I.J. Modern Education and Computer Science, **2023, 6, 73-89**
Published Online on December 8, 2023 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2023.06.06

Optimized Feature Selection and Transformations for Early Stage Prediction of Autism Using Supervised Machine Learning Models

Praveena K N*
Assistant Professor, Bio-intelligence Lab, Department of Computer Science and Engineering, PresidencyUniversity, Itkalpur, Rajanukunte, Bengaluru
Email: praveenakn1988@gmail.com
ORCID iD: https://orcid.org/0000-0003-3047-8729
*Corresponding Author

Mahalakshmi R
Professor, Bio-intelligence Lab, Department of Computer Science and Engineering, Presidency University, Itkalpur, Rajanukunte, Bengaluru
Email: mahalakshmi@presidencyuniversity.in
ORCID iD: https://orcid.org/0000-0001-9368-7224

Manjunath C
Assistant Professor, School of Mechanical Engineering, REVA University, Yelahanka, Bengaluru
Email: manjunath.c@reva.edu.in
ORCID iD: https://orcid.org/0000-0003-1935-1030

Ahmad Faiz Zubair
Senior Lecturer, School of Mechanical Engineering, College of Engineering, Universiti Teknologi Mara, Kampus Pulau Pinang, 13500 Permatang Pauh, Pulau Pinang, Malaysia
Email: ahmadfaiz@uitm.edu.my
ORCID iD: https://orcid.org/0000-0001-6524-7299

P. Karthikeyan
Post Doctor Researcher, Dept. of Computer Science and Information Engineering, National Chung Cheng University, Chiayi, Taiwan-62102
Email: karthi@ccu.edu.tw
ORCID iD: https://orcid.org/0000-0001-8977-5520

Received: 02 January, 2023; Revised: 25 March, 2023; Accepted: 25 May, 2023; Published: 08 December, 2023

Abstract: Autism Spectrum Disorder (ASD) is a neurodevelopmental syndrome which cannot be curable but can be predicted in early stage. Early prediction and cure may help to diagnose the autism. In existing methods, prediction of best feature is not identified for detecting the autism in early stage. In this proposed research, prediction of ASD has been done by identifying the best feature transformation technique with best ML classifier and finding out the most significant feature for diagnosis of autism in early age. Early-detected ASD datasets pertaining to toddler and child are collected and applied few Feature transformation techniques, comprising log, power-box-cox and yeo-Johnson transformations to these datasets. Then, using these ASD datasets, several classification approaches were applied, and their efficiency was evaluated. Adaboost given 100% accuracy for toddler dataset and whereas, Random forest showed 98.3% accuracy for child datasets. The feature transformations ensuing the best prediction was Log, Power- Box cox and Yeo-Johnson Transformation for toddler and Log transformation for children datasets. After these exploration, various feature selection techniques like univariate (UNI) and recursive feature elimination (RFE) are applied to these transformed datasets to recognize the most significant ASD risk feature to predict the autism in early stage for toddler and child data. It is found that A5 feature is most significant feature for toddler, A4 stands most significant feature for child based on univariate and RFE. This benefits the doctor to provide the suitable diagnosis in their early stage of life. The results of these logical methodologies show that ML methods can yield precise predictions of ASD when they are accurately optimised. This shows that using these models for early ASD detection may be feasible.

Index Terms: Autism, AQ-10 dataset, ML algorithms, Feature transformation, Feature selection technique, predictive model**.**

1. Introduction

Autism Spectrum disorder (ASD) is a neurological developing syndrome [1,2]. Communication between individuals is impacted by ASD, and autistic individuals rarely communicate with others. Usually the symptoms will occur during their childhood. This disorder will be there for life long and will not be curable. A study says that around 33% of kids with difficulties except ASD have certain common signs of ASD which leads to wrong prediction. It is helpful for the doctors to detect ASD in early stage so that the patients can get suitable treatment or medication required thus reduces the long term cost for diagnosis. The very young children and poor in Communication not able to answer properly. This may lead to score around 25% of total ADI-R tests [2, 3]. So the doctors are in need of efficient ASD screening methods with less time consuming and easily available ASD screening methods which can predict whether the person is having ASD and should convey them to take further diagnosis [4]. Currently the ASD datasets are very less and connected with experimental analysis which is generally inherited in environment, e.g., AGRE [5, 6], National Database of Autism Research (NDAR) [5-7] and Boston Autism Consortium (AC) [5, 8]. At present machine learning is used to detect many psychological syndromes like depression, ASD etc. The main intension of using machine learning is to enhance analysis accurateness and decrease diagnosing period [9]. In the meantime, the diagnosis process of the case consists of exact class (ASD, Not ASD) depends on data available, so this method can be considered as a predictive process in machine learning [10].

The drawbacks of existing methods are,

1 The main drawback of existing method is not identifying the important feature for prediction of ASD in early stage.
2 To overcome this, in our proposed research, prediction of ASD has been done by identifying the best feature transformation technique with best ML classifier and finding out the most significant feature for prediction of autism in early age.

The major research objectives are:

1 Early stage prediction of Autism for toddler and child datasets.
2 Three different feature transformation (FT) methods are applied for these datasets in order to boost our model performance. And then, ML algorithms are applied.
3 Then, two feature selection techniques [FST] are applied to find the best significant features in toddler and child, so that it will be helpful for the doctors to diagnose the disorder in early stage.

Therefore, this work represents that combination of Feature transformation and ML [10] algorithms can be used to find the ASD risk factors and FST's are used to find out the most significant feature for predicting the autism in early stage [11, 12, 13].

The purpose of applying machine learning classification algorithms is to achieve better-quality precision, recall and accuracy. AQ 10 dataset is chosen for our analysis and classification of ASD. From this dataset, we have found few points as follows:

1. Analysing features of Toddler and child datasets.
2. Applied Feature Transformation techniques before passing into the classifier.
3. Comparing other classification algorithms and identifying best classifier which is suitable for autistic datasets.
4. Finding the optimal Feature Transformation technique and classifier for toddler and child dataset.
5. Show the most significant feature using Feature Transformation techniques.

This paper is planned as follows: Section 2 provides us the related work, section 3 represents materials and methods used. In section 3, the description of dataset and exploratory analysis of ASD were described and it tells about the analysis of classification algorithms and also described the Feature transformation techniques and also discussed the FST's. Comparing the results which is obtained using various algorithms is analysed in section 4. Finally, paper is concluded with some future workings in section 5.

2. Related Work

To advance the diagnosis of ASD, numerous studies have incorporated machine learning techniques. The primary motivation for the use of ML models on ASD is to decrease recognition times, enabling quicker access to medical care administrations and an increase in analysis accuracy.

T Akter et al. [1] described and analysed AQ10 dataset for toddler, child, adult and adolescent by applying various Classifiers and different Feature transformation (FT) techniques. And finally analysed the best classifier based on different FT's.

Allison et al. [2] described a small measureable worksheet that can be used at numerous phases of the patient's life, including toddlers, kids, youths and adults.

Thabtah et al.[3] described A smartphone application called ASDTests, and established on the Q-CHAT and AQ-10 instruments, aids in the early diagnosis of ASD [1]. Additionally, they collected ASD information by means of these mobile application and updated it as an open source dataset to Kaggle and UCI repository.

Thabtah and Peebles [4] explained many investigations to distinguish and analyse ASD using an assortment of ML techniques. To survey the ASD properties, suggested a Rules-based ML (RML).

Satu et al.[5] established separate large highlights of typical and autism youngsters in Bangladesh using Decision tree [1].

Abbaset et al. [6] described about how to address the difficulties of scarcity, information loss, and shortage, so he combined the ADI-R and ADOS ML [1] approaches into a particular assignment.

Fadi Thabtah et al. [7] proposed Support vector machine (SVM) [1], Decision tree (DT) and Logistic regression (LR) were used to propose a computational insight (CI) technique known as Variable Analysis (VA), which revealed feature to-class and feature-to-feature correlations.

