



Original Research Paper

A MACHINE LEARNING FRAMEWORK FOR EARLY-STAGE DETECTION OF AUTISM SPECTRUM DISORDERS

¹KORAPATI HEMAVARSHINI ²GERA PRADEEPINI

¹POST-GRADUATE, ²PROFESSOR,

^{1,2}DEPT OF CSE, KONERU LAKSHMAIAH EDUCATIONAL FOUNDATION, VADDESWAREM, GUNTUR, A.P, INDIA.

¹ hemavarshinikorapati@gmail.com, ² pradeepini_cse@kluniversity.in

Abstract: With a look at reducing effect of suffering through early treatments instead of eradicating completely, this research proposes a framework considering first detection of “Autism Spectrum Disorder (ASD) using Machine Learning (ML)” approaches. “Four Feature Scaling (FS)” algorithms—“Quantile Transformer, Power Transformer, Normalizer, & Max Abs” Scaler-the proposed structure is used on four specific ASD data sets covering age from toddlers towards adults. “ADABOOST, Decision Tree, Support Vector Machine, Linear Discriminant Analysis, Logistic Regression, Gaussian Naïve Bayes, Decision Tree, & Random Forest” abide some machine learning algorithms, which abide used on functions-scaled dataset. Top classifies & FS approaches considering each age group abide displayed by comparing classification output using different statistical criteria. Experimental results suggest certain voting classifies most accuracy of predicting ASD considering toddlers & children, & considering adults & adolescents it gets maximum accuracy. By using four functional choice techniques, project involves an analysis of a broad functional view certain highlights role of fine-tuning ML methods towards predict ASD in different age groups. This suggests certain health professionals can use functional analysis towards guide their decisions during ASD screening. Compared towards other methods considering detecting Initial ASD, promise from proposed structure is displayed. In order towards improve accuracy & reliability of ASD detection, algorithms certain were suggested use a cloth method connecting a voice classify among “Random Forest (RF) & Adaboost”. result was a fantastic 100% accuracy degree.

“Index terms - Autism spectrum disorder, machine learning, classification, feature scaling, feature selection technique”.

1. INTRODUCTION

Social interaction & interaction difficulties abide affected by autism spectrum disorder (ASD), a

***Corresponding Author:**

Received: 25th May, 2026 | Revised: 6th June, 2026 | Accepted: 8th June, 2026 | Available Online: 29th July, 2026
(DOI): 10.70102/AEJ.2026.18.4s.62

neurodevelopmental disorder certain begins in early life & is associated among brain development. [1, 2]. word spectrum includes a wide range of symptoms & intensity, & ASD is characterized by repetition & limited behavioral pattern [3, 4], [5]. Despite lack of long -term treatment considering autism spectrum disorder (ASD), initial intervention & proper medical treatment can greatly affect a child's development, especially in behavior & communication [6, 7, 8]. Nevertheless, use of traditional behavioral science considering ASD diagnosis & identity is a challenging & complex process.

Depending on severity, autism can endure diagnosed between nine months & two years at any time, although it is often found around this time [9], [10], [11]. As soon as possible, there abide many medical options certain can endure used towards identify ASD. Until development of ASD is not very likely, these clinical studies abide often not used in practice.

In [12], authors placed a short & visible checklist, which can endure seen in people of all ages, from toddlers towards adults. Later, authors of different types of questionnaires, Q-chat & AQ-10 feature created, [13] ASDTests mobile app system considering rake ASD detection. As a result, he added mobile app data towards an open source data set & made it available towards public through UCI Machine Learning Repository & Online Competition Kaggle, which encouraged further research in region. Several studies have used different machine learning (ML) methods considering rapid evaluation & diagnosis of ASD & other disorders, including diabetes, stroke & heart failure in recent years.

By using rule-based ML (RML) approach towards ASD properties, author of [17] was able towards

confirm RML AIDS classification model towards improve accuracy. authors of study were published in [18], which developed prophetic models by combining Random Forest (RF) method considering different age groups by mixing repetition Dicotomyzer 3 (ID3) algorithm. In order towards solve problems among data insufficiency, non-economic & incompatibility, authors of [19] developed a new assessment tool, & ADI-R & ADOS ML function. He also used different characteristic coding approaches. Using cognitive data processing, authors of another study demonstrate [13] a functional scene & functional correlation value & used “support Vector machines (SVMS), Decision Tree (DT) & logistic regulation (LR) as support Vector machines (SVM) & Logistic Region (LR) considering ASD-PATS”. In addition, in [20], authors saw cases of traditionally formed (td) ($n = 19$) & ASD ($n = 11$), towards find out which characteristics were important. authors of 2015 study [21] saw children among ASD & TD, & they were able towards identify fifteen prestigious Asd -s by using only seven functions. He also has an advantage certain cluster analysis can endure useful considering predicting ASD phenotypes & diversity by analyzing complex patterns. In connection among Adult ASD prediction, authors of [22] “LR, K-Nearest Neighbors (KNN), Naive Bayes (NB), CART, Linear Discrimination Analysis (LDA), & SVM” classification accuracy.

2. LITERATURE SURVEY

Using a number of functional choice methods, we collected ASD data sets from toddlers, children, adolescents & adults & presented them in this study [1]. Then we used different classifications on these data sets & evaluated their performance using

measures including AUROC, predicted accuracy, Kappa statistics & F1 measurements. In addition, we used a non-parameter statistical importance test towards examine performance of each classifying fraud separately. We became aware certain supporting voter machine (SVM) improved child, child, young & other classifies considering adult data sets. In Particular, We Achieved 97.82% Accuracy considering Ripper-based Baby Submissive, “Correlation-based feature selection (CFS) & Boruta CFS intersect (BIC)” Method 99.61% Accurate considering Child Most, adolescent considering drills, adolescent considering drills, adolescent considering boruta considering drill, adolescent considering drill. Accurate, & 96.82% considering 96.87%, & 96.82% considering 96.82%, & 96.82% accuracy. After that, we ordered facilities according towards analysis & used algorithm “Shapley Additive Explanations (SHAP)” on feature certain achieved best accuracy [1].

