JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 503-510
https://publishoa.com
ISSN: 1309-3452

Advance Hybrid RF-GBC-RFE Wrapper-Based Feature Selection Techniques for Prediction of Autistic Disorder

C.Radhika1 , Dr N.Priya2
Shrimathi Devkunwar Nanalal Bhatt Vaishnav College for women,
University of Madras, Tamilnadu, India
Corresponding author's Email: radhamegha24@gmail.com

Abstract— Autistic disorder is a premature developmental ailments characterized by impaired societal interaction and persistent verbal exchange with stereotyped conduct. Detecting autistic ailments at an early stage is time consuming and very expensive. Machine learning classifiers play an imperative role in the early detection of autism spectrum disorders. The intention of this article is to make people aware of the early deduction of ASD in affected children. We provide a new hybrid technique to select the Feature-RF-GBC-RFE model in this work using the feature-based recursive feature elimination (RFE) ensemble of the Random Forest (RF) and the Gradient Boosting Classifier (GBC). Feature selection is a system that derives a subset of the perfect capabilities of a predictive modeling dataset. The feature in the ASD dataset is analysed and reduced by age category in this article. The hybrid RF-GBC-RFE feature selection technique, ML techniques such as Random Forest, Support Vector Machine, Gradient Boosting Classifier, and AdaBoost are used to study the reduced feature set. The model's overall performance can be categorized into precision and sensitivity metrics. A hybrid RF-GBC-RFE feature selection strategy is proposed in a unique way that improves data classification accuracy.

Index Terms— ASD, ML Techniques, RFE, Wrapper-based.

I. INTRODUCTION

Autism spectrum disorder, also known as ASD, is a neurological development condition that is frequently associated with expensive medical expenses and time-consuming tests. The early identification of characteristics associated with ASD can assist in slowing the evolution of the condition [1]. ASD is a gathering of neuro-developmental inabilities that are not treatable but rather might be improved by early intercessions [2]. ASD is a complicated, highly genetic disorder in which a number of natural factors interact with inherited features to raise the risk and lead to a variety of clinical manifestations and outcomes [3]. Mental imbalance range issues incorporate a gathering of neurodevelopmental anomalies with comprehensively differing degrees and signs, for the most part, start in youth, and described by issues in friendly correspondence and connection, alongside conduct issues, for example, limited interests and dreary behaviors [4]. The various types of Autism disorder are Asperger's disorder**:** A person with Asperger's strength be amazingly astute and prepared to manage their ordinary day by day presence. They might be truly centered around points that interest them and examine them relentlessly. In any case, they have significantly harder time within society [5][6][7]. PDD-NOS**:** This significant piece of examination included most adolescents whose mental imbalance was more genuine than Asperger's condition, however not as extreme as an autistic disorder [5][6][7]. Autistic ailment**:** This more established term is further along with the mental imbalance range than Asperger's and PDD-NOS. It fuses similar styles of appearances, anyhow at a extra outrageous stage [5][6][7]. Childhood disintegrative syndrome**:** This was the most uncommon and most extreme severe part of the spectrum. It portrayed children who develop customarily and later on rapidly lose several social, languages, and mental competencies, for the maximum part between a long time 2 and 4. Much of the time, these children additionally constructed up a seizure disorder [5]. Rett ailment: Kids with Rett circumstance regularly have rehearsed like autism, and experts cluster it amongst a variety of problems. However, for the reason that it is recognized to be added approximately via a hereditary transformation [5]. Autism (ASD) involves a collection of neurodevelopmental anomalies that start in youth albeit the principal finding may now and again happen at some point depicted by issues in correspondence and social behavior [8]. Computer-based intelligence can improve suggestive and intervention research within the sociologies, and may be mainly important in tests consisting of the outstandingly everyday and heterogeneous state of mental awkwardness range disorder [9].

