale-2 (CARS-2), and Screening Tool for Autism in Toddlers and Young Children (STAT) [14]. Big data utilisation in ASD research is still lacking due to the unavailability of proper datasets, but recent developments in data collecting [5]. ASD diagnosis, genetic knowledge, and successful intervention development are all being investigated with machine learning [5].



In line with global trends, men had a higher prevalence of ASD than women.

The study, however, contradicted other studies that suggested more severe social impairment among males with ASD and showed no significant link between gender and ASD.[Image Source:https://tinyurl.com/mpk49he7]

This groundbreaking Indian study, unlike others, uncovered a lower autism prevalence (0.15%) in 1-10-year-olds compared to Western/Asian research (1-2%). Interestingly, rural children in India had higher rates, which might suggest different factors affecting diagnosis or prevalence in diverse settings.

Moderate autism was more prevalent in older children (4-10 years), suggesting potential delays in diagnosis linked to speech and motor development. This is concerning, as early detection of ASD is possible even before age 2. Furthermore, socioeconomic disparities were evident, with rural "upper class" and urban/tribal "middle class" children showing higher rates. The study suggests limited awareness in lower SES groups might contribute to these disparities, leading to later identification and difficulties accessing needed support. Additionally, it emphasizes the critical gap in addressing developmental disorders within child health programs in India, creating significant challenges for children with ASD due to the lack of effective identification, referral, and support systems.

Diagnoses and knowledge of the biochemical and genetic underpinnings of intellectual and developmental disabilities (IDDs), such as autism spectrum disorder (ASD), are extremely difficult [12]. Biological understanding of IDDs is being improved by developments in high-throughput sequencing, imaging, and AI technologies [12]. The neurodevelopmental disorders Attention-Deficit/Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD) are common in children [20]. Because clinical psychology and neuroscience are only recently being integrated, it is difficult to pinpoint the precise brain regions that are linked to these illnesses [20].

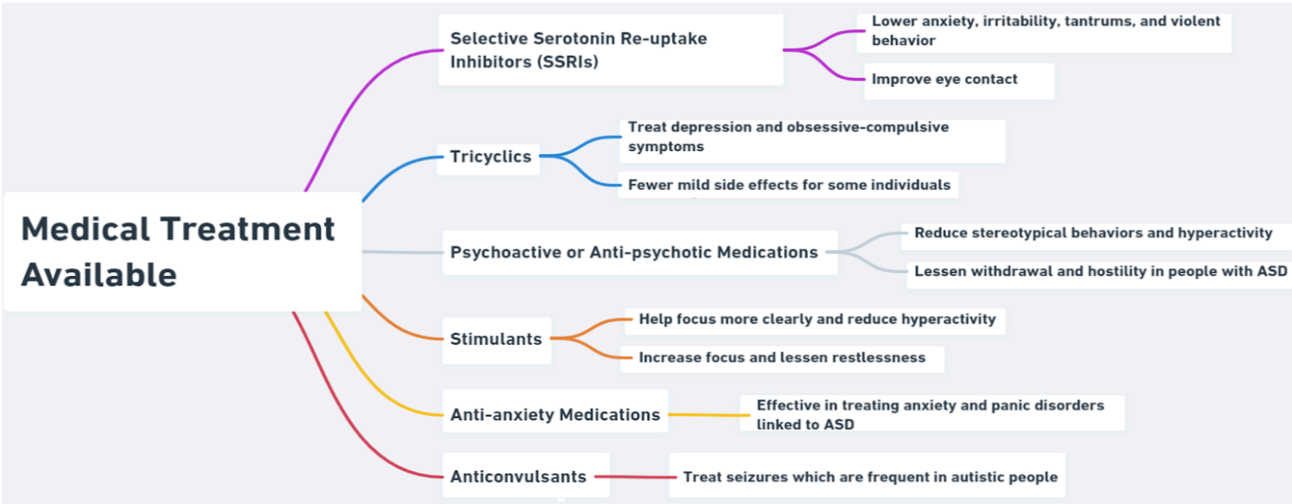


Figure 3. Various Medical Treatments Available for Autistic Persons

2. Machine Learning and Deep Learning Algorithms

Inspired by the human brain, deep learning is a potent branch of machine learning that uses artificial neural networks that are networked to understand complex patterns and relationships from massive volumes of data. Deep learning is beneficial for applications like picture identification, natural language processing, and even autism prediction since it excels at automatically extracting this information, unlike classical machine learning, which requires custom features. This enables it to spot minute trends in eye-tracking data, behavioural observations, or brain scans that other techniques might overlook. To guarantee that its promise for autism prediction translates into real-world benefits, however, issues like data constraints, interpretability, and ethical considerations demand responsible development and cautious integration with clinical competence.