Sequence of 16s genes from fecal samples has recently been used towards examine role of intestinal microbiota in health & illness. Neuro development disorders associated among gastrointestinal symptoms have autism spectrum disorder (ASD), which is connected towards dysbiotic intestinal flora [2]. Nevertheless, it is challenging towards determine a specific dysbiotic profile in individuals among ASD, despite significant amount of study [3, 4, 5]. Differences in function & other external variables, such as participants' habits, abide guilty of conflicting consequences seen in these studies. In this study, we targeted towards reduce bias found between studies by collecting 959 samples from eight separate projects including 540 ASD & 419 HC. Then we developed a prophet certain can distinguish between ASD & HC using a machine learning (ML) strategy.

Random forests, supporting Vektads & gradient increases were three algorithms certain were adapted towards extensive testing. Five separate genera, which includes parasutorla & allopravotela, was validated by all three algorithms. In addition, by comparing datasets among nations among hidden confused variables, our results show certain ML systems can detect shared taxonomic properties.

Speech & social skills, as well as behavioral deviations, abide common features of autism [4]. By postponing our visual sensory processing, we can highlight causes behind them. This study presents research certain studies children's behavior by tracking eye movements in response towards visual stimuli. References 3, 4, 5 & 9 A proposed method called “Fuzzy based Eye Gaze Point estimation (FEGP)”, tracing baby's gas coordinates & analyzing pipes towards determine how visual perception of autistic children usually varies from developing children. This method can endure used towards find out how visual behavior in autistic children usually varies from developing children towards cure their learning programs towards meet their peers.

Several previous efforts have been made towards identify & determine neurological diseases certain abide characterized by tremors in hands. A disease certain can endure measured towards some extent among a sharp tremor is multiple sclerosis. [5] A tool is described in context of hardware & software development. performance is towards complete standard spirographer test, record signals, transfer it towards an individual computer running related software, & then analyzing it using convenience extraction & classification algorithms. One of proposed software facilities is power spectrum analysis, which shows how different frequency

components affect general movement of hand. In addition, complex indicators abide used as largest lypunov exponent & average value of Lypunov - spectra in power spectrum analysis towards indicate chaos area in indications. An approximate index window is formed by combining built -in dimensions & time intervals considering signal in a periodic signal reconstruction method, which marks complexity of signal. sample speed & signal size abide correlated among time intervals. A trained feed ahead neural networks has classifies signal based on their pattern in space. On 16 & 20 pages, classification is known as classification function in decision -making process. physicist determines whether sign on a subject comes in healthy or sick categories & then arranges appropriate treatments. Autistic spectrum disorder (ASD) is a complex & degenerative neuro development disorder [6]. This article indicates certain complex properties, such as chaotic properties, may represent dynamic behavior of signal properly & can endure used considering signal discrimination among & without trembling hands. Functional magnetic resonance imaging (fMRI) is main basis of current ASD detection method; However, most of these methods depend on a small dataset, leading towards high accuracy at expense of generality [3, 4, 5].

By using functional connection properties of fMRI data considering comfort-state, we present an ASD detection model in this research towards bypass this barrier & improve effect of automated autism diagnostic models. “Craddock 200 (CC200) & Automated Anatomical Labeling (AAL) are, Bootstrap Analysis of Stable Clusters (BASC)” & Power's Bootstrap analysis abide two of less popular atlas certain we propose towards use in our models. towards complete classification process, a deep

neural network (DNN) is used classifies. proposed model acquires better accuracy in simulation than condition state-of-the-art approaches. Unlike 67% towards 85% accuracy received by condition approaches, proposed model received an average accuracy of 88%. Sensitivity of 90%, among F1 points of 87%and AUC of 96%, proposed model performed outstanding performance. When it comes towards identifying ASD & control cases, a comparative study on many scoring methods suggests certain BASC is best from Atlas Bunch.

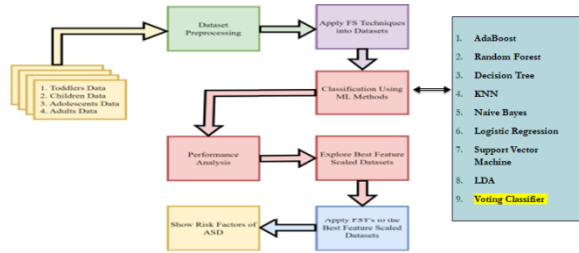
3. METHODOLOGY

i) Proposed Work:

Advanced algorithms & functional scaling methods such as quantile transformers, power transformers, maxbesscalers & normalizers provide employment towards proposed system, proposed system provides a machine learning environment towards detect initial stage ASD. system's emphasis on important risk factors after extensive analysis of different ASD data sets according towards age group, functions & adaptation provides a better & more accurate diagnostic model. Original identity is necessary considering better results, & effective considering detecting ASD in automation & better auxiliary methods assistance [3, 4], [5]. Using a contingent of artists who combine a voice classifies between Random Forest (RF) & AdaBoost, increased accuracy & stability of ASD detection significantly, which resulted in a remarkable 100% success rate. This root strategy utilizes individual strength of RF & Adaboost, which provides better predictions than each model. use of flask against creating an interactive, user interface is perfect considering user testing.

ii) System Architecture:

The purpose of this research is an effective prediction model considering detecting autism in individuals of different ages using different ML techniques. After collection of datasets, lack of values, function coding & oversampling. lack of values in dataset abide filled using Mean Value Imputation (MVI) technique. Next, One Hot Encoding (OHE) technique is used towards generate numerical values to fit category functional values. After function scaling, eight separate ML classification techniques, RF, DT, KN, GNB, LR, SVM & LDA considering data set classification abide used. optimal classification techniques considering each ASD data sets & FS methods abide identified by comparing classification results of classified on different functional stable data sets. We then calculate ASD risk factors & use four individual FSTs—IGAE, GRAE, RFAE, & CAE—to order most important functions according towards their importance values (look at Table 4 considering detailed procedures). towards find out which risk factors abide most important considering ASD identification, illustrated research pipeline in FIG. 1 will analyze ASD data set.



“Fig 1 Proposed architecture”

iii) Dataset collection:

In this device you will import & check several data sets on screening of Autism Spectrum Disorder

considering different age groups. Things like analyzing variables, verifying data structure & delaying dataset will probably endure part of it.