II. RELATED WORK

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 503-510
https://publishoa.com
ISSN: 1309-3452

Machine Learning techniques assume an imperative part for prognosis ASD in the investigation of the dataset. This segment essentially centers around predict ASD utilizing different Machine Learning techniques. Shahamiri et al. [1] developed a unique new autism framework that replaces the standard scoring function in traditional screening procedures with deep learning techniques. The CNN classification method detects the possibility of psychologically unbalanced qualities in a variety of datasets quickly. Akter et al. [2] have implemented attribute selection and ranking strategies inside the different stages of ASD datasets and used diverse classifiers to research that converted statistics and determined significant functions that are extraordinarily predictive for ASD datasets. Padmapriya.S, et al. [10] have proposed a new relevant feature selection approach that carries the performance of the Chi-Squared characteristic choice and IG attribute selection methodologies to lessen the scale of the ASD dataset. Raj et al. [11] have proposed CNN primarily based version may be applied for the recognition of ASD turned into capable of attaining the maximum accuracy result than all of the other Machine learning strategies. Thabtah et al.[12] Variable psychoanalysis is a brand new computational intelligence
technique that evaluates characteristic-to-elegance correlations while reducing feature-to-function correlations. Modified Checklist for Autism in Toddlers is a generally utilized evaluating test for early recognition of children susceptible to emergent ASD [13, 14]. An Artificial intelligent will illustrate promising results in diagnosing ASD with the guide of lessen data dimensionality and picking the ideal and significant autism features ASD [15-20].

III. DATASET

The "ASD screening Data for child Dataset" was retrieved from the UCI repository [21] and is a dataset that has the potential to be utilised in the creation of an accurate predictive model. The AQ-10 is used to decide whether or not a person should be referred for an exhaustive autistic assessment. AQ-10 test addresses center around diverse areas, for example, precision, consideration exchanging, correspondence, creative intelligence, and societal connection. The scoring strategies for the investigation are that alone 1 point can be scored for every one of the 10 inquiries. Datasets of youngsters contain 292 occurrences separately. The datasets contain 21 features which are a blend of mathematical and straight out information, that incorporates: Age, Gender, Ethnicity, presence of Jaundice at the time of birth, Family part already with PDD, Relationships, Nation, Have you ever utilized the screening application before?, Screening tactic kind, Query1-10, and Category. ASD screening child dataset Feature descriptions and details of Feature mapping are presented in Table[I, II].

TABLE I
ASD screening Data for child Dataset descriptions[21]

Age	4-11 yrs
Instances	292
Clean Instance	248
Feature	21
No. of Male & Female	M-174 ,F-74
No. of Children Born Birth with Jaundice	61-Y, 187-N
No. of ASD with family	45-Y, 203-N
Class	Y-126, N-122

TABLE II
Details of variables mapping to the AQ_10 screening Tool [22][23][24]

Feature	ASD screening Data for child Dataset
AQ1	Does your youngster take a gander at you when you call his/her name?
AQ2	How simple is it for you to get an eye-to-eye connection with your kid?
AQ3	Does your kid highlight show that is/he needs something
AQ4	Does your kid highlight share interest with you?
AQ5	Does your youngster daydreamer?
AQ6	Does your kid looking in the same direction as you?
AQ7	Does your child show signs of fading to calm you or another family member who is clearly distressed?
AQ8	Would you depict your youngster's first vocabulary as:
AQ9	Does your youngster utilize basic motions?
AQ10	Does your kid stare at nothing with no obvious reason?
Era	Era in years
Sexuality	{m, f}
Ethnicity	List of nationalities
Hyperbilirubinemia	Regardless of whether the case was brought into the world with jaundice
The family with ASD	whether or not there is a genetic link to a PDD.
Lodging	List of nations
Associations	Parent, self, clinical staff, and so on,
App used before	{0, 1}
credit	0 through 10 are whole number values
Age_ desc	4 to >=18yrs
category	ASD attributes or No ASD characteristics

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 503-510
https://publishoa.com
ISSN: 1309-3452