2.1 Supervised Learning - To predict the diagnosis for new instances, these algorithms learn from labelled data (individuals with ASD diagnoses and typically developing individuals). Define decision boundaries using features to distinguish between people who have been diagnosed and those who have not. Combine many decision trees to provide flexibility and resilience. Brain-inspired models are complicated and able to learn complex patterns from big datasets.

2.2 Unsupervised Learning - These algorithms find hidden structures and patterns in unlabeled data, which may reveal modest ASD signs that are still poorly understood. organises people into groups according to shared qualities, which may reveal ASD spectrum subgroups. helps with data visualisation and analysis by reducing the dimensionality of the data while maintaining critical information.

Supervised/Unsupervised Learning Algorithms	**Usage in ASD Research**
Support Vector Machines (SVM)	Binary predictions related to diagnosis and screening. Used in genetics, neuroimaging, and text-mining studies for biomarker identification and accurate ASD classification.
Alternating Decision Tree (AD Tree)	Proficient in binary predictions and handling categorical data efficiently. Enhances diagnostic processes and provides transparency in ASD research.
Naïve Bayes	Specialized in text mining, particularly in analyzing social media data related to ASD for predicting ASD based on textual samples.
Random Forest (RF)	Featured in text mining and neuroimaging studies. Utilized for outlier detection, monitoring ASD populations, and developing screening tools due to their ensemble learning approach.
Logistic Regression	A reliable method for binary classification tasks in ASD diagnosis due to its ability to fit an optimal curve to data points using a logistic function.
K-Nearest Neighbours (KNN)	Classification of individuals based on their nearest neighbours' features. Used to gain insights into specific patterns and characteristics within ASD populations.

Conditional Inference Forest (CF)	Modification of Random Forest algorithm utilizing statistical inference tests for feature selection in ASD-related genetic data. Identifies key genetic factors linked to ASD, aiding in understanding molecular pathways.

Table 1. Summary of Supervised/Unsupervised Machine Learning Algorithms utilized for Autism Detection

2.3 Implementation Strategies

1. Collect diverse and representative data: Verify that datasets cover a range of age ranges, demographics, and autism presentations to reduce bias and enhance generalizability.

2. Prioritize data quality: To assure accuracy and reduce noise, apply strict data collection, cleaning, and pre-processing procedures.

3. Address data privacy and security: To secure sensitive information, get informed approval, hide data when it can, and put strong security measures in place.

4. Choose appropriate algorithms: Examine the drawbacks of various algorithms and choose the ones best suited to the particular detection task and the properties of the data.

5. Focus on early intervention: Make use of machine learning-driven findings to quickly identify possible scenarios and provide access to remedy resources.

6. Complementary tool, not a replacement: Consider machine learning as a helpful tool for medical professionals rather than as a substitute for thorough clinical assessment and customised decision-making.

2.4 Algorithms Efficiency

Table 2 also highlights the efficiency of usage of various supervised/unsupervised algorithm utilized for Autism Detection as given in literature review. This information highlights the significance of Machine Learning algorithms for Autism Detection.

Algorithms	Objective	Efficiency	Accuracy	R
Logistic Regression	Detection of autism spectrum disorder (ASD) in children and adults using machine learning	High	93.15%	[13]
Naïve Bayes	Detection of autism spectrum disorder (ASD) in children and adults using machine learning	High	97.53%	[13]
Support Vector Machine	Machine Learning Methods for Diagnosing Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder Using Functional and Structural MRI: A Survey	High	94.9%	[20]
Random Forest	Framework for Grading Autism Severity Using Task-Based FMRI	Moderate	72%	[25]
KNN	Machine Learning Methods for Diagnosing Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder Using Functional and Structural MRI: A Survey	Low	66%	[20]
SVM(children)	Detection of autism spectrum disorder (ASD) in children and adults using machine learning	High	93.84%	[13]
RF(rs-fMRI)	Applications of Supervised Machine Learning in Autism Spectrum Disorder Research	High	91%	[5]

Table 2.