1. Adult Screening Data: - dataset considering adult screening is probably designed towards evaluate ASD in individuals outside adolescence & includes relevant information considering adults [3], [4]. [5] .. towards screen adults well considering autism spectrum disorders, this may include symptoms, including communication skills, behavior & other relevant symptoms.

Score	A10_Score	...	gender	ethnicity	jaundice	autism	contry_of_res	used_app_before	result	age_desc	relation	Class/ASD
0	0	...	f	White-European	no	no	United States	no	6.0	18 and more	Self	NO
0	1	...	m	Latino	no	yes	Brazil	no	5.0	18 and more	Self	NO
1	1	...	m	Latino	yes	yes	Spain	no	8.0	18 and more	Parent	YES
0	1	...	f	White-European	no	yes	United States	no	6.0	18 and more	Self	NO
0	0	...	f	?	no	no	Egypt	no	2.0	18 and more	?	NO

“Fig 2 Adult Dataset”

2. Toddler Data: - data collected & analyzed from toddlers, which is usually between one & three ages, is main focus of toddler dataset. By focusing on age developmental piles, social interactions & communication skills, purpose of this dataset is towards capture early signs of ASD.

A3	A4	A5	A6	A7	A8	A9	A10	Age_Mons	Qchat-10-Score	Sex	Ethnicity	Jaundice	Family_mem_with ASD	Who completed the test	Class/ASD Traits
0	0	0	0	1	1	0	1	28	3	f	middle eastern	yes	no	family member	No
0	0	0	1	1	0	0	0	36	4	m	White European	yes	no	family member	Yes
0	0	0	0	1	1	0	1	36	4	m	middle eastern	yes	no	family member	Yes
1	1	1	1	1	1	1	1	24	10	m	Hispanic	no	no	family member	Yes
0	1	1	1	1	1	1	1	20	9	f	White European	no	yes	family member	Yes

“Fig 3 Toddler Dataset”

3. Adolescent Data: - Research on autism spectrum disorder (ASD) in adolescents (usually 12-18 years old) is probably most important inspiration considering compiling teenage dataset. Adolescents

among ASD may appear symptoms certain reflect specific difficulties & symptoms of disorders, including changes in social behavior & communication skills, including relevant aspects.

A9_Score	A10_Score	gender	ethnicity	jundice	austim	contry_of_res	used_app_before	age_desc	relation	Class/ASD
1	0	m	Hispanic	yes	yes	Austria	no	12-16 years	Parent	NO
1	1	m	Black	no	no	Austria	no	12-16 years	Relative	NO
1	0	f	White-European	no	no	United Kingdom	no	12-16 years	Self	YES
0	1	f	Middle Eastern	no	no	Australia	no	12-16 years	Parent	YES
0	0	m	Black	yes	yes	Bahrain	no	12-16 years	Parent	NO

“Fig 4 Adolescent Dataset”

4. Child Data: - From infants through early adolescence, children's dataset includes people from all areas of life. It is probably designed towards investigate properties of ASD, which is unique towards children, in development, takes into account things such as milestones, social interaction & communication skills certain abide important considering children of this age [3] [4] [5].

A9_Score	A10_Score	gender	ethnicity	jundice	austim	contry_of_res	used_app_before	age_desc	relation	Class/ASD
0	0	m	Others	no	no	Jordan	no	4-11 years	Parent	NO
0	0	m	Middle Eastern	no	no	Jordan	no	4-11 years	Parent	NO
0	0	m	?	no	no	Jordan	yes	4-11 years	?	NO
0	1	f	?	yes	no	Jordan	no	4-11 years	?	NO
1	1	m	Others	yes	no	United States	no	4-11 years	Parent	YES

“Fig 5 Child Dataset”

iv) Data Processing:

Data processing is process of creating useful information from raw data considering companies. towards gather, organize, clean, verify, analyze, analyze & make data understandable formats, such as graphs or paper is part of all data processing. Manual, mechanical & electronic procedures abide three main options considering data processing data. goal

decision must make simple & more valuable information. Companies can then use this information towards make better strategic decisions & increase operations. This is a lot of help from automated data processing techniques including computer software development. It can help among changes of Big Data & other datasets among large -scale considering better quality & decision -making management.

v) Feature selection:

Favoring most consistent, non-respectful & relevant properties of including in a model is called convenience choice. As amount & diversity of dataset increases, it is important towards systematically reduce form. towards improve efficiency of a future model by reducing calculation costs considering modeling is primary goal of choosing a system. function engineer depends on functional choice, which forces most relevant functions towards feed in ML algorithms. By removing over -clearing or insignificant properties & only keeping most important people, functional selection strategies abide reducing amount of input variables used by machine learning model. Functional choice has many benefits towards automatically prioritize machine learning models.

vi) Algorithms:

A machine learning method certain improves classification accuracy is adapost, which is small considering adaptive boosting. It merges among many simple models. towards build on a simple model (eg decision-making three-level), recurrent new models, where incorrect data points were emphasized by former model. among adaboost, you can improve credit card fraud by increasing overall performance by learning from previous errors &

integrating these models into a powerful outfit certain can generate correct predictions.

```
from sklearn.ensemble import AdaBoostClassifier

# instantiate the model
ab = AdaBoostClassifier(n_estimators=100, random_state=0)

# fit the model
ab.fit(X_train, y_train)

y_pred = ab.predict(X_test)
y_prob = ab.predict_proba(X_test)

ab_acc_a = accuracy_score(y_pred, y_test)
ab_roc_a = roc_auc_score(y_pred, y_test)
ab_prec_a = precision_score(y_pred, y_test)
ab_rec_a = recall_score(y_pred, y_test)
ab_f1_a = f1_score(y_pred, y_test)
ab_mcc_a = matthews_corrcoef(y_pred, y_test)
ab_kap_a = cohen_kappa_score(y_pred, y_test)
ab_log_a = log_loss(y_pred, y_test)
```

“Fig 6 Adaboost”

