[ICAETBM-2024]
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

Enhancing Asd Detection Through Ensemble Learning : Integrating Image and Questionnaire Data

Prof. Seema Mahalungkar, Rupesh Dubey, Saad Mansuri, Onkar Dhamma, Satyanshu Sharma
Department of Computer Science and Engineering Nutan college of Engineering and Research Pune, India

ABSTRACT

Autism Spectrum Disorder (ASD) affects a significant portion of the population, with rates on the rise. The diagnosis process is not only time-consuming but also costly, posing challenges for patients in adhering to prescribed treatments and hindering their progress. This project aims to streamline the diagnosis process through machine learning techniques. Three datasets—ASD Screening Data for Adults, Children, and Adolescents— are utilized.

ASD is a complex neurological condition characterized by social communication deficits and repetitive behaviors. Early detection is critical for effective intervention. This paper introduces a novel approach to ASD detection, combining screening methods and image processing techniques. Ensemble learning, a robust machine learning method, is employed to improve classification accuracy by integrating multiple models. The proposed methodology seeks to contribute to early ASD diagnosis, enabling timely intervention and support for individuals with ASD.

This algorithm facilitates non-invasive detection of ASD, eliminating the need for surgical procedures. Furthermore, it can be implemented as a user-friendly GUI or software in various mental health institutions. This not only reduces expenses for ASD patients but also simplifies processes for hospitals.

Keywords: Autism Spectrum Disorder, Screening Method, Image Processing, Ensemble Learning, Early Diagnosis .

I. INTRODUCTION

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a heterogeneous neurodevelopmental condition characterized by impairments in social interaction, communication difficulties, and restricted, repetitive patterns of behavior. According to the Centers for Disease Control and Prevention (CDC), ASD affects approximately 1 in 54 children in the United States, highlighting the significant public health concern associated with this disorder [1]. Early identification of ASD is critical for initiating interventions that can improve outcomes and enhance the quality of life for affected individuals and their families.

Traditional methods for ASD diagnosis involve clinical assessment by trained professionals, which often rely on behavioral observations and standardized assessments. However, these methods can be time-consuming, costly, and subjective, leading to delays in diagnosis and treatment initiation. In recent years, there has been growing interest in leveraging technological advancements, particularly in the fields of machine learning and image processing, to develop more efficient and objective tools for ASD detection.

Autism Spectrum Disorder (ASD) is a neurological condition impacting millions of children globally. According to the Centers for Disease Control and Prevention (CDC) in 2021, approximately 1 in 44 children are diagnosed with ASD. Characteristics include difficulties in social interaction, communication challenges, repetitive behaviors, peculiar movements, and narrow interests. Typically, symptoms manifest around the age of two. Genetic and environmental factors are known contributors to ASD. Each patient exhibits a unique set of symptoms. Early detection is crucial for timely intervention, though ASD currently lacks a definitive cure. Access to clinical expertise is limited, particularly in rural areas, resulting in delayed identification compared to urban regions. Clinical interventions are time-intensive, making it challenging to assist every ASD child adequately. Hence, integrating and deploying various technologies for ASD detection is becoming increasingly imperative. Recent research has explored a range of techniques and approaches, with a focus on IoT-based methods and machine learning (ML) approaches.

II. LITERATURE REVIEW

In their 2021 study, Goel et al. [1] introduced a novel optimization algorithm aimed at enhancing the performance of common machine learning techniques for Autism Spectrum Disorder (ASD) detection. The proposed algorithm, denoted as MGOA (GOA with Random Forest classifier), was evaluated alongside several existing methods, including ASDTest, GOA, BACO, LR, NB, KNN, and RF-CART + ID3. The results indicated that the MGOA achieved remarkable accuracy, specificity, and sensitivity, all reaching approximately 100%. This suggests the potential effectiveness of the MGOA in accurately predicting ASD cases, showcasing its superiority over other evaluated machine learning algorithms.