Fadi Thabtah et al. [8] explained about ASD screening tools, using DSM-4.

K. C. Howlader et al.[9] described about the how decision tree is used for classifying the diabetes disorder and found the best decision tree.

Ali Hossain et al.[10] described how machine learning models are used to identify the gene expression patterns for ovary cancer.

Duda et al. [11] explained about the classifiers and found that only five out of the 65 traits were adequate to differentiate autism with attention deficit hyperactivity disorder (ADHD)[1].

Gohet et al. [12] divided patients into normally established and autism spectrum disorder, where a correlation-based feature selection (CFS) [1] was used to assess the significance of features.

Crippa et al. [13] explained the analysis of ASD and TD children in 2015 led to the identification of 15 kindergarten-aged children with ASD utilising seven features. Despite this, they hypothesised that cluster analysis could more effectively capture intricate factors for prediction of ASD phenotype and heterogeneity [1].

References [14-17] described about the dataset used in this research.

Yao Zhang et al. [18] this paper described about the logarithmic transformation [LT]. LT is used to reduce the skewness in the data.

Yanli Liu et. al [19] introduced about random forest algorithm. It is the Machine Learning algorithm used for classification and prediction.

Wang [20] and Md Rafiqul Islam [28] explained about how K- Nearest Neighbour algorithm is used to classify the data.

Bujlow et al. [21] described about C5 algorithm which are used for classification of network traffic using machine learning.

Baoshan Ma et al. [22] explained about Extreme gradient boosting algorithm (XGBoost) for classification of camcers.

Satu et al. [23] described about the classification of sites of protein using machine learning algorithms like logistic regression, random forest, and artificial neural networks.

Praveena K N at al. [24] explained about classification of Autism for eye gaze pattern of dataset using convolutional neural networks.

Satu et al. [25] analysed about the traffic accident patterns for national highways using decision tree algorithm.

Hossain et al. [27] described the comprehensive review for disease prediction using ML algorithms.

Praveena K N et al. [29] explained the overview about autism disease based on supervised machine learning algorithms.

K.S Oma et al. [30], Omar et al. [31] described on machine learning approach for prediction of autism.

H Talabani et al. [32] explained about Support vector machine for classification of autism with respect to child dataset.

Fadi Thabtah [33] proposed ASDTests app can be used by doctors to help the patients and tell them about their diagnosis.

This study obtained ASD datasets from the UCI ML repositories that were related to research on ASD characteristics in toddlers and children [1, 4]. From References [5-11] I have learnt how ML algorithms are used for classification. These datasets were subjected to a variety of feature transformation (FT) procedures, which transformed them into a format appropriate for these studies which is referred in reference 18. Later, several classifiers were applied to these distinct datasets, and we were able to identify high-performing ML techniques in papers 19 to 33. We also looked into the impact of information change on classifier display. After these exploration, various feature selection techniques like univariate (UNI) and recursive feature elimination (RFE) are applied to these transformed datasets to recognize the most significant ASD risk feature to predict the autism in early stage for toddler and child data. These results recommend that ML could be assessed to detect ASD risk factors. Furthermore, we identified the top machine learning (ML) simulations to investigate the predictive risk elements of ASD, discovering that a few ML techniques achieved well for various datasets we used in this research.

3. Materials and Methods

The overall methodology, which is broken up into four modules as shown in Fig. 1, will be briefly described in this section. Fig.1 represents the proposed model of our methodology for predicting ASD cases. Firstly, toddler and child datasets are collected, in detail is given below section i.e. Data collection. This helps us in achieving our objective 1.

3.1 Data collection

Thabtah et al.'s screening and risk factor identification methods for autism. ASDTests program utilized Q-Chat 10 and AQ-10 Apparatuses (Baby, Kid). A positive expectation of ASD is indicated by a final individual score of at least 6 out of 10 on the scale of 0 to 10. The value of each item ranges from 1 to 10. Using ASDTests to sum the datasets, we gathered total 2009 records from repositories [14-17]. These included the following datasets for:

i) Toddler has total number of people i.e. N=1054. In which 319 female around 30.26% and 735 male around 69.73%.
ii) Children has total number of people i.e. N=248. In which 74 female around 29.83% and 174 male around 70.16%.

The feature descriptions of the various datasets used in this study are presented in Table 1 and 2.

Table 1. ASD Dataset Description

Feature	**Data Type**	**Explanation**
Age	Number	years and months of age
Gender	String	Men or Women
Ethnicity	String	Text-based list of common societies
Born with Jaundice	Boolean (yes or no)	Does the individual have jaundice?
Family member with ASD	Boolean (yes or no)	It states that anyone in the family who was not a parent, a caregiver, a clinician, etc. had autism before.
Who completing the test (User)	String	The person may provide a succinct justification for doing the assignment.
Why taken the screening	Meta	Answer is binary.
Used App Before	Boolean (yes or no)	The consumer will provide a description in their native language.
Language Spoken	Boolean (yes or no)	common societies listed in text
Country of residence	String	a text-based list of nations
Used the screening app before	Boolean (yes or no)	What screening apps the user has used
Screening Method Type	Int (0,1,2,3)	Based on age categories (0 = toddler, 1 = child), the type of screening methods is selected.
A1 (Response of Q1)	Binary (0, 1)	See Table 2 for more detail on Q1
A2 (Response of Q2)	Binary (0, 1)	See Table 2 for more detail on Q2
A3 (Response of Q3)	Binary (0, 1)	See Table 2 for more detail on Q3
A4 (Response of Q4)	Binary (0, 1)	See Table 2 for more detail on Q4
A5 (Response of Q5)	Binary (0, 1)	See Table 2 for more detail on Q5
A6 (Response of Q6)	Binary (0, 1)	See Table 2 for more detail on Q6
A7 (Response of Q7)	Binary (0, 1)	See Table 2 for more detail on Q7
A8 (Response of Q8)	Binary (0, 1)	See Table 2 for more detail on Q8
A9 (Response of Q9)	Binary (0, 1)	See Table 2 for more detail on Q9
A10 (Response of Q10)	Binary (0, 1)	See Table 2 for more detail on Q10
Scoring Results	Integer	See Table 2 for more details
ASD	Boolean	Toddlers and child diagnosed with ASD

Table 2. Details of Variables mapping

Variable in Dataset	QCHAT-10 Toddler Features (18-36 Months)	AQ-10-Child Features (4-11 Years)
Q1	Does your child face you when you call him or her by name?	He or she frequently hears tiny noises that others do not.
Q2	How easy do you find it to look your child in the eye?	In most cases, they are more concerned with the big picture than with the nitty-gritty.
Q3	Does your child point while expressing a desire for something? (For instance, an out-of-reach toy)	He or she can readily follow the conversations of multiple persons in a social group.
Q4	Does your youngster show signs of sharing your interests? (For instance, pointing to a captivating scene)	He or she has no trouble switching back and forth between activities.
Q5	Does your kid act out?	He or she struggles to maintain a discussion with their Peers.
Q6	Does your youngster follow where you're looking when you do things like take care of dolls or use a toy phone?	He or she is skilled at small talk.
Q7	Is your child wandering around when you or another member of your family is clearly upset? Ex: giving them a hug and petting their hair)	When reading a story, the individual finds it challenging to figure out the character's thoughts or emotions
Q8	Does the phrase "does your child employ basic gestures" appear in your child's first words? Does your child stare aimlessly and blankly when you wave goodbye, for instance?	He or she liked to play games when they were in preschool.
Q9	Does your kid act out?	includes playing pretend with other kids
Q10	When you take care of dolls or use a toy phone, does your child follow where you are looking?	He or she finds it simple to determine what someone is feeling or thinking.