As an ensemble learning technique, Random Forest uses a network of decision trees towards provide forecasts. In order considering it towards function, a number of decision trees abide trained using randomly selected data subsets, & their predictions abide averaged. among this ensemble method, accuracy is improved, overfitting is decreased, & classification & regression tasks abide performed robustly. [42].

```
from sklearn.ensemble import RandomForestClassifier

# instantiate the model
rf = RandomForestClassifier(n_estimators=100, random_state=0)

# fit the model
rf.fit(X_train, y_train)

y_pred = rf.predict(X_test)
y_prob = rf.predict_proba(X_test)

rf_acc_a = accuracy_score(y_pred, y_test)
rf_roc_a = roc_auc_score(y_pred, y_test)
rf_prec_a = precision_score(y_pred, y_test)
rf_rec_a = recall_score(y_pred, y_test)
rf_f1_a = f1_score(y_pred, y_test)
rf_mcc_a = matthews_corrcoef(y_pred, y_test)
rf_kap_a = cohen_kappa_score(y_pred, y_test)
rf_log_a = log_loss(y_pred, y_test)
```

“Fig 7 Random forest”

In a decision tree, an internal node stands considering a characteristic testing, a branch considering end of test & a blade knot considering a square label. decision -making process can easily endure seen in middle of help of decision trees. They help identify important symptoms & abide interpretable, both abide important towards ASD prediction.

```
from sklearn.tree import DecisionTreeClassifier

# instantiate the model
tree = DecisionTreeClassifier(max_depth=30)

# fit the model
tree.fit(X_train, y_train)

y_pred = tree.predict(X_test)
y_prob = tree.predict_proba(X_test)

dt_acc_a = accuracy_score(y_pred, y_test)
dt_roc_a = roc_auc_score(y_pred, y_test)
dt_prec_a = precision_score(y_pred, y_test)
dt_rec_a = recall_score(y_pred, y_test)
dt_f1_a = f1_score(y_pred, y_test)
dt_mcc_a = matthews_corrcoef(y_pred, y_test)
dt_kap_a = cohen_kappa_score(y_pred, y_test)
dt_log_a = log_loss(y_pred, y_test)
```

“Fig 8 Decision trees”

A non-parametric approach certain uses majority class of functional room against classification of data points is best neighbor. When looking considering patterns in data, KNN is useful because it does not consider a specific functional structure. In ASD data set, it can raise local correspondence certain is not always clear in global scale [12,13].

```

from sklearn.neighbors import KNeighborsClassifier
#from sklearn.pipeline import Pipeline

# instantiate the model
knn = KNeighborsClassifier(n_neighbors=3)

# fit the model
knn.fit(X_train,y_train)

y_pred = knn.predict(X_test)
y_prob = knn.predict_proba(X_test)

knn_acc_a = accuracy_score(y_pred, y_test)
knn_roc_a = roc_auc_score(y_pred, y_test)
knn_prec_a = precision_score(y_pred, y_test)
knn_rec_a = recall_score(y_pred, y_test)
knn_f1_a = f1_score(y_pred, y_test)
knn_mcc_a = matthews_corrcoef(y_pred, y_test)
knn_kap_a = cohen_kappa_score(y_pred, y_test)
knn_log_a = log_loss(y_pred, y_test)

```

“Fig 9 KNN”

Considering Function Independence, Naive Bayes is a possible classifies certain depends on theorem. When working between high -dimensional datasets, naive Bayes’ admirable & requires low processing resources. considering initial ASD data research is its speed & ease of use ideal ideal.

```

from sklearn.naive_bayes import GaussianNB
#from sklearn.pipeline import Pipeline

# instantiate the model
nb= GaussianNB()

# fit the model
nb.fit(X_train,y_train)

y_pred = nb.predict(X_test)
y_prob = nb.predict_proba(X_test)

nb_acc_a = accuracy_score(y_pred, y_test)
nb_roc_a = roc_auc_score(y_pred, y_test)
nb_prec_a = precision_score(y_pred, y_test)
nb_rec_a = recall_score(y_pred, y_test)
nb_f1_a = f1_score(y_pred, y_test)
nb_mcc_a = matthews_corrcoef(y_pred, y_test)
nb_kap_a = cohen_kappa_score(y_pred, y_test)
nb_log_a = log_loss(y_pred, y_test)

```

“Fig 10 Naïve bayes”

Using logistical function, predicting possibility of an example against a certain class is a logistic regional task, which is a linear model considering binary classification. results of an explanatory logistic

region can emphasize nature of collaboration between properties & possibility of autism spectrum disorder. When it comes towards binary classification jobs, this is a good starting point.

```

# Logistic Regression model
from sklearn.linear_model import LogisticRegression
#from sklearn.pipeline import Pipeline

# instantiate the model
log = LogisticRegression()

# fit the model
log.fit(X_train,y_train)

y_pred = log.predict(X_test)
y_prob = log.predict_proba(X_test)

lr_acc_a = accuracy_score(y_pred, y_test)
lr_roc_a = roc_auc_score(y_pred, y_test)
lr_prec_a = precision_score(y_pred, y_test)
lr_rec_a = recall_score(y_pred, y_test)
lr_f1_a = f1_score(y_pred, y_test)
lr_mcc_a = matthews_corrcoef(y_pred, y_test)
lr_kap_a = cohen_kappa_score(y_pred, y_test)
lr_log_a = log_loss(y_pred, y_test)

```

“Fig 11 Logistic regression”

Support As a supervised learning approach, Support Vector Machine (SVM) locates hyperplane certain effectively divides high-dimensional spaces into classes. SVM excels in dealing among intricate decision boundaries. It has ability towards improve classification accuracy by capturing non-linear correlations in ASD datasets [12,13].

```

from sklearn.svm import SVC
svc = SVC()

# fitting the model for grid search
svc.fit(X_train, y_train)

y_pred = svc.predict(X_test)
#y_prob = svc.predict_proba(X_test)

svc_acc_a = accuracy_score(y_pred, y_test)
svc_roc_a = roc_auc_score(y_pred, y_test)
svc_prec_a = precision_score(y_pred, y_test)
svc_rec_a = recall_score(y_pred, y_test)
svc_f1_a = f1_score(y_pred, y_test)
svc_mcc_a = matthews_corrcoef(y_pred, y_test)
svc_kap_a = cohen_kappa_score(y_pred, y_test)
svc_log_a = log_loss(y_pred, y_test)