IV. RESEARCH METHODOLOGY

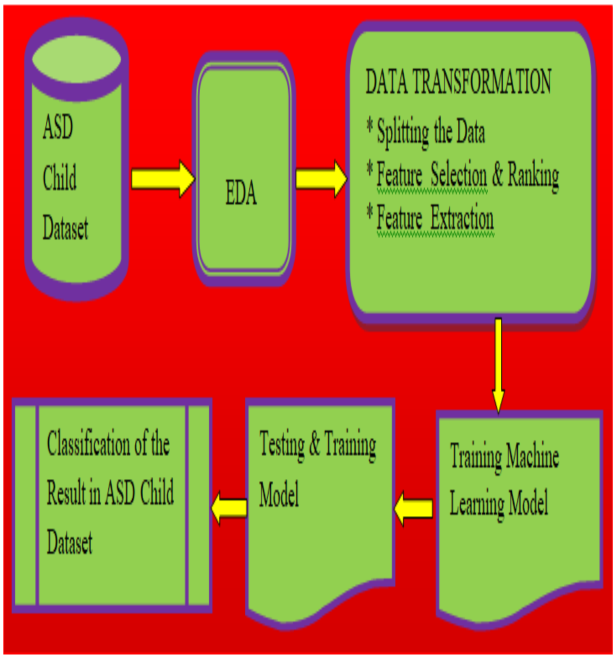

Fig. 1. The Framework Of The Proposed Methodology

In Machine Learning techniques, data pre-processing is a critical advancement. The proposed framework has been carried out with the persuasive pre-processing technique for Machine Learning strategies in identifying ASD for beginning phase recognition. The stream outline of the proposed procedure appears in Fig.1. The Machine Learning strategies for a prescient errand that incorporates SVM, Random Forest, Gradient Boosting Classifier, AdaBoost, with the Hybrid RF-GBC-RFE feature selection techhnique. These techniques are contrasted and diverse execution measurements and utilized for better dynamics.

A. Procedure for methodology

Step1: Load the ASD Child dataset.
Step2: Exploratory Data Analysis (EDA) by the way of Handle Missing value, Normalization & Scaling and apply One-Hot encoding.
Step3: Partition ASD child dataset into training and testing.
Step4: Implementing the Feature selection technique.
  Step4.1: Train Feature selection and Feature ranking list.
  Step4.2: Construct a model with Feature selection- GBC-RFE.
  Step4.3:Estimate Feature significance.
  Step4.4:Apply PCA for Feature extraction
Step5: Construct a machine learning Classifiers with Hybrid RF-GBC-RFE feature selection technique.
Step6: Performance evaluation and Metrics.
Step7: Select the effectual Machine Learning technique for the ASD child dataset.

B. Exploratory Data Analysis

The gathered information was orchestrated to eliminate insignificant attributes. For instance, the ID section was contemptuous to build up a forecast model; accordingly, it was taken out. In ASD screening Data for child Dataset, there are relatively few features they do presently that don't give any advantage of our evaluation. Results showed dropping Ethnicity, Residence, Relations, Utilize app before, Score, Age_desc, Class features would bring about more precise classification thus those sections were dropped. An outline of the incorporated datasets has appeared in Table II. There are insufficient values in the age, nationality, and relationships components in ASD databases. When we use the Handle Missing Value technique on the ASD dataset, all missing occurrences get a value of 'unknown.' We forecast contemporary realities to find requesting conditions and replies for knowledge classifiers after pre-preparing the missing instance. Following that, one-hot encoding strategies are used to generate dummy attributes for each instance of the genuine traits. ASD datasets must be split into test and training data after producing dummy _features. The data can be randomly assigned to training sets, with 30% of the data going to the testing set. In the scaler, various traits are changed over appropriately into standardized scale information. This proposed framework makes utilizes Standard Scaler () strategies for scaling information [25].

C. Feature selection based on Hybrid RF-GBC-RFE

Ensembles of diverse techniques in machine learning tactics may produce better overall performance than any single method. As a result, we can anticipate that an ensemble of RFE will outperform a single RFE. In this paper, we present a new hybrid-based feature extraction method that combines Random forest with GBC-RFE, a recursive feature elimination-based wrapper. The proposed technique is applied to the ASD dataset to determine the most significant features for predicting autistic traits. The exact outcomes establish that the proposed method selects significant features amongst other algorithms. The structural design of the proposed hybrid RF-GBC-RFE feature selection has appeared in fig 2. The most significant reduce final features set are shown in Table III.

TABLE III
Feature Significant of ASD Child Dataset[22][23][24]

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 503-510
https://publishoa.com
ISSN: 1309-3452

Model	Selected Feature Based On The Ranking
GBC-RFE	A1,A3,A6,A8,A10,A5
ABC-RFE	A2,A3,A5,A7,A9,A1
RF-RFE	A4,A3,A5,A8,A10,A6
SVM-RFE	A3,A6,A1,A10,A7,A5
HYBRID RF-GBC-RFE	A4,A5,A7,A9,A3,A6