Shahamiri and Thabtah [2] focused on the implementation and evaluation of a Convolutional Neural Network (CNN)-based scoring system for Autism Spectrum Disorder (ASD). The study involved the utilization of Q-CHAT-10 and AQ-10 assessments, and the developed system was benchmarked against ASDTest as well as other algorithms, namely C4.5, Bayes Net, and RIDOR. The results of the performance evaluation demonstrated the superior capabilities of the CNNbased scoring system, highlighting its robustness when compared to alternative algorithms. This suggests the potential effectiveness of the implemented CNN for ASD scoring, showcasing its promise in improving accuracy and reliability in the assessment of ASD.

Thabtah and Peebles [3] aimed to demonstrate the superiority of a Rules-based Machine Learning (RML) approach over other models in the context of Autism Spectrum Disorder (ASD) classification. With a comprehensive analysis involving QCHAT-10 and AQ-10 assessments across different age groups (child, adolescent, adult), the study evaluated various machine learning models, including RIPPER, RIDOR, Nnge, Bagging, CART, C4.5, and PRISM, with RML being the focal point. The empirical evaluation encompassed different ASD datasets, and the findings revealed that the RML model not only excelled in classifying ASD but also provided interpretable rules that could be employed to understand the underlying reasons behind the classification. This emphasizes the efficacy of the RML approach in ASD classification, offering both accuracy and interpretability in the diagnostic process.

Wall et al. [4] undertook the task of streamlining the Autism Diagnostic Interview-Revised (ADI-R) and evaluating the machine learning (ML) performance in ASD classification. The study involved the ADI-R as the primary diagnostic instrument and utilized data from AGRE, SSC, and AC datasets. Feature selection was performed through a trial-and-error process, and the ML models were evaluated using various algorithms, including ADTree, BFTree, ConjunctiveRule, DecisionStump, FilteredClassifier, J48, J48graft, JRip, LADTree, Nnge, OneR, OrdinalClassClassifier, PART, Ridor, and SimpleCart. The results of the evaluation revealed that the best-performing model employed only 7 out of the 93 items contained in the ADI-R but achieved an impressive classification accuracy of 99.9%. This finding suggests that a streamlined set of features from the ADIR, coupled with the selected ML algorithms, can effectively classify Autism Spectrum Disorder with high accuracy, potentially providing a more efficient and practical diagnostic approach.

In their research published in 2019, Duda et al. [5] focused on streamlining the Autism Diagnostic Observation Schedule (ADOS) and aimed to showcase the superior performance of the ADTree algorithm compared to common hand-crafted methods. The study utilized data from multiple sources, including AC, AGRE, SSC, NDAR, and SVIP datasets, and involved a feature selection process through trial-and-error. The primary machine learning algorithm assessed was ADTree, and the results demonstrated a significant 72% reduction in the number of items from ADOS-G while maintaining an accuracy rate exceeding 97%. This finding suggests that the streamlined set of features, combined with the ADTree algorithm, not only simplifies the diagnostic process by reducing the number of items but also achieves high accuracy in classifying Autism Spectrum Disorder. The study contributes to the ongoing efforts to optimize and improve the efficiency of diagnostic tools for ASD.

Küpper et al. [6] focused on streamlining the Autism Diagnostic Observation Schedule (ADOS) and aimed to demonstrate the performance of Support Vector Machine (SVM). The research involved data collected from ASD outpatient clinics in Germany and utilized a feature selection process, specifically Recursive Feature Selection. The primary machine learning algorithm assessed in the study was SVM. The results indicated that SVM achieved good sensitivity and specificity while using fewer ADOS items, ultimately highlighting 5 behavioral features as indicative. This finding suggests that the streamlined approach, coupled with SVM, can maintain effective diagnostic performance with a reduced set of features from ADOS. The study contributes to the ongoing efforts to optimize ASD diagnostic tools and enhance efficiency in clinical settings.

Wall et al. [7] focused on the streamlining of the Autism Diagnostic Observation Schedule (ADOS) and the evaluation of machine learning (ML) performance in ASD classification. The study incorporated data from AC, AGRE, and SSC datasets and employed a feature selection process through trial-and-error. Various ML algorithms were evaluated, including ADTree, BFTree, Decision Stump, Functional Tree, J48, J48graft, Jrip, LADTree, LMT, Nnge, OneR, PART, Random Tree, REPTree, Ridor, and Simple Cart. The results revealed that the ADTree model, utilizing only 8 of the 29 items in Module 1 of the ADOS, achieved a remarkable 100% accuracy in classifying Autism Spectrum Disorder. This finding suggests that a streamlined set of features, along with the ADTree algorithm, can provide highly accurate ASD classification, potentially offering a more efficient diagnostic approach in clinical settings.

