International Journal of
HRM and Organizational Behavior

www.ijhrmob.com

editor@ijhrmob.com

International Journal of
HRM and Organizational Behavior
ISSN 2454 - 5015

EARLY-STAGE AUTISM SPECTRUM DISORDER DETECTION USING MACHINE LEARNING

G SWAPNA[1], THERU CHARANYALAKSHMI [2], MADHURI[3], K BHASKAR[4]
[1]Associate Professor Department of Pharmaceutics Apollo Institute of Pharmaceutical sciences, Email:swapnagv111@gmail.com
[2]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email: tcharanya3@gmail.com
[3]Assistant Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email:madhuri.yandra94@gmail.com
[4]Associate Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email: bhaskark.mca@gmail.com

Abstract: This project needs an ML algorithm to diagnose Autism Spectrum Disorder early. ASD is difficult to minimize, however the business tries to decrease stimulation early in implantation. Four feature selection (FS) programs—Quantile Transformer, Power Transformer, Normalizer, Max Abs Scaler—will be tested against four common ASD datasets from children to females. The scaled datasets will be processed using ML algorithms including AdaBoost, RF, DT, KNN, GNB, LR, SVM, LDA, etc. Most ideal decision classifier and FS plan each gathering of same status are driven using numerical assessments. The polling classifier envisions ASD accompanying the best accuracy in infants, children, adolescents, and adults. An in-depth examination of feature significance using a four-determinant solving scheme highlights the need of adjusting ML approaches to predict ASD across age groups and helps healthcare providers find alternatives to camouflage. Distinguished to existent methods for early discovery of ASD, the submitted scheme acts well. To further help the flexibility and veracity of ASD discovery, an ensemble approach utilizing voting classifiers accompanying RF and AdaBoost completed 100% accuracy.

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

1. INTRODUCTION

ASD is an early neurodevelopmental disorder of the brain that affects social interaction and collaboration. [1], [2] The continuum of ASD side effects and seriousness incorporates bound and repetitive standards of conduct [3], [4], [5]. There is no enduring therapy for ASD, albeit early intercession and great clinical consideration can work on a kid's way of behaving and correspondence capacities [6], [7], [8]. ASD location and determination are as yet extreme and complex using standard conduct research. Autism is frequently analyzed at two years of age however can be analyzed later relying upon seriousness [9], [10], [11]. ASD recognition might be done expeditiously with a few medicines. Before a high gamble of ASD, these demonstrative techniques are seldom performed.[46]

The authors of [12] created a brief, simple approach for toddlers, children, adolescents, and adults. [13] created the ASDTests smartphone app to quickly assess ASD using the results, Q-Talk, and AQ-10.

They sent the multi-use application data to the University of California, Irvine (UCI) ML repository and Kaggle for study in an open-source dataset. As of late, different investigations have utilized ML procedures to quickly evaluate and analyze ASD and different infections like diabetes, stroke, and cardiovascular breakdown [14], [15], [16].

[17] employed rule-based ML (RML) to examine ASD characteristics and found that RML enhanced accumulation accuracy. In [18], the authors used RF and ID3 algorithms to build prognostic models for children, adolescents, and adults. To address missing data, nonlinearity, and inconsistency, [19] proposed a new assessment tool that integrates ADI-R and ADOS-ML techniques and uses an alternative method for coding components. One more concentrate on mental thinking [13] utilized SVM, DT, and LR as pointers and prescient classifiers for ASD, featuring elements and classes and adding relationship values to featuring [17]. In [20], the creators examined TD (N = 19) and ASD cases (N = 11) utilizing relationship-based highlight decision making to assess the significance of elements. Seven standards were utilized to recognize 15 preschoolers with ASD in 2015 [21]. They also showed how group analysis may predict ASD numbers and variety from sophisticated models.Adult ASD prediction approaches include ANN, LR, LDA, CART, NB, and SVM [22].

2. LITERATURE SURVEY

We collected ASD datasets from toddlers, children, adolescents, and adults[1] using several object choice procedures. We then, at that point, applied various classifiers to these datasets and assessed their presentation utilizing prescient accuracy, kappa measurement, f1 measure, and AUROC. The presentation of individual classifiers was additionally tried utilizing non-parametric real significance tests. SVM performed better than other classifiers on infant, child, adolescent, and adult datasets with an accuracy of 97.82% for the RIPPER-based neonatal subgroup and 99.61% for correlation-based feature-based selection (CFS). The pediatric subgroup used Boruta CFS Intersect (BIC) method with an accuracy of 95.87% for the Boruta-based young adult subgroup and 96.82% for the CFS-based adult sub. The Shapley Additive Explains (SHAP) approach was applied to include the subgroups with the highest accuracy after validation [1].