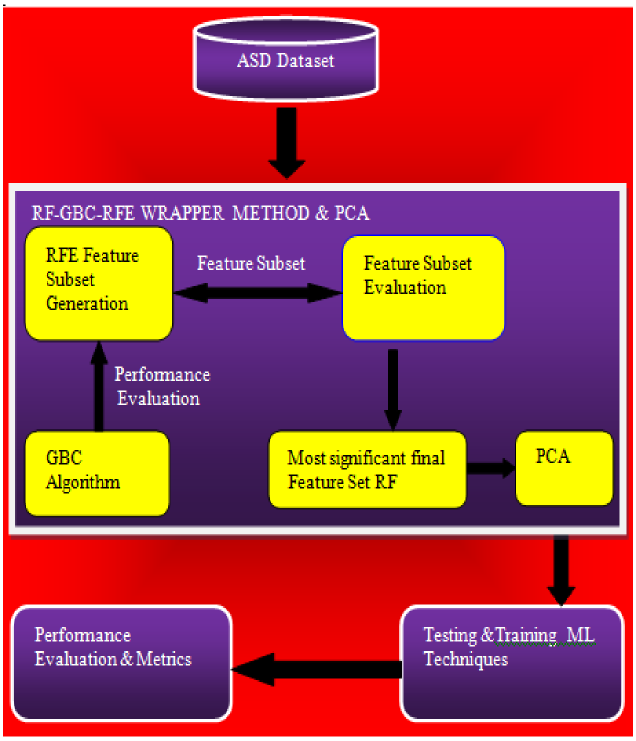

Fig. 2. Design outline of proposed Hybrid RF-GBC-RFE

V. COMPARISON OF DIFFERENT MACHINE LEARNING CLASSIFIER WITH HYBRID RF-GBC-RFE FEATURE SELECTION TECHNIQUE

A. Recursive Feature Elimination

In general, RFE is a wrapper-kind attribute selection technique that still makes use of filter-based attribute selection internally. This implies that a diverse ML technique is given and applied inside the center of the strategy, is wrapped by way of RFE, and used to assist select feature. In other words, RFE works via searching out a subset of attributes by starting with all of the attributes within the preparation dataset and effectively eliminating attributes prior to the ideal range of remaining parts[26,27].

B. RF-RFE

A random forest is a classifier comprising of an assortment of tree-structured classifiers { g(x,θh), k=1,...} where θh are indistinguishably disseminated arbitrary vectors and each tree makes a element preference for the maximum properly-liked class at input x [28]. The RF-RFE technique comprised of (a) consecutively RF to establish preliminary significance ratings, (b) eliminating the base 3% of factors with the least significance rankings from the informational collection, and (c) appointing positions to eliminated factors as per the request where they were taken out and their latest significance scores. This was performed iteratively utilizing the diminished informational collection made in sync two until 3% of the number of residual factors rounds to nothing [29, 30]. In this article performance evaluation metric for RF-RFE is exposed in Table IV.

C. *SVM-RFE*

The SVM-RFE is a type of rearward elimination strategy which begins with a absolute set of all elements, and afterward eliminates the maximum insignificant attributes individually. The pinnacle-ranked attributes eliminated in the final iteration of SVM-RFE are the maximum significant, even as the base positioned ones are the most informative and eliminated in the first iteration [26, 31-34]. SVM is an incredible classification strategy however it has no feature selection technique. Therefore, SVM-RFE generates a rating of features by means of figuring statistics gather for the duration of iterative in backward feature elimination. In each iterative, SVM-RFE sorts the attributes ineffective set within the request for a contrast of the objective capacities and eliminates an element with the base distinction. Characterizing IG(n) as data acquire when n-th include is eliminated[35,36]. In this study performance evaluation metric for SVM-RFE is revealed in Table IV.

D. *Gradient Boosting Classifier*

GBC-primarily based RFE (GBC-RFE) utilizes the gradient boosting strategies to prepare the classifier within the RFE technique. GBM makes use of boosting that is every other delegate ensemble approach. The proposal of boosting is to prepare a stable blend classifier with better precision from an outfit of feeble classifiers with decrease characterization exactness. GBA utilizes gradients inside the loss characteristic that is a movement demonstrating how exceptional the model's coefficients are at fitting the basic facts. The Gini list, which utilizes recurrence to assess the exactness of a tree-based calculation, is utilized to assess highlight significance (W_{gt}) in this examination. If the Gini list has a huge worth, it implies that the component is significant. The category variable is meant as r & f_i indicates the proportion of the number of perceptions in each elegance at a given node[29,37]. In this paper performance estimation metric for GBC-RFE is appeared in Table IV.

$$Wgt = 1 - \sum_{i}^{r} f^2 i \qquad (1)$$

E. Adaptive Boosting Classifier

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 503-510
https://publishoa.com
ISSN: 1309-3452