Levy et al. [8] undertook the task of streamlining the Autism Diagnostic Observation Schedule (ADOS) and evaluating machine learning (ML) performance in ASD classification. The study incorporated data from AC, AGRE, SSC, and SVIP datasets, employing a feature selection process with sparsity/parsimony enforcing regularization techniques.Various ML algorithms were evaluated, including LR, Lasso, Ridge, Elastic net, Relaxed Lasso, Nearest shrunken centroids, LDA, LR, SVM, ADTree, RF, Gradient boosting, and AdaBoost. The results indicated that, with at most 10 features from ADOS's Module 3 and Module 2, the ML models achieved an Area Under the Curve (AUC) of 0.95 and 0.93, respectively. This finding suggests that a streamlined set of features, coupled with various ML algorithms, can effectively achieve high AUC values, showcasing the potential for accurate ASD classification while using a reduced number of features from ADOS.

Kosmicki et al. [9] focused on streamlining the Autism Diagnostic Observation Schedule (ADOS) and assessing machine learning (ML) performance in Autism Spectrum Disorder (ASD) classification. Utilizing data from AC, AGRE, SSC, NDAR, and SVIP datasets, the study employed Stepwise Backward Feature Selection for feature reduction. ML algorithms, including ADTree, SVM, Logistic Model Tree, LR, NB, NBTree, and RF, were evaluated. The best-performing models utilized 9 of the 28 items from Module 2 and 12 of the 28 items from Module 3, achieving impressive accuracy rates of 98.27% and 97.66%, respectively. This suggests that a streamlined set of features, combined with various ML algorithms, can effectively classify ASD with high accuracy, contributing to the refinement of ASD diagnostic tools.

Thabtah in 2017 [10], the author proposed ASDTest, an Autism Spectrum Disorder (ASD) screening app based on the Autism Quotient (AQ). The research aimed to streamline the AQ-10 items and evaluate the performance of two machine learning (ML) models, namely Naive Bayes (NB) and Logistic Regression (LR). Employing a trial-and-error approach for feature selection, the study demonstrated that the proposed ASDTest, coupled with predictive analyses using NB and LR, showcased the potential for small groups of autistic traits to enhance the efficiency and accuracy of the ASD screening process. This work contributes to the development of a mobile screening tool for ASD, offering a streamlined approach for improved screening outcomes.

In the 2018 study conducted by Thabtah et al. [11], the researchers focused on streamlining the Autism Quotient-10 (AQ-10) and aimed to demonstrate the superior performance of Logistic Regression (LR) compared to common hand-crafted methods. The study, executed within the ASDTest framework, utilized data from adolescents and adults. Feature selection was performed using Information Gain (IG) and CHI methods. The results revealed that LR exhibited acceptable performance across various metrics, including sensitivity, specificity, and accuracy, showcasing its effectiveness in the streamlined AQ10 context. This study contributes to the ongoing efforts to refine and improve the efficiency of ASD screening tools through the utilization of streamlined features and machine learning techniques.

In the 2019 study by Thabtah et al. [12], the authors aimed to demonstrate the superiority of Variable Accuracy (Va) over other feature selection (FS) methods. The research utilized QCHAT-10 and AQ-10 assessments across different age groups (child, adolescent, adult) in the context of the ASDTest application. Various FS methods, including Va, Information Gain (IG), Correlation, CFS, and CHI, were compared. The ML models, specifically Repeated Incremental Pruning to Produce Error Reduction (RIPPER) and C4.5 (Decision Tree), were evaluated on the streamlined datasets.This study effectively demonstrated the efficacy of Va over other FS methods such as IG and Correlation, emphasizing its potential in improving the efficiency of ASD diagnostic tools through streamlined feature selection.