3.2 Data Exploration and Data Pre-processing:

Our objective is to classify the ASD using different machine learning classification algorithms and at last finding the optimal performing algorithm and optimal Feature transform technique which is suitable for this dataset. Since, this

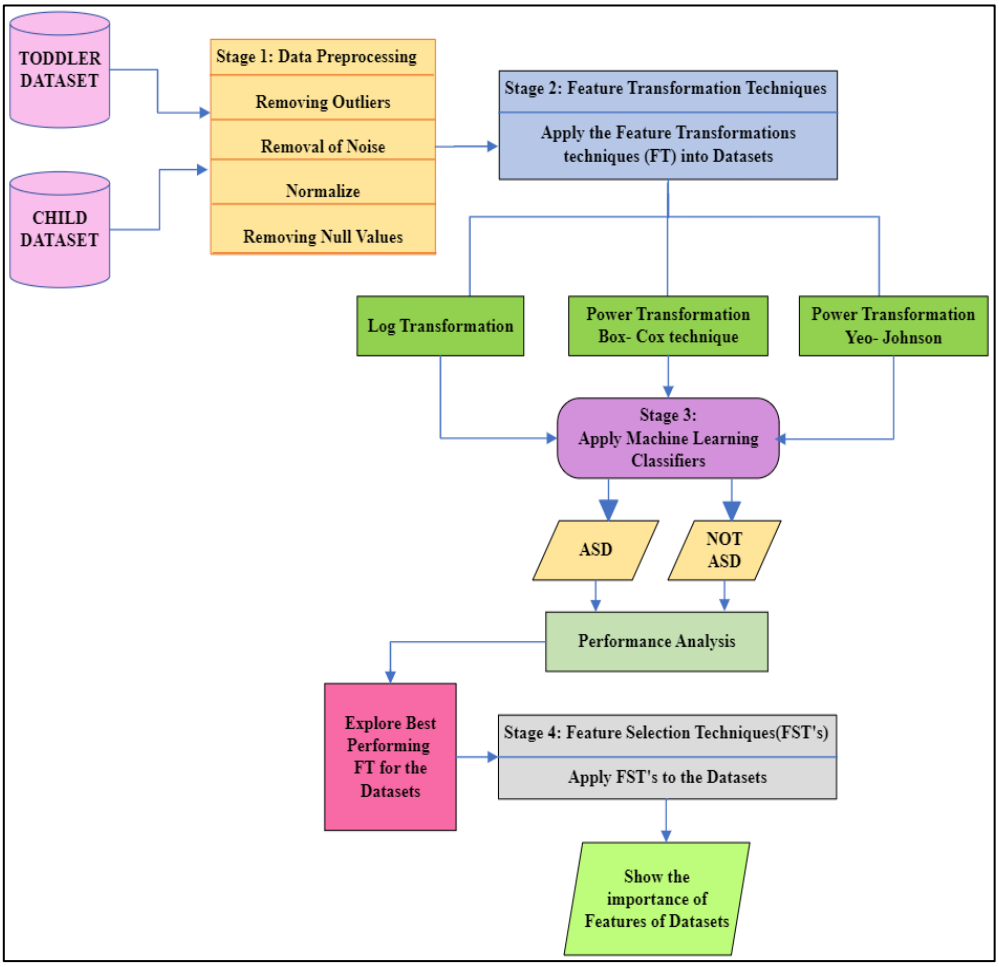


Fig.1. Proposed model to predict ASD at early stage

research is mainly focused on early prediction of autism therefore, the ASD dataset for toddler and children datasets are analysed by considering the accuracy, recall, precision, and F-measure metrics. Data pre-processing has been done for the dataset by eliminating the attributes that have misplaced qualities and furthermore eliminated the features which do not offer any advantage during the analysis. Since the dataset is categorical, we have applied pre-processing techniques such as Label Encoder. In order to make the labels machine readable, it is utilised to transform the labels into numeric data. Following that, machine learning algorithms can choose the optimum method for processing these labels [16].

3.3 Feature Transformation Technique:

After encoding, various Feature Transformation techniques were applied to decrease skewness, make equivalence, linearity and for the feature that have zero's and negative values. Some common methods were applied for these datasets like Log Transformation, Power Transformation i.e. Box-Cox method and Yeo-Johnson. These methods are used to achieve objective 2.

The details of these transformation is given in the Table 3.

Table 3. Brief Description of different Feature Transformations (FT)

FT's	Details	Mathematical Representation
Logarithmic	It converts excessively skewed density into a near Gaussian density. [18]	$y = c\ log_b(1+x)$
Power Transformation (Box-Cox) [22]	The transformation of non-normal dependent variables into their normal form is known as a Box Cox transformation [22].	$y(\lambda) = \frac{y^{\lambda}-1}{\lambda}, if\ \lambda \neq 0$
Power Transformation (Yeo-Johnson) [26]	This transformation allows for zero and negative values of y [26]. λ can be any real number.	$y^{(\lambda)}{}_i = \frac{\left(y_i \pm 1^{\lambda-1}\right)}{\lambda}, if\ \lambda \neq 0, y \geq 0$ $y^{(\lambda)}{}_i = \log(y_i \pm 1), if\ \lambda = 0, y >= 0$

3.4 Machine Learning Classifiers

Machine Learning algorithms are utilized for classification of autism. So, total 6 classifiers are used and in that Random forest showed 98.3% accuracy for child and Adaboost has given 100% for toddler datasets. After feature transformation ML classifiers were applied for these transformed datasets which is used to achieve objective2. Out of which Logistic regression does not showing any changes in any Feature Transformations (FTs) Therefore LR is not considered for this dataset. The remaining classifiers that we have considered are Random Forest, K-Nearest Neighbour, Support vector machine, Decision tree [31], Gradient Boosting and Adaboost. Fig.1 represents the steps to analyse the dataset.

Brief description on different classifiers are signified as follows:

- **Random Forest (RF):** A random forest is a tree-structured classifier using the following formula: h(x, Θk), where k = 1, 2, etc.

Θk -> independent random vectors with identical distributions [7]

x-> feed data

After k times running, the classifier sequence is obtained as follows $\{h_1(x), h_2(x), \ldots. h_k(x)\}$ which constitutes more than one classification models [18][30]. Therefore the final result of this RF is given in equation (1):

$$H(x) = \arg\max \sum_{i=1}^{k} I\left(h_i(x) = Y\right) \quad (1)$$

Where:

H(x) is combination of classification model,

h_i is a single decision tree model,

Y is the output variable,

I() is an indicator function [15].

It is made up of many Decision Trees and operates as an ensemble. An ensemble contains a group of models that are used to predict the result, instead of specific model. In random forests, each decision tree calculates a class result and the class result with the maximum amount of polls becomes the calculation of random forests. For accurate calculations, the Decision Trees should be least correlated with each other.

- **K nearest neighbour algorithm (KNN):** It uses nearest neighbour methodology for prediction by assuming that objects of neighbours have same predictive data [29]. The basic idea behind the nearest neighbour algorithm is to locate the k points in the multidimensional space Rn that are nearest to the unidentified sample, and then to classify the unexplained sample according to the groups of the k points [4]. These k sites are the k closest neighbours of the unidentified samples. The algorithm assumes that each example relates to a point in

a multidimensional space. Based on the usual Euclidean distance [20,4], the closest neighbour of an example is identified. Let x's eigenvector is, where ar(x) signifies the instance x's rth characteristic value. Equation (2) states that the distance amongst the x_i and x_j is defined as d (x_i, x_j):

$$d\left(x_i, x_j\right) = \sqrt{\sum_{r=1}^{n}\left(a_r\left(x_i\right) - a_r\left(x_j\right)\right)^2} \tag{2}$$

The distinct entity classification function in KNN is represented by the formula f: $R_n \to S$, where S is a finite set of different category sets (s1, s2,... ss). According to the quantity and degree of dispersion in each type of sample, the nearest neighbour k value is selected, and alternative k values can be proposed for different applications.