Glmboost(toddlers,children,adults)	Detection of autism spectrum disorder (ASD) in children and adults using machine learning	High	97%	[13]
Elastic Net(rs-FMRI)	Automatic diagnosis of autism based on functional magnetic resonance imaging and elastic net	High	83.33%	[10]
Sparse LR(rs-fMRI,sMRI)	Diagnostic Classification for Human Autism and Obsessive-Compulsive Disorder Based on Machine Learning From a Primate Genetic Model	High	82.14%	[15]

Algorithmic Efficiency of various Machine Learning Algorithms for Autism Detection

2.5 Limitations of Machine Learning and Deep Learning in Autism Detection

1. Data Bias and Imbalance: Algorithm training datasets may overrepresent certain autism presentations or be biased towards specific factors, which could result in predictions that are not accurate for a variety of groups.

2. Interpretability and Explainability: It's important to know how algorithms arrive at their results, particularly in complex fields like healthcare. However complicated models can be clear, which makes it challenging to understand their logic and spot any biases.

3. Limited Generalizability: To guarantee greater applicability, algorithms trained on particular datasets may not generalise well to other populations or circumstances, needing extensive validation and modification.

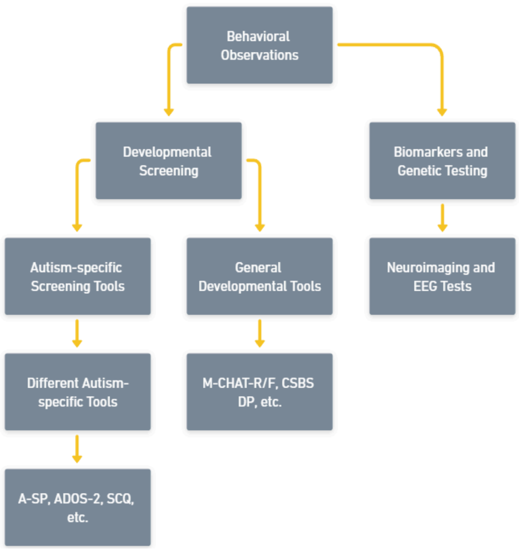
4. Overdiagnosis and Missed Diagnoses: If algorithms are too sensitive, people who don't fit all the criteria may be overdiagnosed with ASD, and people with less common presentations may go unnoticed.

5. Lack of Clinical Integration: Clinical judgement and skill should be added, not replaced, by machine learning and deep learning technologies. It is essential to combine with established diagnostic procedures and clinical assessment.

2.6 Techniques for Prediction

Medical diagnosis of autism spectrum disorder (ASD) entails identifying probable symptoms and risk factors using a variety of procedures and evaluations. Here are some techniques commonly used for predicting autism

1. Developmental Screening: To determine a child's developmental milestones, paediatricians frequently perform developmental screenings at well-child visits. Further assessment may be warranted if there are speech, motor, or social development delays.



2. Developmental Surveillance: During checks, medical professionals frequently observe a child's development, enabling them to track any concerns over time.

3. Autism-Specific Screening Tools: The Modified Checklist for Autism in Toddlers (M-CHAT) is one of several screening instruments made specifically to determine the likelihood of autism. Positive outcomes from these tools suggest that more testing is required.

4. Genetic Testing: An elevated risk of autism is linked to specific genetic abnormalities. These mutations may be found in people through genetic testing, particularly in those who have a family history of ASD.

5. Brain Imaging: Electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) are sometimes used to detect variations in brain activity in people with ASD[25]. These methods are more frequently used in research settings.

6. Biomarker Research: ASD risk may be indicated by biomarkers that have been found in blood, urine, or cerebrospinal fluid, according to ongoing studies. However, research in this field is still in its infancy.

7. Eye-Tracking Technology: Studies using eye-tracking technology have shown that people with ASD interpret visual information differently. As a potential tool for early detection, this technique is now being investigated.

2.7 Attributes for Prediction

1. Behavioural Observations: A child's behaviour, speech, and social interactions are evaluated by qualified professionals. Specific behavioural patterns may indicate a higher risk of autism.

2. Parental Concerns: Parents frequently become aware of unusual behaviours in their kids first. Medical examinations are prompted in large part by their findings and concerns.

3. Discussion

Considering the difficulties encountered by autistic children, especially their inclination towards playing alone, this project argues for multiple methods that help their social and developmental growth. Using a child's particular interests and attention span, games therapy, a specialised form of play therapy, becomes an important tool for developing social skills, communication, and emotional expression through lighthearted activities. Through carefully constructed games, therapists help kids in making the transition from alone time to social engagement. This lets kids explore their feelings, make sense of their environment, and form relationships with friends and family. Furthermore, game therapy concepts can be learned by parents and guardians, who can then use these insights to build stronger family bonds, promote general well-being, and actively participate in their child's growth.