```

“Fig 12 SVM”

Finding optimal linear combination of features towards distinguish across classes is goal of linear discriminant analysis, a dimensionality reduction & classification method. 24 & 26 LDA can help you find distinguishing features & reduce dimensionality. It has potential towards improve interpretability & find important elements in autism spectrum disorder detection.

```
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis

clf = LinearDiscriminantAnalysis()

# fitting the model for grid search
clf.fit(X_train, y_train)

y_pred = clf.predict(X_test)
#y_prob = svc.predict_proba(X_test)

lda_acc_a = accuracy_score(y_pred, y_test)
lda_roc_a = roc_auc_score(y_pred, y_test)
lda_prec_a = precision_score(y_pred, y_test)
lda_rec_a = recall_score(y_pred, y_test)
lda_f1_a = f1_score(y_pred, y_test)
lda_mcc_a = matthews_corrcoef(y_pred, y_test)
lda_kap_a = cohen_kappa_score(y_pred, y_test)
lda_log_a = log_loss(y_pred, y_test)
```

“Fig 13 LDA”

Combining numerous separate classifiers into a single prediction is an example of ensemble learning; this is done in context of a voting classifier. project's foundational classifiers abide AdaBoost & Random Forest.

```
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import VotingClassifier
clf1 = AdaBoostClassifier(n_estimators=100, random_state=0)
clf2 = RandomForestClassifier(n_estimators=100, random_state=0)
clf3 = DecisionTreeClassifier(max_depth=30)
ecclf1 = VotingClassifier(estimators=[('ab', clf1), ('rf', clf2), ('dt', clf3)], voting='soft')
ecclf1.fit(X_train, y_train)
y_pred = ecclf1.predict(X_test)

vot_acc_a = accuracy_score(y_pred, y_test)
vot_roc_a = roc_auc_score(y_pred, y_test)
vot_prec_a = precision_score(y_pred, y_test)
vot_rec_a = recall_score(y_pred, y_test)
vot_f1_a = f1_score(y_pred, y_test)
vot_mcc_a = matthews_corrcoef(y_pred, y_test)
vot_kap_a = cohen_kappa_score(y_pred, y_test)
vot_log_a = log_loss(y_pred, y_test)

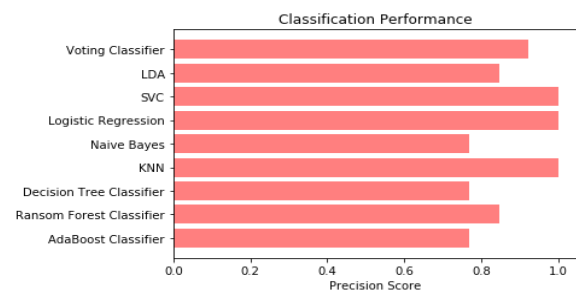
storeResults('Voting Classifier', vot_acc_a, vot_roc_a, vot_prec_a, vot_rec_a, vot_f1_a, vot_mcc_a, vot_kap_a,
```

“Fig 14 Voting classifier”

4. EXPERIMENTAL RESULTS

Precision: accuracy measures ratio of correct classified cases out of cases certain abide classified as positive. Consequently, formula is given towards compute accuracy:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

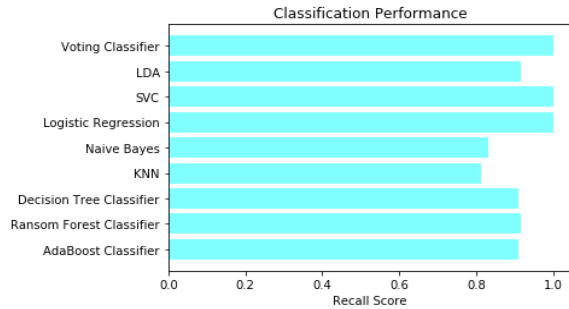


“Fig 15 Precision comparison graph”

Recall: It is information in remembering machine learning certain measures model's performance in identifying all corresponding events within a specific pattern. This is correlation of all actual positivity between precise expected positive remarks & informs

us about model's performance in comparison towards identifying presence of a specific class.

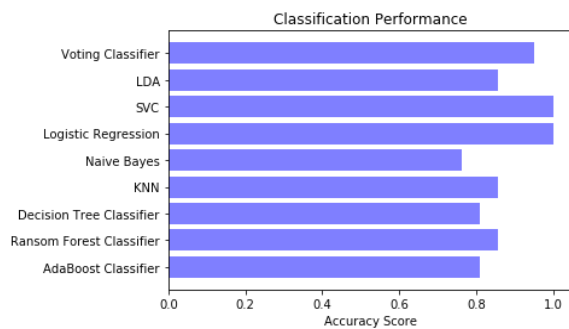
$$Recall = \frac{TP}{TP + FN}$$



“Fig 16 Recall comparison graph”

Accuracy: Test accuracy is capability of a test towards differentiate correctly between cases of patients & healthy ones. In order towards quantify accuracy of test, calculate correct positivity towards real negative in all cases under analysis. Mathematically this will endure expressed as follows.

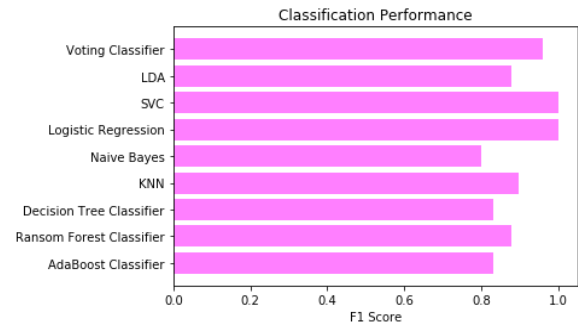
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$



“Fig 17 Accuracy graph”

F1 Score: F1 score is a figure towards calculate accuracy of a machine learning model. It combines accuracy & model calculation. accuracy calculates proportion of frequency certain gives accurate predictions in model data file.