- **Support vector machine (SVM):** Using supervised learning and a kernel-based classifier, the SVM [9] classifier divides the data into two or more categories. For binary classification, SVM is primarily intended. SVM creates a model, maps each class's decision boundary, and specifies the hyper plane that divides the classes that are not similar during the training phase. The accuracy of classification can be enhanced by increasing the distance between classes by raising the hyper plane margin.
- **Decision Tree (C5.0):** Divide and conquer recursion is used in C5.0, an upgraded version of C4.5. It fixes issues with fault pruning, overfitting or underfitting, and robustness to noise and missing data [21]. Using equation 3's expression for the purity of a sample's entropy, we can:

$$P\left(e\right) = \sum_{i}^{N} \frac{\left(p + n_i\right)}{p + n} .I\left(p_i, n_i\right) \tag{3}$$

Where, P is the number of values of positive.
The value n indicates how many records are negative.
The entropy function is I(p, n).

- **Extreme Gradient Boosting (XGBoost):** Regression tree XGBoost [6,22] contains decision criteria that are identical to those of decision tree [6]. It is employed in both classification and regression. This technique is a practical and affordable variation of the gradient boosting machine (GBM) that proven to be broadly applicable in computer vision, data mining etc. [6]. Recent developments have mainly focused on two elements of XGBoost as a gradient boosting machine: rapid the tree creation and designing a novel dispersed algorithm for searches [30]. The fundamental idea of XGBoost is to raise the importance of the objective task. Given a dataset D = (xi,yi), where xi stands for the ASD profile and yi for the associated binary value.

Assuming that the XGBoost model [6] comprises of K decision trees, the optimization objective function is given by equation (4):

$$y_i = \sum_{k=1}^{K} fk\left(x_i\right),\ fk \in F \tag{4}$$

Where each fk correlates to a separate tree with outside scores.
F is the regression tree's space.

- **Adaboost:** This is a boosted classification tree constructed a pseudocode which decreases misclassification mistakes by iterating pseudocode [23]. It can furthermore take care of disappeared accounts, and increases many classifiers which are good at performance. Consider (a_1, b_1) be taken as the initial and (a_n,b_n) as n^{th} training illustrations. Later, it considered all loads of sample $W_1(k) = \frac{1}{n}$ for k= 1,2,…..n.

Where W denotes the weight of the samples for k^{th} training sample. Then, it trains weak learners using W_t distribution and the hypothesis is given in equation (5).

$$h_t : X \in \left\{-1, 1\right\} \tag{5}$$

Then α_t is chosen, where α represents the weight for this classifier [2]. N_t is selected as normalized factor and W_{t+1} is selected as a distribution in order to enhance the weight i.e. α and is given in equation (6).

$$W_{t+1(k)} = \frac{W_t\left(k\right) e^{-\alpha tykht(xk)}}{N_t} \tag{6}$$

Later, the classifier model is denoted in equation (7):

$$W_{t+1(k)} = sign\left(\sum_{t=1}^{T} atht(a)\right) \tag{7}$$

In order to characterize the various classifiers conclusions and performance based on these metrics, there were a number of valuation metrics, including accuracy, precision, recall, the F1 measure, and AUROC. The metrics were denoted by finding the true positive (TP) [1], true negative (TN)[1], false positive (FP) [1] and false negative (FN) [1] values which is given in the table 4 [1]. After evaluation, we investigated the greatest classifiers leads to highest results for all datasets. Additionally these datasets are examined to figure out the best classifiers that are giving the best results in this analysis. Our Proposed methodology overview is briefly explained in algorithm 1.

This algorithm explains about prediction of autism. Toddler and child datasets are taken for prediction in the early stage of ASD. In Data pre-processing, outliers, misplaced values are detached and then FT's are applied to these datasets to make the data normalised by removing skewness. After that ML models are implemented on these datasets for predictions. And also FST's are applied to find the most significant feature for prediction of ASD, which helped doctors for diagnosis. For example, without FT's we achieved less accuracy like 56% using KNN, 78% using RF. Later, in this proposed method, I identified this gap and applied FT's on these datasets, then we achieved best results with 98% for child and 100% for toddler's dataset and found the important feature i.e. A5 feature is most significant feature for toddler, A4 stands most significant feature for child based on univariate and RFE.

Algorithm 1*: Algorithm For Prediction Of ASD*

Step 1: Start

Step 2: //Input
Upload Dataset Ds for Toddler and Child. The input is in the form of QChat

Step 3: //Output
Prediction of ASD or no ASD in early stage of the life.

Step 4: ## Stage 1 Data Pre-processing
if (noise==T || outliers==T || NULL values==T || Missing values==T)
remove noise outliers NULL values Missing values from Input D_s
end if
Stage 2 Applying Feature Transformation Technique (FT's)
1.Log Transformation
import FunctionTransformer from sklearn
declare variable var
var = FunctionTransformer(func = np.log1p)
fit the model
2. Power transformation (PT)
 # (a) Box-Cox PT
import Powertransformation from sklearn
declare pow
pow = PowerTransformation(method = 'Box-cox', standardize=true)
fit the model
 # (b) Yeo-Johnson
import Powertransformation from sklearn
declare pow
pow = PowerTransformation(method = 'Yeo-johnson, standardize=true)
fit the model

Step 5: ## Apply ML classifiers for the transformed data

Let us declare i , j // i represents FT's
// j represents the number of classifier's
for i ← 1 to 3
for j ←1 to 6
for each FT_i apply six C_j and explore the data and find the best FT_i with respect to best C_j's
end
end
end

Step 6: ## Stage 4 Apply Feature selection technique's (FST)
call univariate function
for i 1 ← n // n is number of features
select best features i.e f_i in both toddler and child D_s

Step 7: End

3.5 Feature Selection Techniques

After classification, we then identified the most important features, which are required to predict the ASD in early stage that will be useful for doctors to diagnose autism in early stage. So two different Feature selection techniques (FST) such as Univariate FST and Recursive feature elimination (RFE) FST's are used to achieve objective 3 which are given in the table 5.

Table 4. Evaluation metrics

Metrics	Details	Formula
Accuracy	It is defined as adding of TP and TN divided by total number of data points [23].	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	It is a ratio of TP and the sum of TP and FP	$pre = \frac{TP}{TP + FP}$
Recall	It is defined as the ratio of TP to FN's total.	$recall = \frac{TP}{TP + FN}$
F1 Score	It is a harmonic mean of precision and recall.	$F1score = 2 * \frac{precision * recall}{precision + recall}$
AUC	Area under curve. It measures the whole 2D area under the whole ROC curve. It is TP versus FP rate at decision threshold rate [25].	$TPR = \frac{TP}{TP + FN}$ $TPR = \frac{FP}{FP + TN}$

Table 5. Brief description of different FST's

FST	Abbreviation	Details
Univariate	UNI	Univariate feature selection studies every feature separately to determine the power of the connection of the feature with the response variable
Recursive feature elimination	RFE	It selects the most related features from the dataset. RFE is a wrapper-type feature selection algorithm.

4. Experimental Results

The results achieved through evaluation of experiment using machine learning models with three Feature Transformations were discussed in this section. Additionally, the comparative result analysis of the optimal Feature transformation with Random forest, KNN, SVM, Decision tree, Gradient Boosting and Adaboost were discussed. Furthermore, the optimal feature using Feature selection method is identified for toddler and child to predict the autism in early stage.

4.1 Experiment analysis of Toddler Dataset

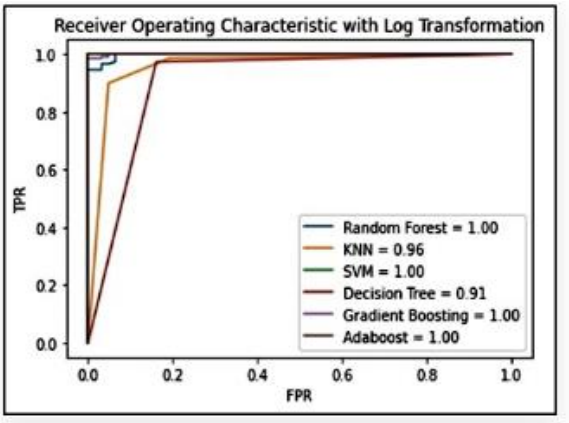
Different ML algorithms were implemented with three transformation such Log Transformation, Power Transformation-Box cox and Power transformation-Yeo- Johnson feature technique. The Adaboost classifier has given highest accuracy of 100% for all transformation functions.