The Adaptive Boosting techniques by Freund and Schapire, Given weighed level, the outcomes are joined with numerous characterizations. Every component is weighed and another theory is learned in every emphasis and the attributes are weighed. The vital information on the AdaBoost method is to contain an allocation or set of weights upon the preparation set. The substance of the transmission on preparing data a_i on overweight wt is addressed as $D(a_i)wt$. Whole weights are instated. Be that as it may, on every cycle, the weights of dishonestly ordered examples are developed, and accordingly the weak learner is desired to accentuate on the preparation set. The feeble learner is to recognize a feeble hypothesis f(x) = {..1,..1}ti, reasonable for the appropriation ti D[39]. The weight of preparing information is instated as w1=1/no of elements. The miscalculation rate is determined by eqn2 and the weight can be modernized[38,39]. In this article Classification result for ABC-RFE is proven in Table IV.

$$err = \sum_{j=1}^{i} f(x) \qquad (2)$$

F. *Hybrid RF-GBC-RFE*

The goal of this research is to develop a new feature selection method called Hybrid RF-GBC-RFE, which combines RF, PCA, and GBC-RFE feature evaluation techniques. The pseudo-code of the Proposed Hybrid RF-GBC_RFE model can be expressed as Table V.

TABLE IV
Categorization result with ASD screening Data for child Dataset before implementation of the proposed model

Evaluation Measure	Gradient Boosting Classifier -RFE	Adaboosting Classifier - RFE	Random Forest Classifier - RFE	Support Vector Method - RFE
Accuracy	94%	83%	91%	89%
F1-rating	0.96	0.88	0.93	0.93
Meticulous ness	0.95	0.87	0.88	0.87
Recall	0.97	0.89	0.99	1.00
Error Rate	0.06	0.17	0.09	0.11

TABLE V
The pseudo-code of Proposed Hybrid RF-GBC-RFE

```
Data: ASD dataset.
Input: ASD Training dataset.
Output: Predicting ASD using the Feature selection method
based on ranking_score.
Begin
    ASD← {Set of ASD features}
    While( ASD* is not empty) do
        Train the Gradient Boosting Classifier using ASD*
        Compute the weight vector of the
                        GBC_ASD(W_ASD1,..,W_ASDn);
        Rank the features in ASD by
                        GBC_ASD(W^2_ASD1,..,W^2_ASDn);
        Find the bottom-ranked feature
        Eliminate the feature with the bottom-ranked feature
            (ASD*← ASD* - {the bottom rank feature})
        Fs ←ranking_feature
    End while
        AFs←Fs
   While (P is not empty) do
       Train AFs by RF with PCA
       Construct the Weight Vector of the
                        RF_AFs(W_AFs1,.., W_AFs n);
       Calculate the classification accuracy
   End while
End
```

VI. EXPERIMENT RESULT

In this segment, the performance evaluation metric for every strategy is acquired on the preferred features from the RFE approaches. To evaluate classification accuracy, we compared the preceding RFE approaches (SVM-RFE, RF-RFE, GBM-RFE, Adaboost-RFE) with the new Hybrid RF-GBC-RFE feature selection strategy. The validation of the proposed Hybrid RF-GBC-RFE feature selection approach in phrases of assessment metrics, prognostic accuracies, and diagnostic plot performance analysis in comparison with RF, SVM, GBC, AdaBoost machine learning algorithms are found to be propounding satisfying exposed in Table VII. In our proposed framework, the general performance evaluation of the classification utilizing every one of the twenty one attributes beforehand pre-preparing with 70 : 30 preparing and test datasets correspondingly appeared in TablE VI. The proposed Hybrid RF-GBC-RFE model is implemented to select among the twenty one best six, crucial features are established on basically dependent on interconnection, communication, facial features exposed in Table III. The ASD screening technique is a binary classifier dilemma given that persons are classified as ASD or Not ASD behavior. Thusly, implementation assessment systems that align with the binary class problem in ML techniques were utilized. The Classification results with ASD screening Data for child Dataset after implementation of Hybrid RF-GBC-RFE model of Test data have appeared in Table VI. The classification algorithm is implemented with k-fold cross-

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 1, 2022, p. 503-510
https://publishoa.com
ISSN: 1309-3452

validation to evaluate its overall performance, and the confusion matrix is used to determine stand-out evaluation measurements, such as miscalculation rate, accuracy, f1-rating, meticulousness, and recall, in order to record the overall performance of algorithms Table VI. A trial instance can be assigned to a predictable magnificence using the confusion matrix during the screening's class stage. As can be seen from the comparison effects, the new Hybrid RF-GBC-RFE model achieved all of the extra high precision while contrasting with different approaches as shown in Table VII and Fig. 3.