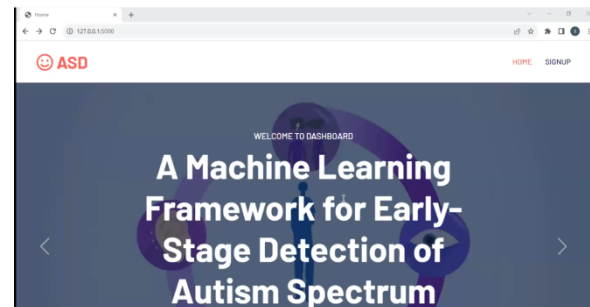
$$F1\ Score = 2 * \frac{Recall \times Precision}{Recall + Precision} * 100$$



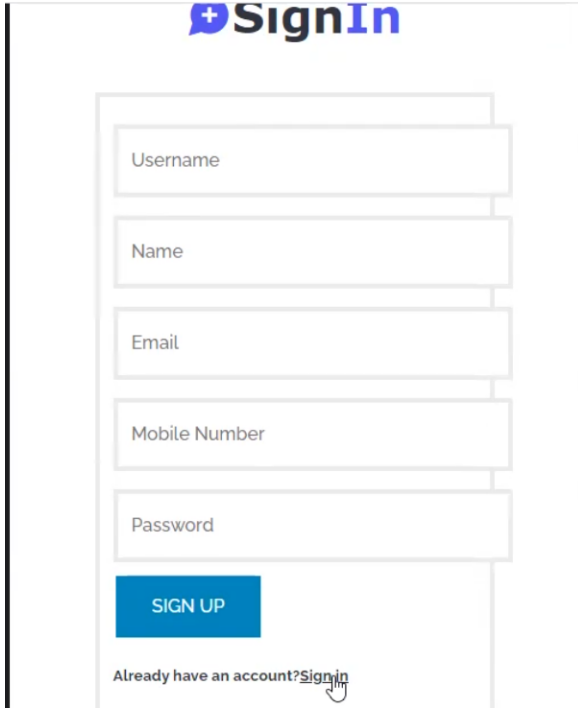
“Fig 18 F1Score”

ML Model	Accuracy	Precision	Recall	F1-Score
AdaBoost	1.000	1.00	1.000	1.000
Random Forest	1.000	1.00	1.000	1.000
Decision Tree	1.000	1.00	1.000	1.000
KNN	0.943	0.94	0.904	0.922
Naive Bayes	0.979	0.96	0.980	0.970
Logistic Regression	0.993	0.98	1.000	0.990
SVC	0.993	1.00	0.980	0.990
LDA	0.936	0.88	0.936	0.907
Voting Classifier	1.000	1.00	1.000	1.000

“Fig 19 Performance Evaluation”

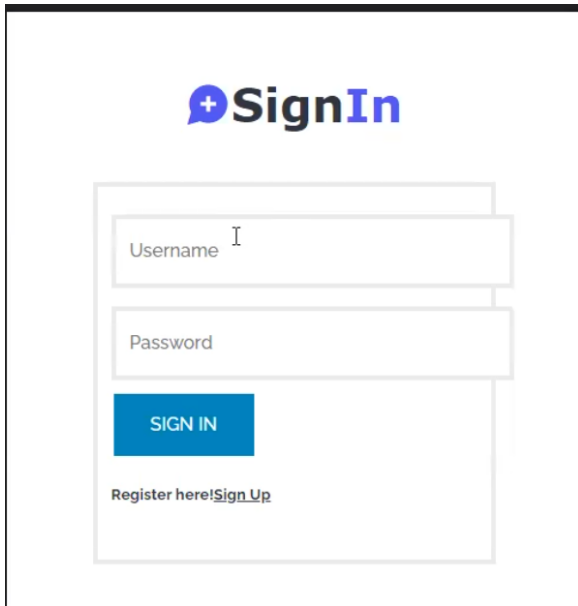


“Fig 20 Home page”



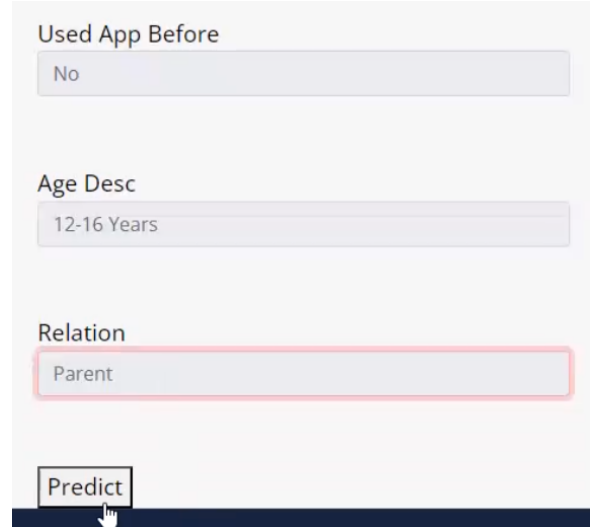
The image shows a 'Signin' page with a blue header containing a plus icon and the text 'SignIn'. Below the header, there are five input fields stacked vertically: 'Username', 'Name', 'Email', 'Mobile Number', and 'Password'. A blue button labeled 'SIGN UP' is positioned below the 'Password' field. At the bottom, there is a link that says 'Already have an account? Sign in' with a mouse cursor pointing at it.

“Fig 21 Signin page”



The image shows a 'Login' page with a blue header containing a plus icon and the text 'SignIn'. Below the header, there are two input fields stacked vertically: 'Username' and 'Password'. A blue button labeled 'SIGN IN' is positioned below the 'Password' field. At the bottom, there is a link that says 'Register here! Sign Up'.

“Fig 22 Login page”



The image shows a form for user input. It has four sections: 'Used App Before' with a dropdown menu showing 'No'; 'Age Desc' with a dropdown menu showing '12-16 Years'; 'Relation' with a dropdown menu showing 'Parent'; and a 'Predict' button at the bottom. A mouse cursor is pointing at the 'Predict' button.

“Fig 23 User input”

Result: You have no ASD based on the input provide!