4.1.1 Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Log Transformation

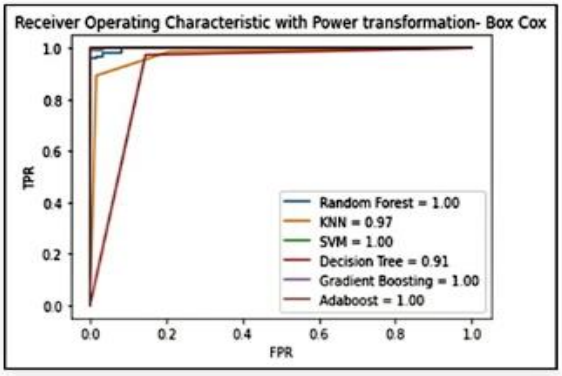
Random forest, KNN, SVM, DT, Gradient boosting and Adaboost are implemented with Log Transformation feature technique, and the results obtained on the test data is compared with different machine learning models and also results are represented in Table 6. Fig.2a. represents the AUC for different machine learning models with Log Transformations. The model with highest accuracy was Adaboost with 1.00 AUCROC score.

Table 6. Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Log Transformation for Toddler's

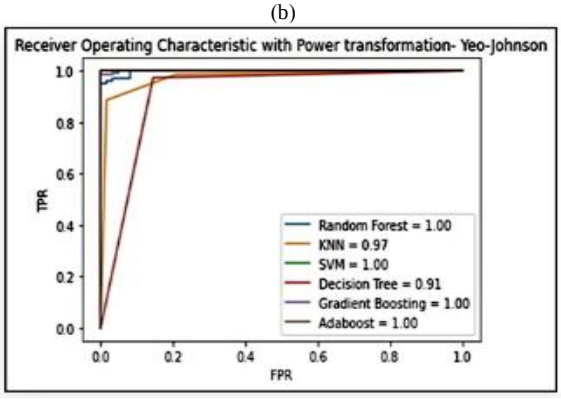
Model Description	Accuracy	Precision	Recall	F1-Score	AUCROC
RF	96	94	100	97	1.00
KNN	91	98	90	94	0.96
SVM	90	92	99	90	1.00
DT	93	94	97	96	0.91
Gradient Descent	99	98	100	99	1.00
Adaboost	100	100	100	100	1.00



(a)



(b)



(c)

Fig. 2. AUC scores of Toddler dataset for different Machine Learning Algorithms w.r.t (a) Log Transformation (b) Power Transformation- Box – Cox Transformation (c) Power Transformation- Yeo – Johnson Transformation

4.1.2 Evaluation of Random forest, K nearest neighbour, support vector machine, Decision tree (DT), Gradient boosting and Adaboost with Power Transformation- Box – Cox Transformation

Random forest, KNN, SVM, DT, Gradient boosting and Adaboost are implemented with Power Transformation- Box – Cox Transformation feature technique, and the results obtained on the test data is compared with different machine learning models and also results are represented in Table 7. Fig.2b. represents the AUC for different machine learning models with Power Transformation- Box – Cox Transformation. Adaboost was found to be the model with the highest accuracy, with a score of 1.00 AUCROC.

Table 7. Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Power Transformation-Box-Cox Transformation for Toddler's

Model Description	Accuracy	Precision	Recall	F1-Score	AUCROC
RF	96	93	100	96	1.00
KNN	92	89	90	94	0.97
SVM	97.15	96	100	98	1.00
DT	94.3	95	97	96	0.91
Gradient Descent	99	98	100	99	1.00
Adaboost	100	100	100	100	1.00

4.1.3 Evaluation of Random forest, K nearest neighbour, support vector machine, Decision tree, Gradient boosting and Adaboost with Power Transformation- Yeo – Johnson Transformation

Random forest, KNN, SVM, DT, Gradient boosting and Adaboost are implemented with Power Transformation- Yeo – Johnson Transformation feature technique, and the results obtained on the test data is compared with different machine learning models and also results are represented in Table 8. Fig.2c. represents the AUC for different machine learning models with Power Transformation- Yeo – Johnson Transformation. Adaboost was found to be the model with the highest accuracy, with a score of 1.00 AUCROC.

Table 8. Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Power Transformation- Yeo – Johnson Transformation for Toddler's

Model Description	Accuracy	Precision	Recall	F1-Score	AUCROC
RF	96	95	100	97	1.00
KNN	91.4	99	89	94	0.97
SVM	98	97	100	98	1.00
DT	93.35	94	97	95	0.91
Gradient Descent	99	98	100	99	1.00
Adaboost	100	100	100	100	1.00

4.2 Experiment analysis of Child Dataset

Different ML algorithms were implemented with three transformation such as Log Transformation, Power Transformation-Box cox and Power transformation-Yeo- Johnson feature technique. It was found that, the Adaboost model has given highest accuracy of 100% for all transformation functions.

4.2.1 Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Log Transformation

Random forest, KNN, SVM, DT, Gradient boosting and Adaboost are implemented with Log Transformation feature technique, and the results obtained on the test data is compared with different machine learning models and also results are represented in Table 9. Fig.2a. represents the AUC for different machine learning models with Log Transformations. With 0.81 and 1.00 AUCROC scores, Random forest and Adaboost were found to be the most accurate models.

Table 9. Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Log Transformation for Child

Model Description	Accuracy	Precision	Recall	F1-Score	AUCROC
RF	98.3	97	100	98	0.81
KNN	92	96	87	91	0.50
SVM	90	90	90	90	1.00
DT	95	94	97	95	0.50
Gradient Descent	97	94	100	97	0.83
Adaboost	98	100	100	100	1.00

4.2.2 Evaluation of Random forest, K nearest neighbour, support vector machine, Decision tree (DT), Gradient boosting and Adaboost with Power Transformation- Box – Cox Transformation

Random forest, KNN, SVM, DT, Gradient boosting and Adaboost are implemented with Power Transformation- Box – Cox Transformation feature technique, and the results obtained on the test data is compared with different machine learning models and also results are represented in Table 10. Fig.3b. represents the AUC for different machine learning models with Power Transformation- Box – Cox Transformation. It is found that, the model with highest accuracy was Random Forest and Adaboost with 0.99 and 1.00 AUCROC score.

Table 10. Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Power Transformation-Box-Cox Transformation for Child

Model Description	Accuracy	Precision	Recall	F1-Score	AUCROC
RF	98	97	100	98	0.99
KNN	92	96	87	91	0.97
SVM	93	96	90	93	1.00
DT	95	94	97	95	0.93
Gradient Descent	97	94	100	97	1.00
Adaboost	98	100	100	100	1.00

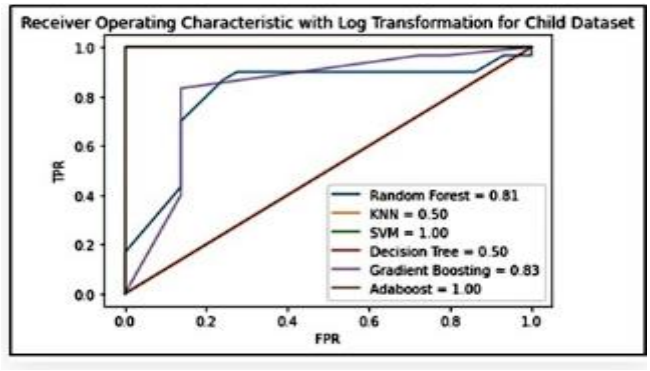
4.2.3 Evaluation of Random forest, K nearest neighbour, support vector machine, Decision tree, Gradient boosting and Adaboost with Power Transformation- Yeo – Johnson Transformation

Random forest, KNN, SVM, DT, Gradient boosting and Adaboost are implemented with Power Transformation- Yeo – Johnson Transformation feature technique, and the results obtained on the test data is compared with different

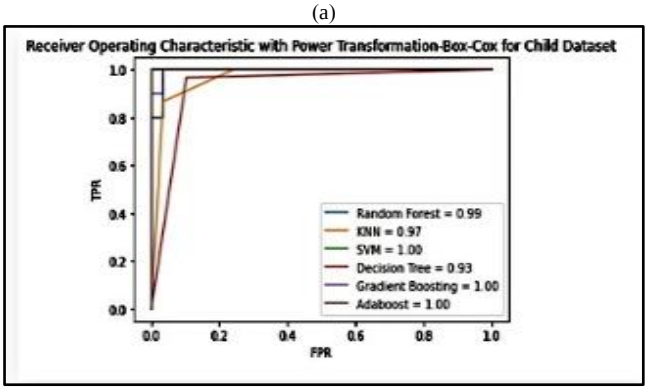
machine learning models and also results are represented in Table 11. Fig.3c. represents the AUC for different machine learning models with Power Transformation- Yeo – Johnson Transformation. It is found that, the model with highest accuracy was Adaboost and Gradient Descent with 1.00 AUCROC score.