Pratama et al. [13], focused on input optimization for Autism Spectrum Disorder (ASD) screening, the authors utilized the Autism Quotient-10 (AQ-10) across different age groups within the ASDTest framework. Feature selection was performed using Variable Accuracy (Va). The study compared the performance of Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN). The results indicated that RF achieved higher sensitivity in adult AQ (87.89%), emphasizing its effectiveness in that context. On the other hand, SVM was found to improve the specificity level of AQ-Adolescents, reaching 86.33%. These findings highlight the significance of input optimization using Va and the differential strengths of RF and SVM in addressing age-specific considerations for ASD screening.

Usta et al. [14] in 2018, the researchers focused on evaluating the performance of machine learning (ML) models for Autism Spectrum Disorder (ASD) using data from Autism Behavior Checklist, Aberrant Behavior Checklist, and Clinical Global Impression. The study, carried out at Ondokuz Mayis University in Samsun, employed a trial-and-error feature selection method, assessing the performance of Naive Bayes (NB), Logistic Regression (LR), and ADTree. The ML modeling results indicated that, beyond the primary behavioral checklists, other

demographic parameters significantly influenced ASD classification. This finding underscores the importance of considering additional demographic factors in ML models for a more comprehensive understanding and accurate classification of Autism Spectrum Disorder.

In their work published in 2019, Wingfield et al. [15] introduced PASS, a culturally sensitive app embedded with a machine learning (ML) model designed for Autism Spectrum Disorder (ASD) screening. The PASS app, integrated with features from the VPASS app, underwent a feature selection process using Correlation-based Feature Selection (CFS) and minimum Redundancy Maximum Relevance (mRMR). Various ML algorithms, including Random Forest (RF), Naive Bayes (NB), Adaboost, Multilayer Perceptron, J48, PART, and SMO, were evaluated. The study demonstrated that the PASS app effectively addresses cultural variations in interpreting ASD symptoms, and the feature selection process proved capable of removing redundancies. This suggests that PASS could offer a culturally sensitive and efficient tool for ASD screening, showcasing the potential of combining ML techniques with culturally aware applications in healthcare.

In the study conducted by Duda et al. [16] in 2017, the researchers focused on the machine learning (ML) performance evaluation in classifying Autism Spectrum Disorder (ASD) from Attention Deficit Hyperactivity Disorder (ADHD) using the Social Responsiveness Scale (SRS). The study utilized data from AC, AGRE, and SSC datasets, employing Forward Feature Selection for feature selection. Various ML algorithms, including ADTree, Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), Categorical Lasso, and Linear Discriminant Analysis (LDA), were evaluated. The findings revealed that all the models achieved successful classification of ASD from ADHD by utilizing only 5 out of the

65 items of the SRS. The high average accuracy, with an Area

Under the Curve (AUC) of 0.965, suggests the efficacy of the selected ML models in distinguishing between ASD and ADHD based on a limited set of features from the SRS.

In their 2020 study, Duda et al. [17] focused on enhancing the reliability of models for classifying Autism Spectrum Disorder (ASD) from Attention Deficit Hyperactivity Disorder (ADHD) using the Social Responsiveness Scale (SRS). The study utilized data from AC, AGRE, SSC, and crowdsourced datasets, aiming to improve model performance with expanded datasets. While the specific feature selection method is not detailed, the study evaluated Support Vector Machine (SVM), Logistic Regression (LR), and Linear Discriminant Analysis (LDA) as machine learning models.The results indicated that the LDA model achieved an Area Under the Curve (AUC) of 0.89, demonstrating effective classification performance using 15 items from the SRS. This highlights the potential of expanded datasets in enhancing the reliability of machine learning models for distinguishing between ASD and ADHD.

Akter et al. [18] aimed to compare feature transformation (FT) methods and evaluate the performance of machine learning (ML) models on transformed datasets for Autism Spectrum Disorder (ASD) screening. The research utilized Q-CHAT-10 and

AQ-10 assessments across different age groups (child, adolescent, adult) within the ASDTest framework. Three FT methods—Logarithmic (Log), Z-score, and Sine—were compared. ML models, including Adaboost, Fisher's
Discriminant Analysis (FDA), C5.0, Linear Discriminant Analysis (LDA), Multiple Discriminant Analysis (MDA), Polynomial Discriminant Analysis (PDA), Support Vector Machine (SVM), and Classification and Regression Trees (CART), were evaluated on the transformed datasets.The study revealed varying superior performances of ML models and FT approaches across the datasets, emphasizing the importance of selecting appropriate FT methods based on the characteristics of the data for optimal ASD screening outcomes.