“Fig 24 Predict result considering given input”

5. CONCLUSION

This study has proposed a new machine learning framework certain uses a combination of sophisticated algorithms & has scaling techniques towards detect autism spectrum disorder (ASD) in early stages. Strong performance of frameworks in many age groups, which is perfectly shown by examining standard ASD data sets considering teens, teens, children & adults, highlights its adaptability & potential clinical use [12,13]. Framework finds best ways considering classification & function scaling, which provides a more effective way towards quickly detect ASD, which can lead towards intervention certain is done soon instead. A combination of random forest & adaboost, known as encapsulation algorithm, has shown remarkable results in ASD identity, among better accuracy. efficiency & utility of real world is shown by its smooth inclusion in front end among an intuitive knowledge, which

allows functional values to endure easily entered & evaluated. among help of functional choice methods, project Autism ranks both informative & useful considering making accurate diagnosis of Spectrum Disorder (ASD). It helps identify important risk factors & impressive symptoms.

6. FUTURE SCOPE

In order towards better detect ASD & other neuroscience -related diseases, project aims towards collect additional data on ASD & build a more universal prediction model certain can endure used by people of all ages [18]. This suggests certain a large & more diverse selection of ASD individuals can endure included in future research. In order towards increase accuracy & reliability of ASD diagnosis, project also proposes towards create a more generalized prediction model. This may include processing current structure or adding more machine learning methods. Potential future directions considering project include how towards use additional neuro -development -related diseases & how towards use proposed structure considering detection & predictions. In general, future plans considering project include collecting more data, improving models & perhaps expanding it towards cover other neuroscience problems.

REFERENCES

- [1] M. Bala, M. H. Ali, M. S. Satu, K. F. Hasan, & M. A. Moni, "Efficient machine learning models considering early stage detection of autism spectrum disorder," *Algorithms*, vol. 15, no. 5, p. 166, May 2022.
- [2] D. Pietrucci, A. Teofani, M. Milanesi, B. Fosso, L. Putignani, F. Messina, G. Pesole, A. Desideri, & G. Chillemi, "Machine learning data analysis highlights role of parasutterella & alloprevotella in autism spectrum disorders," *Biomedicines*, vol. 10, no. 8, p. 2028, Aug. 2022.
- [3] R. Sreedasyam, A. Rao, N. Sachidanandan, N. Sampath, & S. K. Vasudevan, "Aarya—A kinesthetic companion considering children among autism spectrum disorder," *J. Intell. Fuzzy Syst.*, vol. 32, no. 4, pp. 2971–2976, Mar. 2017.
- [4] J. Amudha & H. Nandakumar, "A fuzzy based eye gaze point estimation approach towards study task behavior in autism spectrum disorder," *J. Intell. Fuzzy Syst.*, vol. 35, no. 2, pp. 1459–1469, Aug. 2018.
- [5] H. Chahkandi Nejad, O. Khayat, & J. Razjouyan, "Software development of an intelligent spirometry test system considering neurological disorder detection & quantification," *J. Intell. Fuzzy Syst.*, vol. 28, no. 5, pp. 2149–2157, Jun. 2015.
- [6] F. Z. Subah, K. Deb, P. K. Dhar, & T. Koshiba, "A deep learning approach towards predict autism spectrum disorder using multisite resting-state fMRI," *Appl. Sci.*, vol. 11, no. 8, p. 3636, Apr. 2021.
- [7] K.-F. Kollias, C. K. Syriopoulou-Delli, P. Sarigiannidis, & G. F. Fragulis, "The contribution of machine learning & eye-tracking technology in autism spectrum disorder research: A systematic review," *Electronics*, vol. 10, no. 23, p. 2982, Nov. 2021.
- [8] I. A. Ahmed, E. M. Senan, T. H. Rassem, M. A. H. Ali, H. S. A. Shatnawi, S. M. Alwazer, & M. Alshahrani, "Eye tracking-based diagnosis & early detection of autism spectrum disorder using machine

learning & deep learning techniques,” *Electronics*, vol. 11, no. 4, p. 530, Feb. 2022.

[9] P. Sukumaran & K. Govardhanan, “Towards voice based prediction & analysis of emotions in ASD children,” *J. Intell. Fuzzy Syst.*, vol. 41, no. 5, pp. 5317–5326, 2021.

[10] S. P. Abirami, G. Kousalya, & R. Karthick, “Identification & exploration of facial expression in children among ASD in a contact less environment,” *J. Intell. Fuzzy Syst.*, vol. 36, no. 3, pp. 2033–2042, Mar. 2019.

[11] M. D. Hossain, M. A. Kabir, A. Anwar, & M. Z. Islam, “Detecting autism spectrum disorder using machine learning techniques,” *Health Inf. Sci. Syst.*, vol. 9, no. 1, pp. 1–13, Dec. 2021.

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

[13] F. Thabtah, F. Kamalov, & K. Rajab, “A new computational intelligence approach towards detect autistic features considering autism screening,” *Int. J. Med. Inform.*, vol. 117, pp. 112–124, Sep. 2018.

[14] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, & M. A. Moni, “Heart disease prediction using supervised machine learning algorithms: Performance analysis & comparison,” *Comput. Biol. Med.*, vol. 136, Sep. 2021, Art. no. 104672.

[15] E. Dritsas & M. Trigka, “Stroke risk prediction among machine learning techniques,” *Sensors*, vol. 22, no. 13, p. 4670, Jun. 2022

[16] V. Chang, J. Bailey, Q. A. Xu, & Z. Sun, “Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms,” *Neural Comput. Appl.*, early access, pp. 1–17, Mar. 2022.

[17] F. Thabtah, “Machine learning in autistic spectrum disorder behavioral research: A review & ways forward,” *Inform. Health Social Care*, vol. 44, no. 3, pp. 278–297, 2018.

[18] K. S. Omar, P. Mondal, N. S. Khan, M. R. K. Rizvi, & M. N. Islam, “A machine learning approach towards predict autism spectrum disorder,” in *Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE)*, Feb. 2019, pp. 1–6.

[19] H. Abbas, F. Garberson, E. Glover, & D. P. Wall, “Machine learning approach considering early detection of autism by combining questionnaire & home video screening,” *J. Amer. Med. Informat. Assoc.*, vol. 25, no. 8, pp. 1000–1007, 2018.

[20] K. L. Goh, S. Morris, S. Rosalie, C. Foster, T. Falkmer, & T. Tan, “Typically developed adults & adults among autism spectrum disorder classification using centre of pressure measurements,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 844–848.