Table 11. Evaluation of Random forest (RF), K nearest neighbour (KNN), support vector machine (SVM), Decision tree (DT), Gradient boosting and Adaboost with Power Transformation- Yeo – Johnson Transformation for Child

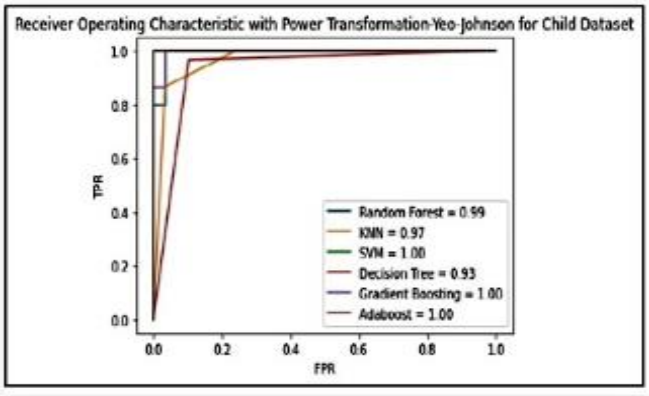
Model Description	Accuracy	Precision	Recall	F1-Score	AUCROC
RF	98	97	100	98	0.99
KNN	92	96	87	91	0.97
SVM	93	96	90	93	1.00
DT	92	90	93	92	0.93
Gradient Descent	97	94	100	97	1.00
Adaboost	97	100	100	100	1.00



(a)



(b)



(c)

Fig. 3. AUC scores of Child dataset for different Machine Learning Algorithms w.r.t (a) Log Transformation (b) Power Transformation- Box – Cox Transformation (c) Power Transformation- Yeo – Johnson Transformation

4.3 Feature Ranking for Toddler and Child

We implemented two different Feature selection techniques (FST) approaches like univariate (UNI) and recursive feature elimination (RFE) to recognize the important features of both toddler and child datasets which is given in table 12. A5 is the most significant feature to predict autism for toddler based on all FST's. A4 and A10 is the most

significant feature to predict autism for child based on all FST's given in Fig.4. The Feature ranking is given in Table 12.

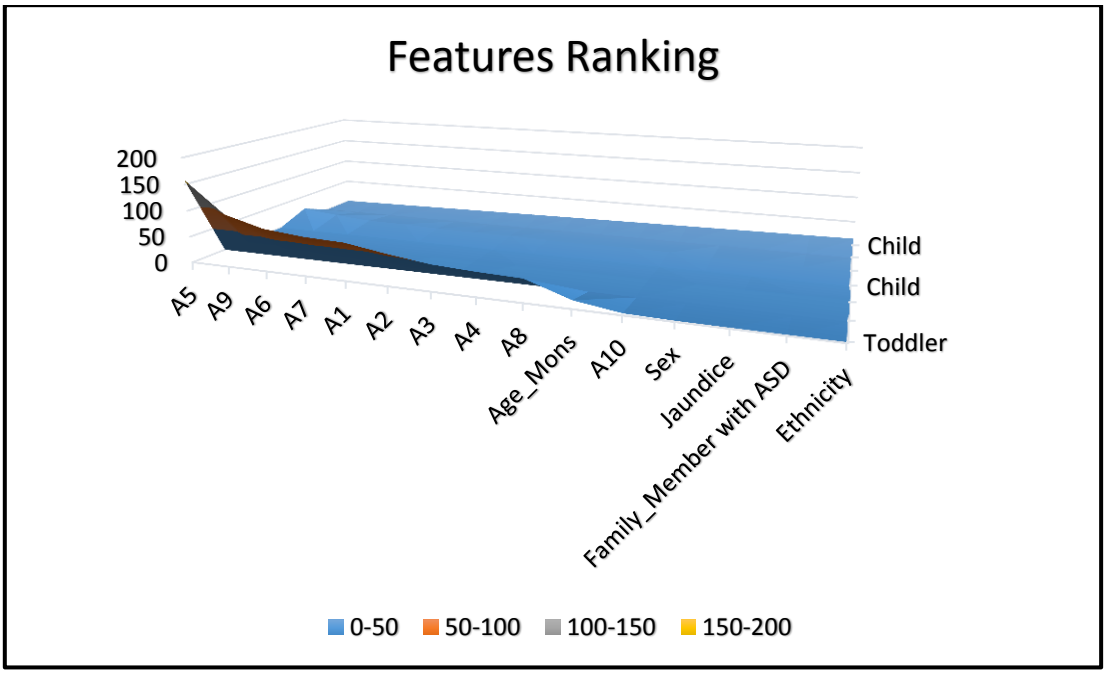


Fig.4. Features Ranking for Toddler and Child dataset

Table 12. Feature Ranking

Toddler				Child			
Features	**Univariate**	**Features**	**RFE**	**Features**	**Univariate**	**Features**	**RFE**
A5	158.838	A5	1	A4	22.031	A4	1
A9	100.09	A7	1	A10	14.510	A10	2
A6	80.272	A9	1	A6	8.502	A9	3
A7	73.263	A6	1	A2	8.298	A8	4
A1	70.575	A1	1	A1	5.878	A1	5
A2	58.578	A2	2	A5	5.446	A2	6
A3	48.57	A4	2	A3	5.0913	A3	7
A4	44.566	Family_Member with ASD	3	Age	4.343	Used_app before	8
A8	42.192	A8	3	A7	4.147	Autism	9
Age_Mons	15.13	A3	3	A8	4.139	A5	10
A10	3.44	A10	3	Jaundice	3.448	A6	11
Sex	1.815	Jaundice	4	Gender	1.691	Age	12
Jaundice	1.562	Sex	4	Used app before	1.495	Gender	13
Family_Member with ASD	0.678	Ethnicity	5	Autism	1.176	A7	14
Ethnicity	0.171	Age_Mons	5	Ethnicity	0.358	Ethnicity	15

4.4 Comparative Evaluation of Machine Learning Models

Omar et al. [31] proven an ASD detection through Random Forest-CART (RF-CART) and Random Forest-ID3 (RF-ID3) [1] and detected ASD with 92% accuracy for child dataset. Talabani et al. [39] used SVM for child ASD datasets and they attained the accuracy of 95.54%. Thabtah [33] used AQ10 dataset for predictive analysis using several classification algorithms. They acheived the highest results for Logistic Regression as 98% accuracy for the child, 94% accuracy for adolescent and 99.85% accuracy, for adult. Our study has given highest accuracy of 100% and 98.3% for Toddler and child as shown in Fig.5. The proposed model is also compared to earlier studies. However, several earlier research did not use the toddler's analyses (look table 13). Additionally, we used numerous FSTs to analyze and rank the features in toddler and child datasets and based on that, we discovered the elements that are most crucial for predicting ASD, which were not adequately displayed in earlier studies [30, 32].