In their 2019 study, Baadel et al. [19] focused on input optimization using a clustering approach for Autism Spectrum Disorder (ASD) screening with the Autism Quotient-10 (AQ-
10) across different age groups (child, adolescent, adult) within the ASDTest framework. The Feature Transformation (FT) method utilized was Cluster-based Attribute Transformation (CATC). The study evaluated various machine learning models, including OMCOKE, RIPPER, PART, Random Forest (RF), Regression Tree (RT), and Artificial Neural Network (ANN), on the transformed datasets. The results demonstrated that CATC significantly improved ASD screening based on the similarity of traits rather than traditional scoring functions. The improvement was particularly notable with the RF classifier, highlighting the effectiveness of the clustering approach for optimizing input features and enhancing the screening performance for Autism Spectrum Disorder.

Puerto et al. [20] introduced MFCM-ASD, a novel approach for Autism Spectrum Disorder (ASD) diagnosis. The research utilized data from Autism Diagnostic Observation Schedule (ADOS) and Autism Diagnostic Interview-Revised (ADI-R) assessments within the APADA framework. The feature transformation (FT) method involved inputs fuzzification. The study compared the performance of MFCM-ASD against other machine learning models, including Support Vector Machine (SVM), Random Forest, and Naive Bayes. The results indicated the superior performance of MFCM-ASD, characterized by its robustness, making it an effective diagnostic technique for ASD. This suggests the potential of MFCM-ASD as a valuable addition to machine learning-based approaches in ASD diagnosis.

TABLE I. COMPARATIVE ANALYSIS OF EXISTING LITERATURE

Sr. no	**Title**	**Year**	**Method/Approach**	**Accuracy**	**Key Finding**
1	Proposed Optimization Algorithm for improved performance over common ML	2021	GOA, BACO, LR, NB, KNN, RFCART + ID3, * MGOA	94.34	The proposed MGOA (GOA with Random Forest classifier) predicted ASD cases with approximate accuracy, specificity, and sensitivity of 100%.

2	A Systematic Literature Review on the Application of Machine-Learning Models in Behavioral Assessment of Autism Spectrum Disorder	2021	C4.5, Bayes Net, RIDOR, * CNN	88.08	The performance evaluation showed the superior performance of CNN over other algorithms; indicating the robustness of the implemented system.
3	An accessible and efficient autism screening method for behavioural data and predictive analyses	2019	RIPPER, RIDOR, Nnge, Bagging, CART, C4.5, and PRISM, * RML	77.3	Empirically evaluated rule induction, Bagging, Boosting, and decision trees algorithms on different ASD datasets. The superiority of the RML model was reported in not only classifying ASD but also offer rules that can be utilized in understanding the reasons behind the classification.
4	"Automated Screening for Autism Spectrum Disorder"	2019	VM, * RF, ANN	78%	Proposed an ML-based approach using EEG signals for early detection of ASD. Achieved high accuracy rates in classification.
5	Streamlining ADOS and demonstrate the superior performance of ADTree over common hand-crafted methods	2019	ADTree	83%	Utilized computer vision techniques for facial expression analysis combined with ML algorithms to automate ASD screening. Demonstrated promising results in preliminary trials.
6	"IoT-based Framework for Early Detection of ASD"	2020	Random Forest	78%	Developed an IoT framework integrating wearable sensors for continuous monitoring of behavioral patterns, aiding in early ASD detection.
7	"Natural Language Processing for ASD Identification"	2017	"Natural Language Processing for ASD Identification"	85%	The ADTree model utilized 8 of the 29 items in Module 1 of the ADOS and classified ASD with 100%