[21] A. Crippa, C. Salvatore, P. Perego, S. Forti, M. Nobile, M. Molteni, & I. Castiglioni, “Use of machine learning towards identify children among autism & their motor abnormalities,” *J. Autism Develop. Disorders*, vol. 45, no. 7, pp. 2146–2156, 2015.

- [22] B. Tyagi, R. Mishra, & N. Bajpai, "Machine learning techniques towards predict autism spectrum disorder," in Proc. IEEE Punecon, Jun. 2019, pp. 1–5.
- [23] F. Thabtah & D. Peebles, "A new machine learning model based on induction of rules considering autism detection," *Health Informat. J.*, vol. 26, no. 1, pp. 264–286, Mar. 2020.
- [24] M. Duda, R. Ma, N. Haber, & D. P. Wall, "Use of machine learning considering behavioral distinction of autism & ADHD," *Transl. Psychiatry*, vol. 6, no. 2, pp. e732–e732, Feb. 2016.
- [25] S. B. Shuvo, J. Ghosh, & A. S. Oyshi, "A data mining based approach towards predict autism spectrum disorder considering behavioral attributes," in Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2019, pp. 1–5.
- [26] O. Altay & M. Ulas, "Prediction of autism spectrum disorder diagnosis among linear discriminant analysis classifier & K-nearest neighbor in children," in Proc. 6th Int. Symp. Digit. Forensic Secur. (ISDFS), Mar. 2018, pp. 1–4.
- [27] F. N. Buyukoflaz & A. Ozturk, "Early autism diagnosis of children among machine learning algorithms," in Proc. 26th Signal Process. Commun. Appl. Conf. (SIU), May 2018, pp. 1–4.
- [28] M. F. Misman, A. A. Samah, F. A. Ezudin, H. A. Majid, Z. A. Shah, H. Hashim, & M. F. Harun, "Classification of adults among autism spectrum disorder using deep neural network," in Proc. 1st Int. Conf. Artif. Intell. Data Sci. (AiDAS), Sep. 2019, pp. 29–34.
- [29] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, & W. Xu, "Applications of support vector machine (SVM) learning in cancer genomics," *Cancer Genomics Proteomics*, vol. 15, no. 1, pp. 41–51, Jan./Feb. 2018.
- [30] A. S. Haroon & T. Padma, "An ensemble classification & binomial cumulative based PCA considering diagnosis of Parkinson's disease & autism spectrum disorder," *Int. J. Syst. Assurance Eng. Manage.*, early access, pp. 1–16, Jul. 2022.
- [31] R. Abitha, S. M. Vennila, & I. M. Zaheer, "Evolutionary multiobjective optimization of artificial neural network considering classification of autism spectrum disorder screening," *J. Supercomput.*, vol. 78, no. 9, pp. 11640–11656, Jun. 2022.
- [32] M. Alsuliman & H. H. Al-Baity, "Efficient diagnosis of autism among optimized machine learning models: An experimental analysis on genetic & personal characteristic datasets," *Appl. Sci.*, vol. 12, no. 8, p. 3812, Apr. 2022.
- [33] S. P. Kamma, S. Bano, G. L. Niharika, G. S. Chilukuri, & D. Ghanta, "Cost-effective & efficient detection of autism from screening test data using light gradient boosting machine," in *Intelligent Sustainable Systems*. Singapore: Springer, pp. 777–789, 2022.
- [34] U. Gupta, D. Gupta, & U. Agarwal, "Analysis of randomization-based approaches considering autism spectrum disorder," in *Pattern Recognition & Data Analysis among Applications*. Singapore: Springer, pp. 701–713, 2022.

- [35] T. Akter, M. Shahriare Satu, M. I. Khan, M. H. Ali, S. Uddin, P. Lio, J. M. W. Quinn, & M. A. Moni, “Machine learning-based models considering early stage detection of autism spectrum disorders,” *IEEE Access*, vol. 7, pp. 166509–166527, 2019.
- [36] Kaggle. (2022). Autism Spectrum Disorder Detection Dataset considering Toddlers. [Online]. Available: <https://www.kaggle.com/fabdelja/autism-screeningfor-toddlers>
- [37] UCI. (2022). UCI Machine Learning Repository: Autistic Spectrum Disorder Screening Data considering Adolescent Data Set. [Online]. Available: <https://shorturl.at/fhxCZ>
- [38] UCI. (2022). UCI Machine Learning Repository: Autism Screening Adult Data Set. [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Autism+Screening+Adult>
- [39] UCI. (2022). UCI Machine Learning Repository: Autistic Spectrum Disorder Screening Data considering Children Data Set. [Online]. Available: <https://shorturl.at/fiwLU>
- [40] D. Singh & B. Singh, “Investigating impact of data normalization on classification performance,” *Appl. Soft Comput.*, vol. 97, Dec. 2020, Art. no. 105524.
- [41] D. Mease, A. J. Wyner, & A. Buja, “Boosted classification trees & class probability/quantile estimation,” *J. Mach. Learn. Res.*, vol. 8, no. 3, pp. 409–439, 2007.
- [42] Q. Wang, W. Cao, J. Guo, J. Ren, Y. Cheng, & D. N. Davis, “DMP_MI: An effective diabetes mellitus classification algorithm on imbalanced data among missing values,” *IEEE Access*, vol. 7, pp. 102232–102238, 2019.
- [43] S. M. M. Hasan, M. A. Mamun, M. P. Uddin, & M. A. Hossain, “Comparative analysis of classification approaches considering heart disease prediction,” in *Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (ICME)*, Feb. 2018, pp. 1–4.
- [44] D. Ramesh & Y. S. Katheria, “Ensemble method based predictive model considering analyzing disease datasets: A predictive analysis approach,” *Health Technol.*, vol. 9, no. 4, pp. 533–545, Aug. 2019.
- [45] A. Arabameri & H. R. Pourghasemi, “Spatial modeling of gully erosion using linear & quadratic discriminant analyses in GIS & R,” in *Spatial Modeling in GIS & R considering Earth & Environmental Sciences*. Amsterdam, Netherlands: Elsevier, pp. 299–321, 2019.