There is more to helping autistic children than just therapeutic techniques. This research highlights the need for early awareness efforts that inform communities about the weak signs of ASD and highlight the transformative power of early diagnosis and intervention. These campaigns aim to create a more inclusive society. The complete plan envisions a time when young people with autism are not only recognised and assisted, but also given the tools they need to flourish through strong support systems, complementary therapies, ongoing research, and a seamless transition into adulthood.

4. Future Work

We envision a cutting-edge software system using predictive analytics. This platform will examine trends and patterns across many data sets, including emotional nuance and health metrics, and analyse data from government-approved health exams. By dissecting these patterns, we hope to get insight into the emotional and behavioural patterns of autistic kids, possibly identify physical vulnerabilities, and even forecast the likelihood of specific diseases. There is a great deal of promise for proactive treatment and enhanced well-being for autistic people with this technology.

We place a high priority on empowering families by providing them with seminars and easily accessible resources because we know that caring, knowledgeable carers provide a nurturing home environment for children with ASD. To build an inclusive culture from the bottom up, we collaborate with educational institutions to develop specialised programmes and support inclusive curriculums in mainstream schools. This fosters acceptance and empathy among peers in addition to understanding.

Our goal goes beyond data analysis to make accessible autism treatment. Our software will forecast possible problems and direct preventive care by evaluating behavioural, emotional, and health data. We intend to collaborate with governments and healthcare facilities to lower the cost and increase accessibility of this instrument, particularly for individuals facing difficult financial circumstances. To ensure that everyone, regardless of poverty, may benefit from proactive, predictive autism care, this project addresses the high expenses associated with private therapy.

5. Conclusion

Because autism spectrum disorder (ASD) has complicated symptoms and insufficient diagnostic tools, diagnosing the condition can be difficult. There is potential for improvement in traditional procedures, which rely on laborious psychological testing and observations.

The diagnosis of ASD could be revolutionised by machine learning. ASD is highly accurately predicted by algorithms such as Support Vector Machine and Logistic Regression, which may help with diagnosis, screening, and understanding genetic underpinnings. Nonetheless, issues like the interpretability of the model and the lack of open-source data must be resolved. There is great potential for quicker, more accurate diagnosis and a better understanding of the molecular causes of ASD by combining machine learning and deep learning techniques. Gaining confidence and a wider acceptance of these models in clinical contexts depends on ensuring their interpretability.

References:

[1] K. S. Omar, P. Mondal, N. S. Khan, Md. R. K. Rizvi, and M. N. Islam, “A Machine Learning Approach to Predict Autism Spectrum Disorder,” in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Feb. 2019, pp. 1–6. doi: 10.1109/ECACE.2019.8679454.

[2] A. Singh, N. Thakur, and A. Sharma, “A review of supervised machine learning algorithms,” in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, Mar. 2016, pp. 1310–1315. Available: https://ieeexplore.ieee.org/abstract/document/7724478

[3] I. Jamwal, D. Malhotra, and M. Mengi, “A systematic study of intelligent autism spectrum disorder detector,” *International Journal of Computational Vision and Robotics*, vol. 13, no. 2, pp. 219–234, Jan. 2023, doi: 10.1504/IJCVR.2023.129435.

[4] J. Licinio and M.-L. Wong, “Advances in autism research, 2021: continuing to decipher the secrets of autism,” *Mol Psychiatry*, vol. 26, no. 5, Art. no. 5, May 2021, doi: 10.1038/s41380-021-01168-0.

[5] “Applications of Supervised Machine Learning in Autism Spectrum Disorder Research: a Review | Review Journal of Autism and Developmental Disorders.”. Available: https://link.springer.com/article/10.1007/s40489-019-00158-x

[6] “ARFID Genes and Environment (ARFID-GEN): study protocol | BMC Psychiatry.” Available: https://link.springer.com/article/10.1186/s12888-023-05266-x

[7] I. Jamwal, D. Malhotra, and M. Mengi, “Autism Spectrum Disorder Detection Using ASD_sfMRI,” in *Computer Vision and Robotics*, J. C. Bansal, A. Engelbrecht, and P. K. Shukla, Eds., in