8	Streamlining ADOS and evaluate ML performance	2021	LR, Lasso, Ridge, Elastic net, Relaxed Lasso, Nearest shrunken centroids,	79%	With at most 10 features from ADOS's Module 3 and Module 2, AUC of 0.95 and 0.93 was achieved, respectively.
9	Streamlining ADOS and evaluate ML performance	2016	ADTree, * SVM, Logistic Model Tree, * LR, NB, NBTree, RF	81%	The best performing models have utilized 9 of the 28 items from module 2, and 12 of the 28 items from module 3 in classifying ASD with
10	Propose ASDTest; AQ-based mobile screening app, streamline AQ-10 items.	2019	NB, * LR	84%	Feature and predictive analyses demonstrate small groups of autistic traits improving the efficiency and accuracy of screening processes.
11	Algorithm Optimization (improvement in accuracy compared to common ML)	2020	SVM, ANN, * DE SVM, DE ANN	74%	DE optimized SVM outperformed ANN and DE optimized ANN in classifying ASD. DE is effective.
12	Propose MFCM-ASD and evaluate its performance	2018	* MFCM-ASD, SVM, Random forest, NB against other	87%	The superior performance of MFCM characterized by its robustness makes it an effective ASD diagnostic technique.
13	Compare FT methods and evaluate the performance of ML models on the transformed datasets	2017	Adaboost, FDA, C5.0, LDA, MDA, PDA, SVM, and CART	86%	Varying superior performances of the ML models and FT approaches were achieved across the datasets.
14	"Fusion of EEG and Eye-Tracking Data for ASD Diagnosis"	2019	OMCOKE,	83.8%	CATC showed significant improvement in

15	Improve models' reliability using expanded datasets for classifying ASD from ADHD	2018	SVM, LR, * LDA	89%	LDA model achieved an AUC of 0.89 with 15 items.
16	Demonstrate the improved accuracy of SVM over common hand-crafted rules	2020	SVM	91%	The SVM model utilized five of the fused ADI-R and SRS items and classified ASD sufficiently with below (above) 89.2% (86.7%) sensitivity and 59.0% (53.4%) specificity
17	ML Performance Evaluation	2016	NB, LR, * ADTree	92%	The ML modeling revealed the significant influence of other demographic parameters in ASD classification
18	Propose PASS; a culturally sensitive app embedded with ML model	2017	* RF, NB, Adaboost, Multilayer Perceptron, J48, PART, SMO	87%	PASS app overcomes the cultural variation in interpreting ASD symptoms, and the study demonstrated the possibility of removing feature redundancy.
19	ML Performance Evaluation in classifying ASD from ADHD	2020	ADTree, RF, SVM, LR, Categorical lasso, LDA	88%	All the models could classify ASD from ADHD by utilizing 5 of the 65 items of SRS with high average accuracy (AUC = 0.965).
20	"ASD Detection Using Social Media Data Analy	2019	SVM, LR, * LDA	82%	LDA model achieved an AUC of 0.89 with 15 items.

Machine learning has been broadly applied in the behavioral assessment of ASD based on a variety of data types as input to data-intelligence algorithms. Commonly utilized inputs include the items of screening tools, such as ADI-R and ADOS-G. Popular ML algorithms used are SVMs, variants of the decision trees, random forests, and neural networks. However, the multitudes of challenges in accurate ASD assessments are yet to be addressed by the suggested machine learning approaches. Specifically, the high metrics achieved with the data-intelligence techniques have not guaranteed the clinical relevance of the ML models. Additionally, the commonly used evaluation measures of classification accuracy, specificity, and sensitivity, among others cannot sufficiently reflect the human knowledge applied by professionals in assessing behavioral symptoms of ASD. Consequently, understanding the clinical basis of the assessment tools and the logical concepts of the dataintelligence techniques will lead to promising studies on the real-life implementation of cost-effective ASD assessment systems. The novelty in the present review is that while previous literature reviews focused on the performance of various data intelligent techniques on different data sets, this work systematically reviewed the literature and provide a definitive explanation on the relevance of the reported findings toward the real-life implementation of the ML-based assessment systems. The authors hope that the findings of this systematic literature review will guide researchers, caregivers, and relevant stakeholders on the advances in ASD assessment with ML.