TABLE VI
Comparison of Different Machine Learning Classifier with proposed Hybrid RF-GBC-RFE Feature Selection Technique

Evaluation Measure	Gradient Boosting Classifier with hybrid RF-GBC-RFE	Adaboosting Classifier with hybrid RF-GBC-RFE	Random Forest Classifier with Hybrid RF-GBC-RFE	SVM with hybrid RF-GBC-RFE
Accuracy	97%	95%	99%	93%
F1-rating	0.99	0.96	0.99	0.91
meticulousness	1.00	0.93	0.98	0.94
Recall	0.98	0.98	1.00	0.88
Error Rate	0.03	0.05	0.01	0.07

TABLE VII
Comparison of accuracy with different feature selection technique.

MACHINE LEARNING CLASSIFIER	TRAINING	TESTING-30%
GBC with Hybrid RF-GBC-RFE	0.95	0.97
ABC with Hybrid RF-GBC-RFE	0.94	0.95
RF with Hybrid RF-GBC-RFE	0.94	0.99
SVM with Hybrid RF-GBC-RFE	0.91	0.93

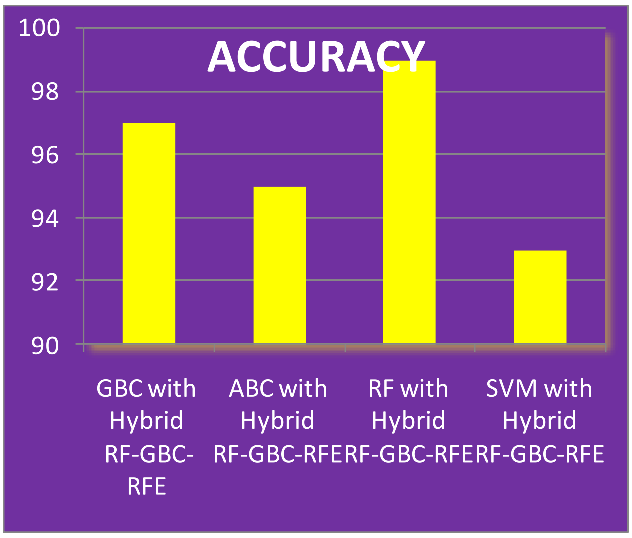

Fig. 3. Performance analysis of classifiers

VII. CONCLUSION

In this study, we have proposed the utilization of the Hybrid RF-GBC-RFE model as effective dimensionality reduction strategies for Feature selection and extraction. Feature determination in choosing a subset of the innovative Feature and Feature extraction to change the data onto another element subspace. Under the limited number of instances, it is important to choose a significant feature for preparing. The rule of the proposed arrangement of attributes choice is to get a minimal partition of attributes without decreasing the exactness of characterization. Initially, we followed the recursive feature elimination dependent on gradient boosting classifier (GBC-RFE) was received to choose the ideal element subset. At last, Feature extraction and random forest strategies were utilized to classify ASD datasets. This paper intended to give valuable and precise ASD screening models to help guardians and individuals rapidly analyze their youngsters' condition. Lamentably, a few families and grown-up patients don't have adequate information on ASD manifestations, so cases of ASD are not predicting early. Machine learning is utilized as of now in most living regions, and their utilization in the field of clinical conclusion adds to a spearheading step in utilizing the accessible information as an apparatus for advancement and progress. In the research of Autistic, indicative execution is indispensable and can be improved to characterize the specific sort of ASD precisely and cost-effectively. This can be refined differently, like expanding prescient precision, looking after affectability, specificity, and legitimacy, and lessening demonstrative time. In this exertion, detection of autistic was attempted utilizing different machine learning procedures. Experiment outcome show that random forest based on

GBC-RFE has the superior exhibition contrasted and different strategies. In our future work, we intend to work with various feature selections that incorporate element appraisal and class by and large for improved execution.

REFERENCES

[1] S.R. Shahamiri & F. Thabtah, "Autism AI: a New Autism Screening System Based on Artificial Intelligence," Cognitive Computation. DOI:10.1007/s12559-020-09743-3, issue 4/2020, Print ISSN:1866-9956, Electronic ISSN:1866-9964, 2020.