Table 13. Comparing a proposed model with existing studies

		Accuracy	Precision	Recall	F1 score	AUROC
Toddler	Kazi et al.[31]	-	-	-	-	-
	Aktar et al.[1]	98.77	99.39	99.39	97.10	99.98
	Thabtah et al.[33]	-	-	-	-	-
	Proposed Model	100	100	100	100	100
Child	Omar et al.[31]	92.26	89.76			
	Aktar et al.[1]	98.77	99.39	99.39	97.10	99.98
	Thabtah et al.[33]	97.80	92.30	92.20	-	-
	Proposed Model	98.3	100	100	100	100

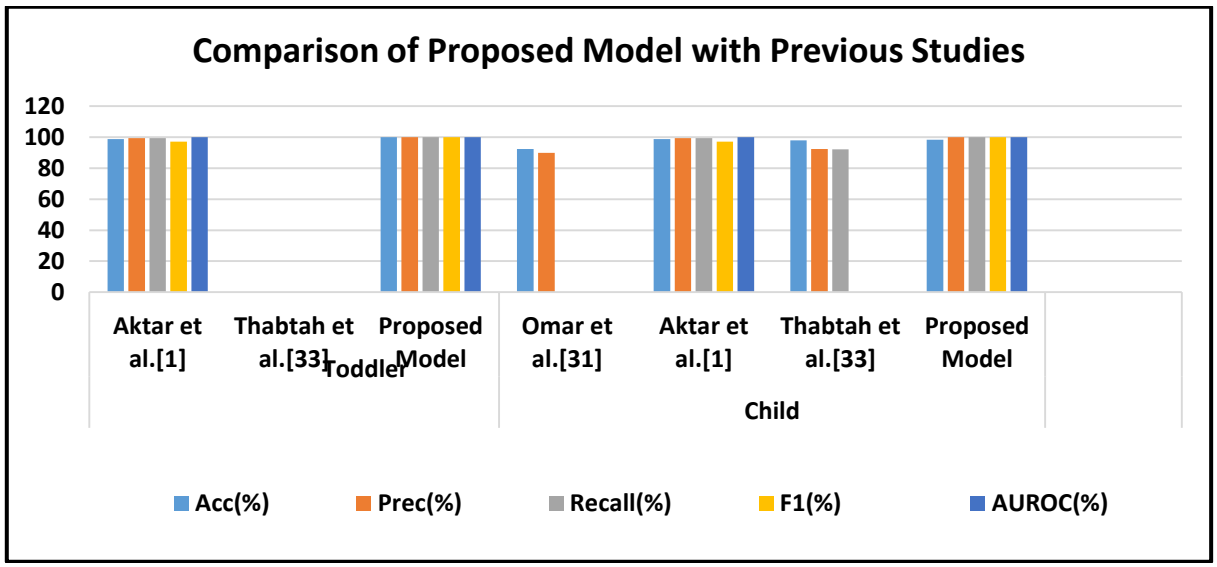


Fig. 5. Comparing a proposed model with existing studies

5. Discussions

Several scholars have executed readings with ASD datasets [1], but ASD prediction yet wants major enhancements [27, 28]. In my study, we are focusing on predicting the ASD in early stages of life. Therefore we have taken toddler and child dataset, since these dataset have less age and analyzed the results by using different classifiers and different feature transformation techniques to discover the important ASD features. We got the 100% outcomes, the top predictions for all accuracy metrics in the random sample dispersal of investigational results and by taking these results we have compared with the earlier readings. After scrutinizing autism dataset through applying various FT approaches and then executed classifiers (RF, KNN, SVM, DT, XGBoost and ADABoost) for transformed datasets in python. After analyzing these datasets, important FT methods were explored that will give the better performance than others is depicted in Fig.2 and Fig.3.

6. Conclusion

In summary, three different FT methods are implemented for different stages of ASD datasets, and then applied several classifiers to examine those transformed data and estimated the performance. And then, the best features were explored which are extremely predictive for ASD with the help of feature selection techniques. This will progress the capability of doctors to identify ASD at the early stage by using our recognized features. Several accuracy metrics, such as accuracy precision, Recall, F1-score and AUCROC curve were used to indicate how well our performance evaluations performed. Adaboost and RF gave good performance for toddler and child dataset. In future, we will find the related drawbacks of this method and study other neuro disorders which are associated with ASD in order to progress the performance in prediction of ASD and other related neurodevelopment syndromes.

References

[1] Akter, Tania, Md Shahriare Satu, Md Imran Khan, Mohammad Hanif Ali, Shahadat Uddin, Pietro Lio, Julian MW Quinn, and Mohammad Ali Moni. "Machine learning-based models for early stage detection of autism spectrum disorders." *IEEE Access* 7 (2019): 166509-166527.

[2] C. Allison, B. Auyeung, and S. Baron-Cohen, ‘‘Toward brief ‘red flags’ for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls,’’J. Amer. Acad. Child Adolescent Psychiatry, vol. 51, no. 2, pp. 202–212, 2012.

[3] F. Thabtah, F. Kamalov, and K. Rajab, ‘‘A new computational intelligence approach to detect autistic features for autism screening,’’Int. J. Med. Inform., vol. 117, pp. 112–124, Sep. 2018.

[4] F. Thabtah and D. Peebles, ‘‘A new machine learning model based on induction of rules for autism detection”, Health Inform. J., 2019, Art. no. 1460458218824711, doi: 10.1177/1460458218824711.

[5] M. S. Satu, F. F. Sathi, M. S. Arifen, M. H. Ali, and M. A. Moni, ‘‘Early detection of autism by extracting features: A case study in Bangladesh,’’in Proc. 1st Int. Conf. Robot., Elect. Signal Process. Techn. (ICREST), Jan. 2019, pp. 87–90.

[6] H. Abbas, F. Garberson, E. Glover, and D. P. Wall, ‘‘Machine learning approach for early detection of autism by combining questionnaire and home video screening,’’ J. Amer. Med. Informat. Assoc., vol. 25, no. 8,pp. 1000–1007, 2018.

[7] F. Thabtah, ‘‘Machine learning in autistic spectrum disorder behavioral research: A review and ways forward,’’ Informat. Health Social Care vol. 44, no. 3, pp. 278–297, 2018.

[8] F. Thabtah, ‘‘Autism spectrum disorder screening: Machine learning adaptation and DSM-5 fulfillment,’’ in Proc. 1st Int. Conf. Med. Health Inform., 2017, pp. 1–6.

[9] K. C. Howlader, M. S. Satu, A. Barua, and M. A. Moni, ‘‘Mining significant features of diabetes mellitus applying decision trees: A case study in Bangladesh,’’bioRxiv, Nov. 2018, Art. no. 481994.

[10] M. A. Hossain, S. M. S. Islam, J. M. Quinn, F. Huq, and M. A. Moni, ‘‘Machine learning and bioinformatics models to identify gene expression patterns of ovarian cancer associated with disease progression and mortality,’’J. Biomed. Inform., vol. 100, Oct. 2019, Art. no. 103313, doi: 10.1016/j.jbi.2019.103313.

[11] M. Duda, R. Ma, N. Haber, and D. P. Wall, ‘‘Use of machine learning for behavioral distinction of autism and ADHD,’’Transl. Psychiatry, vol. 6, no. 2, p. e732, 2016.

[12] K. L. Goh, S. Morris, S. Rosalie, C. Foster, T. Falkmer, and T. Tan, ‘‘Typically developed adults and adults with autism spectrum disorder classification using centre of pressure measurements,’’ in Proc. IEEE Int.Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2016, pp. 844–848.

[13] A. Crippa, C. Salvatore, P. Perego, S. Forti, M. Nobile, M. Molteni, and I. Castiglioni, ‘‘Use of machine learning to identify children with autism and their motor abnormalities,’’ J. Autism Develop. Disorders, vol. 45,no. 7, pp. 2146–2156, 2015.