Algorithms for Intelligent Systems. Singapore: Springer, 2022, pp. 175–189. doi: 10.1007/978-981-16-8225-4_14.
[8] “Autism Spectrum Disorder Diagnosis Assistance using Machine Learning | Revista de Informática Teórica e Aplicada.” Available: https://seer.ufrgs.br//rita/article/view/126309
[9] P. Moridian *et al.*, “Automatic autism spectrum disorder detection using artificial intelligence methods with MRI neuroimaging: A review,” *Frontiers in Molecular Neuroscience*, vol. 15, 2022, Available: https://www.frontiersin.org/articles/10.3389/fnmol.2022.999605
[10] W. Liu, M. Liu, D. Yang, M. Wang, and T. Tao, “Automatic diagnosis of autism based on functional magnetic resonance imaging and elastic net,” in *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, Jun. 2020, pp. 104–108. doi: 10.1109/ITOEC49072.2020.9141766.
[11] “Biomedicines | Free Full-Text | Role of Artificial Intelligence for Autism Diagnosis Using DTI and fMRI: A Survey.”. Available: https://www.mdpi.com/2227-9059/11/7/1858
[12] “Bringing machine learning to research on intellectual and developmental disabilities: taking inspiration from neurological diseases | Journal of Neurodevelopmental Disorders.” Available: https://link.springer.com/article/10.1186/s11689-022-09438-w
[13] M. S. Farooq, R. Tehseen, M. Sabir, and Z. Atal, “Detection of autism spectrum disorder (ASD) in children and adults using machine learning,” *Sci Rep*, vol. 13, no. 1, Art. no. 1, Jun. 2023, doi: 10.1038/s41598-023-35910-1.
[14] K. Vakadkar, D. Purkayastha, and D. Krishnan, “Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques,” *SN COMPUT. SCI.*, vol. 2, no. 5, p. 386, Jul. 2021, doi: 10.1007/s42979-021-00776-5.
[15] “Diagnostic Classification for Human Autism and Obsessive-Compulsive Disorder Based on Machine Learning From a Primate Genetic Model | American Journal of Psychiatry.”. Available: https://ajp.psychiatryonline.org/doi/full/10.1176/appi.ajp.2020.19101091
[16] X. Yang, M. S. Islam, and A. M. A. Khaled, “Functional connectivity magnetic resonance imaging classification of autism spectrum disorder using the multisite ABIDE dataset,” in *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, May 2019, pp. 1–4. doi: 10.1109/BHI.2019.8834653.
[17] “IJMS | Free Full-Text | Autism Spectrum Disorder: Neurodevelopmental Risk Factors, Biological Mechanism, and Precision Therapy.” Available: https://www.mdpi.com/1422-0067/24/3/1819
[18] H. S. Nogay and H. Adeli, “Machine learning (ML) for the diagnosis of autism spectrum disorder (ASD) using brain imaging,” *Reviews in the Neurosciences*, vol. 31, no. 8, pp. 825–841, Dec. 2020, doi: 10.1515/revneuro-2020-0043.
[19] “Machine learning in autistic spectrum disorder behavioral research: A review and ways forward: Informatics for Health and Social Care: Vol 44, No 3.”. Available: https://www.tandfonline.com/doi/abs/10.1080/17538157.2017.1399132
[20] T. Eslami, F. Almuqhim, J. S. Raiker, and F. Saeed, “Machine Learning Methods for Diagnosing Autism Spectrum Disorder and Attention- Deficit/Hyperactivity Disorder Using Functional and Structural MRI: A Survey,” *Frontiers in Neuroinformatics*, vol. 14, 2021, Available: https://www.frontiersin.org/articles/10.3389/fninf.2020.575999
[21] J. Veenstra-VanderWeele and E. H. Cook, “Molecular genetics of autism spectrum disorder,” *Mol Psychiatry*, vol. 9, no. 9, Art. no. 9, Sep. 2004, doi: 10.1038/sj.mp.4001505.
[22] “Primary Care Autism Screening and Later Autism Diagnosis | Pediatrics | American Academy of Pediatrics.” Available: https://publications.aap.org/pediatrics/article/146/2/e20192314/77011?utm_source=TrendMD&utm_medium=TrendMD&utm_campaign=Pediatrics_TrendMD_0?autologincheck=redirected
[23] “Pro Machine Learning Algorithms : A Hands-On Approach to Implementing Algorithms in Python and R | SpringerLink.” Available: https://link.springer.com/book/10.1007/978-1-4842-3564-5

[24] “SciELO - Brasil - Autism spectrum disorders: an overview on diagnosis and treatment Autism spectrum disorders: an overview on diagnosis and treatment.” Available: https://www.scielo.br/j/rbp/a/GXxmGC7gqrG8FMhzLB5RcLw/

[25] Haweel, R., Dekhil, O., Shalaby, A., Mahmoud, A., Ghazal, M., Khalil, A., ... & El-Baz, A. (2020, April). A novel framework for grading autism severity using task-based fmri. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (pp.1404-1407). IEEE.