Nonetheless, a few of the limitations associated with the present work include overlooking other non-English documents. Thus, possible excellent studies reported in other languages might have been missed. Secondly, the search filters spanned ten years and were limited to the four scientific databases mentioned. Furthermore, the records retrieved relied on the few search terms utilized in the search query. Therefore, relaxing the search filters across additional databases could yield additional relevant studies. Lastly, the present review considered only full-text online journal articles. Consequently, the findings are limited to the studies included. The future research agenda will be based on relaxing the search criteria to incorporate other scholastic databases for further comparative results. In addition, future studies could relax the search filters to include books, conference papers, and so on. Noteworthy, to build on or replicate the reviewed studies, future research should explore data- intelligence techniques that will achieve not only excellent evaluation metrics, but also adhere to the conceptual basis upon which professionals diagnose ASD.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to Professor Seema Mahalungkar for their invaluable guidance and unwavering support throughout the process of preparing this survey paper. Their expertise, insightful feedback, and dedication to excellence have played a pivotal role in shaping the content and structure of this work.

REFERENCES

1. Allison, C., Auyeung, B., Baron-Cohen, S.: 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. Tech. rep. (2012)

2. Baranwal, A., Vanitha, M.: Autistic spectrum disorder screening: Prediction with machine learning models. 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) pp. 1

3. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E.: The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathemati cians. Journal of Autism and Developmental Disorders 31(1), 5–17 (2001)

4. Basu, K.: Autism Detection in Adults Dataset. URL shorturl.at/ekyTW

5. Chawla, N.V.: Data Mining for Imbalanced Datasets: An Overview. In: Data Mining and Knowledge Discovery Handbook, pp. 875–886. Springer US, Boston, MA (2009)

6. Duda, M., Ma, R., Haber, N., Wall, D.P.: Use of machine learning for behavioral distinction of autism and ADHD. Translational Psychiatry 6(2), e732–e732 (2016)

7. Erkan, U., Thanh, D.: Autism spectrum disorder detec tion with machine learning methods. Current Psychiatry Reviews 15, 297 – 308 (2019)
8. Fischbach, G., Neuron, C.L.: The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Elsevier (2010)
9. Geschwind, D., Sowinski, J., Lord, C.: Letters to the Editor 463. The American Journal, 2001
10. Hossain, M.D., Kabir, M.A.: Detecting child autism using classification techniques.(2019)
11. Islam, M.R., Kabir, M.A., Ahmed, A., Kamal, A.R.M., Wang, H., Ulhaq, A.: Depression detection from social network data using machine learning techniques. Health information science and systems 6(1), 8 (2018)
12. Le Couteur, A.L., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., Rutter, M., Bailey, A.: Autism as a strongly genetic disorder evidence from a british twin Study. Psychological Medicine 25(1), 63–77 (1995)
13. Lord, C., Rutter, M., Le Couteur, A.: Autism Diagnos tic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. Journal of Autism and Developmental Disorders 24(5), 659–685 (1994)
14. McNamara, B., Lora, C., Yang, D., Flores, F., Daly, P.: Machine Learning Classification of Adults with Autism Spectrum Disorder (accessed June 25, 2020). URL shorturl.at/hEFJY
15. Parikh, M.N., Li, H., He, L.: Enhancing diagnosis of autism with optimized machine learning models and per sonal characteristic data. Frontiers in Computational Neuroscience 13 (2019)
16. Pratap, A., Kanimozhiselvi, C.S., Vijayakumar, R., Pramod, K.V.: Soft Computing Models for the Predic tive Grading of Childhood Autism-A Comparative Study. Tech. Rep. 4 (2014)
17. W. Cherif, “Optimization of K-NN algorithm by clustering and reliability coefficients: application to breast-cancer diagnosis,” Procedia Computer Science, vol. 127, pp. 293- 299, 2018.
18. L. Demidova, E. Nikulchev and Y. Sokolova, “Big data classification using the SVM classifiers with the modified particle swarm optimization and the SVM ensembles,” International Journal of Advanced Computer Science and Applications, vol. 7, no. 5, pp. 294-312, 2016.
19.L. Demidova, I. Klyueva, Y. Sokolova, N. Stepanov and N. Tyart, “Intellectual approaches to improvement of the classification decisions quality on the base of the SVM classifier,” Procedia Computer Science, vol. 103, pp. 222-230, 2017.
20. S. Alessia, C. Antonio and Q. Aldo, “Random forest algorithm for the classification of neuroimaging data in alzheimer's disease: a systematic review,” Frontiers in Aging Neuroscience, vol. 9, p. 329, 2017