[2] Akter, T., Satu, M.S., Khan, M.I., Ali, M.H., Uddin, S., Lio, P., …Moni, M.A., "Machine Learning-Based Models for Early stage Detection of Autism Spectrum Disorders," IEEE Access, 1-1. DOI:10.1109/access.2019.295.2609, 2019.

[3] K. Francis, G. Karantanos, A. A1-Ozairi, S. Alkhadhari, "Prevention in Autism Spectrum Disorder: A Lifelong Focused Approach. Brain Sci," 2021, Jan24:11(2):151. DOI: 10.3390/brainsci11020151. PMID:33498888; PMCID:PMC7911370.

[4] SA. Hosseini, M. Molla, "Asperger Syndrome," 2021 Mar 6. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan- PMID:32491480.

[5] https://www.webmd.com

[6] Naji, Wafa & Qasim Waheeb, Assist. Prof. Mohammed. , "Autism Spectrum Disorder: Review Article," Medico-Legal Update, 20. 324-329, 2020.

[7] Walker, D.R., Thompson, A., Zwaigenbaum, L., Goldberg, J., Bryson, S. E., Mahoney, W.J .. … Szatmari, P., "Specifying PDD-NOS: A comparison of PDD-NOS, Asperger Syndrome, and Autism," Journal of the American Academy of Child & Adolescent Psychiatry, 43(2), 172-180. DOI:10.1097/00004583-200402000-00012, 2004.

[8] M. Samsam, "athophysiology of autism spectrum disorders: Revisitinggastrointestinal involvement and immune imbalance," World Journal of Gastroenterology, 20(29), 9942. DOI:10.3748/wjg.v20.i29.9942, 2014.

[9] D. Bone, MS. Goodwin, MP. Black, CC. Lee, K. Audhkhasi, S. Narayanan , "Applying machine learning to facilitate autism diagnostics: pitfalls and promises," J Autism Dev Disord. 2015 May, 45(5):1121-36. DOI:10.1007/s10803-014-2268-6.PMID:25294649; PMCID:PMC 4390409.

[10] S. Padmapriya, S. Murgan., "A Novel Feature Selection Methos for the Pre-Processing the ASD Dataset," International Journal of Pure and Applied Mathematics. 2018, ISSN(P): 1311-8080, ISSN(e): 1314-3395, Vol 118 No.8 pp.17-25, www.ijpam.eu.

[11] S. Raj & S. Masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques," Procedia Computer Science, 167, 994-1004. DOI: 10.1016/j.procs.2020.03.399, 2020.

[12] F. Thabtah, F. Kamalov, & K. Rajab, "A new computational intelligence approach to detect autistic features for autism screening," International Journal of Medical Informatics, 117, 112-124. DOI: 10.1016/j.ijmedinf. 2018.06.009, 2018.

[13] S. Kattimani, S. Sarkar , B. Bharadwaj, V. Ashvini, S. Mahadevan, "Early Detection of Autism Spectrum Disorders in Children with Attention Deficit Hyperactivity Disorder by Modified Checklist for Autism in Toddlers: a Pilot Study From India," J Compr Ped. 2014: 5(3):e21730. DOI:10.17795/compreped-21730.

[14] DL. Robins, D. Fein, ML. Bartton, JA. Green, "The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders," J Autism Dev Disord. 2001; 31(2): 131-44. DOI:10.1023/a:1010738829569. PMID:11450812.

[15] Rahman MM, Usman OL, Muniyandi RC, Sahran S, Mohamed S, razak RA, "A Review of Machine Learning methods of Feature Selection and Classification for Autism Spectrum Disorder," Brain Sci. 2020 Dec 7;10(12):949. DOI:10.3390/brainsci10120949. PMID:33297436; PMCID:PMC7762227.

[16] R. Vaishali & R. Sasikala, "A machine learning-based approached to classify autism with optimum behavior sets," Int. J. Eng. Technol. 2018, 7, 18. DOI:10.14419/ijet.v7i4.18.14907.

[17] Kosmicki JA, Sochat V, Duda M, Wall DP, "Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning," Transl Psychiatry. 2015 Feb 24; 5(2): e514. DOI:10.1038/tp.2015.7.PMID:25710120; PMCID: PMC4445756.

[18] S. Levy, M. Duda, N. Haber, & D.P. Wall, "Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism," Mol. Autism 2017, 8, 65. https://doi.org/10.1186/s13229-017-0180-6.