[14] Autism Screening Data for Toddlers. Accessed: Sep. 10, 2018. [Online]. Available: https://www.kaggle.com/fabdelja/autism-screening-for- toddlers

[15] UCI Machine Learning Repository: Autistic Spectrum Disorder Screening Data for Children Data Set. Accessed: Sep. 10, 2018. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Autistic+ Spectrum+Disorder+Screening+Data+for+Children++

[16] UCI Machine Learning Repository: Autistic Spectrum DisorderScreening Data for Adolescent Data Set. Accessed: Sep. 10, 2018.[Online].Available:https://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent +++

[17] UCI Machine Learning Repository: Autism Screening Adult Data Set. Accessed: Sep. 10, 2018. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Autism+Screening+Adult

[18] Zhang, Yao, Jianxue Wang, and Xu Luo. "Probabilistic wind power forecasting based on logarithmic transformation and boundary kernel." *Energy conversion and management* 96, pp. 440-451, (2015).

[19] Liu, Yanli, Yourong Wang, and Jian Zhang. "New machine learning algorithm: Random forest." In *International Conference on Information Computing and Applications*, pp. 246-252. Springer, Berlin, Heidelberg, 2012.

[20] Wang, Lishan. "Research and implementation of machine learning classifier based on KNN." In *IOP Conference Series: Materials Science and Engineering*, vol. 677, no. 5, p. 052038. IOP Publishing, 2019.

[21] Bujlow, T. Riaz, and J. M. Pedersen, ‘‘A method for classification of network traffic based on C5.0 machine learning algorithm,’’ in Proc. Int.Conf. Comput., Netw. Commun. (ICNC), Jan./Feb. 2012, pp. 237–241

[22] Ma, Baoshan, Fanyu Meng, Ge Yan, Haowen Yan, Bingjie Chai, and Fengju Song. "Diagnostic classification of cancers using extreme gradient boosting algorithm and multi-omics data." *Computers in biology and medicine* 121 (2020): 103761.

[23] S. Satu, T. Akter, and M. J. Uddin, ‘‘Performance analysis of classifyinglocalization sites of protein using data mining techniques and artificial neu-ral networks,’’ inProc. Int. Conf. Elect., Comput. Commun. Eng. (ECCE),Feb. 2017, pp. 860–865.

[24] Praveena, K. N., and R. Mahalakshmi. "Classification of Autism Spectrum Disorder and Typically Developed Children for Eye Gaze Image Dataset using Convolutional Neural Network." *International Journal of Advanced Computer Science and Applications* 13, no. 3 (2022).

[25] S. Satu, S. Ahamed, F. Hossain, T. Akter, and D. M. Farid, ‘‘Mining traffic accident data of N5 national highway in bangladesh employing decision trees,’’ in Proc. IEEE Region 10 Humanitarian Technol. Conf. (R10-HTC) ,Dec. 2017, pp. 722–725.

[26] M. S. Satu, S. Ahamed, A. Chowdhury, and M. Whaiduzzaman, ‘‘Exploring significant family income ranges of career decision difficulties ofadolescents in Bangladesh applying regression techniques,’’ inProc. Int.Conf. Elect., Comput. Commun. Eng. (ECCE), Feb. 2019, pp. 1–6.

[27] M. E. Hossain, A. Khan, M. A. Moni, and S. Uddin, ‘‘Use of electronic health data for disease prediction: A comprehensive literature review,’’ IEEE/ACM Trans. Comput. Biol. Bioinf., 2019, doi: 10.1109/TCBB.2019.2937862.

[28] M. R. Islam, A. R. M. Kamal, N. Sultana, R. Islam, M. A. Moni, and A. Ulhaq, ‘‘Detecting depression using K nearest neighbors (KNN) classification technique,’’ in Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (IC4ME2), Feb. 2018, pp. 1–4.

[29] Praveena, K.N., Mahalakshmi, R. (2022). A Survey on Early Prediction of Autism Spectrum Disorder Using Supervised Machine Learning Methods. In: Rana, N.K., Shah, A.A., Iqbal, R., Khanzode, V. (eds) Technology Enabled Ergonomic Design. HWWE 2020. Design Science and Innovation. Springer, Singapore. https://doi.org/10.1007/978-981-16-6982-8_2
[30] K. S. Oma, P. Mondal, N. S. Khan, M. R. K. Rizvi, and M. N. Islam, ‘‘A machine learning approach to predict autism spectrum disorder,’’ in Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE) , Feb. 2019,pp.1–6.
[31] Omar, Kazi Shahrukh, Prodipta Mondal, Nabila Shahnaz Khan, Md Rezaul Karim Rizvi, and Md Nazrul Islam. "A machine learning approach to predict autism spectrum disorder." In *2019 International conference on electrical, computer and communication engineering (ECCE)*, pp. 1-6. IEEE, 2019.
[32] H. Talabani and E. Avci, ‘‘Performance comparison of SVM kernel types on child autism disease database,’’ in Proc. Int. Conf. Artif. Intell. Data Process. (IDAP), Sep. 2018, pp. 1–5
[33] F.Thabtah, ‘‘An accessible and efficient autism screening method for behavioural data and predictive analyses”, Health Informat. J., Sep. 2018, Art. no. 1460458218796636.

Authors’ Profiles

Praveena K N is Assistant Professor in Department of Computer Science and Engineering at presidency university, Bengaluru. She is pursuing PhD degree in Computer Science and Engineering with specialization in Machine learning. Her research areas are machine learning, image analysis.

R Mahalakshmi, Associate Professor, Department of Computer science and Engineering at presidency university, Bengaluru. She has done her PhD in Computer science and Engineering with specialization in image processing. Her research areas are image processing, machine learning.

Manjunath C, Assistant Professor, School of Mechanical Engineering, Reva University, Yelahanka, Bengaluru. He is pursuing PhD degree in Mechanical Engineering with specialization in thermal and integrated with AI. His research areas are in CFD’s, Integrated with AI.

Ahmad Faiz Zubair received his Ph.D. in CAD/CAM Engineering from the Universiti Sains Malaysia in 2019. He received his MSc in CAED Engineering from Strathclyde University, Glasgowin 2009and Bachelor of engineering in CAD/CAM Engineering from University of Malaya in 2006. He is currently a Senior Lecturer in the Faculty of Mechanical Engineering Universiti Teknologi MARA,Penang Branch. There he is a member of the Intelligent and Sustainable Manufacturing Centre, and the advisor of Student’s Boiler and Safety Club. His research interests include CAD/CAM, CAPP, Manufacturing Technology and CNC Machining.He is also a register Chartered Engineerof IMechE United Kingdomandan active ProfessionalTechnologist of MalaysiaBoard of Technologist. He is a Certified Solidworks Associated (CSWA) and certified CATIA Mechanical DesignAssociate.

P. **Karthikeyan** obtained his the Bachelor of Engineering (B.E.,) in Computer Science and Engineering from Anna University, Chennai, and Tamil nadu, India in 2005 and received his Master of Engineering (M.E,) in Computer Science and Engineering from Anna University, Coimbatore India in 2009. He has completed Ph.D. degree in Anna University, Chennai in 2018. Skilled in developing projects and carrying out research in the area of Cloud computing and Data science with the programming skill in Java, Python, R and C. He published more than 30 International journals with good impact factor and presented more than 20 International conferences. He was the reviewer of Elsevier, Springer, Inderscience and reputed Scopus indexed journals. He is acting as editorial board members in EAI Endorsed Transactions on Energy Web. The International Arab Journal of Information Technology and Blue Eyes Intelligence Engineering and Sciences Publication journal.

How to cite this paper: Praveena K N, Mahalakshmi R, Manjunath C, Ahmad Faiz Zubair, P. Karthikeyan, "Optimized Feature Selection and Transformations for Early Stage Prediction of Autism Using Supervised Machine Learning Models", International Journal of Modern Education and Computer Science(IJMECS), Vol.15, No.6, pp. 73-89, 2023. DOI:10.5815/ijmecs.2023.06.06