[19] Washington, P.; Paskov, K.M.; Kalantarian, H.; Stockham, N.; Voss, C.; Kline, A.; Patnaik, R.; Chrisman, B.; Varma, M.; Tariq, Q, "Feature selection and dimension reduction of social autism data," Pac. Symp. Biocomput. 2020,25,707-718.

[20] R. Alzubi, N. Ramzan, H. Alzoubi, "Hybrid feature selection method for autism spectrum disorder SNPs," In Proceedings of the 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology(CIBCB), INNSIDE, Manchester, UK, 23-25 August 2019.

[21] UCI Machine Learning Repository. Retrieved http:// Archive.Ics.Uci.Edu/ML/Index.php.

[22] Thabtah F., "Autism Spectrum Disorder Screening: Machine Learning adaptation and DSM-5

fulfillment," ICMHI 17 Proceedings of the 1st International Conference on Medical and Health Informatics. DOI:10.1145/3107514.3107515, 2017.
[23] Thabtah F., "An accessible and efficient autism screening method for behavioral data and predictive analyses," Health Informatics J. 25(4): 1739-1755. DOi: 10.1177/1460458218796636, 2019.
[24] Thabtah F. ASDTests. A mobile app for ASD screening. [Internet]. 2017[cited 2018 Dec 20]. Available from:www.asdtests.com.
[25] Sebastian Raschka, "Python Machine Learning," Sep 2015, ISBN 978-1-78355-513-0, www.packtpub.com
[26] I. Guyon, J. Weston, S. Barnhill, &V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines," Machine Learning, 46(1/3), 389-422. DOI:10.1023/A:1012487302797, 2002.
[27] P.M. Granitto, C. Furlanello, F. Biasioli, & F. Gasperi, "Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products," Chemometrics and Intelligent Laboratory Systems, 83(2), 83-90. DOI:10.1016/j.chemolab.2006.01.007, 2006.
[28] L. Breiman, random Forests. Mach. Learn. 2001,45, 5-32.
[29] H. Jeon, S. Oh, "Hybrid-Recursive Feature Elimination for Efficient Feature Selection," Appl. Sci. 2020,10,3211. https://doi.org/10.3390/app10093211
[30] B.F. Darst, K.C. Malecki, & C.D. Engelman, "Using recursive feature elimination in the random forest to account for correlated variables in high dimensional data," BMC Genetics, 19(S1). DOI:10.1186/s12863-018-0633-8, 2018.
[31] C. Wang, Z. Xiao, B. Wang, & J. Wu, "Identification of Autism Based on SVM-RFE and Stacked Sparse Auto-Encoder," IEEE Access, 7, 118030-118036. DOI:10.1109/access.2019.2936639, 2019.
[32] X. Lin, C.Li, Y. Zhang, B. Su, M. Fan, and H. Wei, "Selecting feature subsets based on SVM-RFE and the overlapping ratio with applications in bioinformatics," Molecules, vol. 23, p. 52, Dec.2017.
[33] J. Ding, J. Shi, and F.-X. Wu, "SVM-RFE based feature selection for tandem mass spectrum quality assessment," Int. J. Data Minin Bioinf., vol. 5, pp.73-88, Feb. 2011.
[34] X. Ding, Y. Yang, E. A. Stein, and T.J. Ross, "Multivariate classification of smokers and nonsmokers using SVM-RFE on structural MRI images," Hum. Brain Mapping, vol. 36, pp. 4869-4879, Dec. 2015.
[35] S.Yoon, S. Kim, "Adaboost-based multiple SVM-RFE for classification of mammograms in DDSM," BMC Med Inform DecisMak 9, S1 ,2009. https://doi.org/10.1186/1472-6947-9-S1-S1.
[36] Y. Lecun, JS. Denker, SA Solla, "Optimal Brain Damage," In advances in Neural Information processing systems Morgan Kaufmann; I990:598-605.
[37] X. Lin, X. Zhang and X. Xu, "Efficient Classification of Hot Spots and Hub Protein Interfaces by Recursive Feature Elimination and Gradient Boosting," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol.17, no.5, pp.1525-1534, 1 Sept.-Oct. 2020, DOI:10.1109/TCBB.2019.2931717.
[38] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," in Proc. Of the Thirteenth Int. Conf. on Machine Learning, 1996.
[39] Y. Freund and R. Shapire, "A Short introduction to boosting," journal of Japanese Society for Artificial Intelligence 14(5)(September 1999) 771-780.