Sequencing the 16S quality from dung has been utilized to concentrate on stomach microbiota and disease lately. ASD, a neurodevelopmental condition with gastrointestinal side effects, has dysbiotic stomach microbiota [2]. In spite of a few exploration, a normal dysbiotic profile in ASD people is as yet hard to characterize [3], [4], [5]. Specialized factors (trial philosophies) and outer components (dietary habits) make these investigations vary. We diminished concentrate on inclination by gathering 959 examples from eight activities (540 ASD and 419 Healthy Controls, HC). We made an ASD-HC indicator utilizing Machine Learning (ML). RF, SVM, and GB Machine were advanced. Five genera, including Parasutterella and Alloprevotella, meant quite a bit to every one of the three calculations. We showed the way that ML frameworks could find normal ordered characteristics by contrasting datasets from nations and secret frustrating elements.[48]

Autism diminishes correspondence, commitment, and conduct [4]. Understanding their visual tangible cycles makes sense of these. The study used image enhancement to analyze children's behavior by observing where they looked and when.[3,4,5,9] FEGP tests children's gaze direction to assess visual comprehension in clinically introverted and average children. The procedure utilizes a presentation level marker, representation, and deductions to recognize medically introverted children's visual conduct contrast and designer their learning projects to match their friends.

A few endeavors have been made to determine and evaluate neurological sicknesses to have hand quake side effects. Different sclerosis can be surveyed by hand quake seriousness. [5] A technique for recording and dissecting Spirography standard test computerized signals is introduced in this examination. We developed hardware and software for a device that performs a typical spirography test, captures signals, sends them to a computer using software, and analyzes them using feature extraction and characterization techniques. The program uses power spectrum analysis to demonstrate how each frequency component affects hand movements. Notwithstanding Power Spectrum Analysis, confounded markers like Biggest Lyapunov Example and mean Lyapunov range esteem are utilized to show signal chaoticity. Implanting aspect and delay structure a rough record window in occasional sign reproduction to reflect signal intricacy. Signal shape and examining rate influence delay. Incorporate space signals are requested by a pre-arranged feed forward mind association. [16,20] Request task is the unique cycle wherein the physicist figures each subject's sign's enlistment to strong and crippled social occasions and arranges matching medicines. In this paper, we show how elements like B. Disturbance parts can precisely recognize people with and without hand quake in ASD, a complex degenerative neuroplastic problem [6]. Most ASD identification techniques utilize functional magnetic resonance imaging (fMRI) with little datasets and give high precision yet low agreement rates [3, 4, 5]. In this review, we address this gap and propose an ASD detection strategy that leverages useful organizational elements of resting-state fMRI data to develop a computational model that provides evidence of psychological problems. New brain atlases Craddock 200 (CC200) and Automatic Anatomical Labeling (AAL), Bootstrap Analysis of Stable Clusters (BASC), and Power are used in our model. We provide a DNN classifier. Replication indicates the suggested model exceeds best approaches in accuracy. The suggested model has 88% accuracy, while the best techniques are 67%-85%. The proposed model has a sensitivity of 90%, a score of 87x, and an AUC of 96%. The BASC card book groups ASDs and manages better than other atlases that use different evaluation strategies.[50]

3. METHODOLOGY

i) Proposed Work:

Quantile Transformer, Power Transformer, MaxAbsScaler, and Normalizer optimize data and enhance early stage ASD diagnosis in the suggested ML architecture. The methodology analyzes multiple ASD datasets across age groups, focuses on key risk factors by identifying and optimizing features to create more accurate diagnostic models. Automation and complex pre-processing techniques improve ASD identity [3], [4], [5] and facilitate early reconciliation to improve results. An ensemble approach using RF and AdaBoost voting classifiers achieved 100% accuracy, further improving the versatility and accuracy of ASD detection. The group approach leverages various features of RF and Adaboost to provide more accurate predictions. Flask can provide

ISSN 2454 - 5015

a smooth and intuitive frontend for your client tests.[52]

ii) System Architecture:

A predictive model for autism diagnosis in distinct age groups is developed using ML methods in this work. To preprocess records after sorting, missing characteristics, add-on encoding, and oversampling are employed. MVI imputes missing dataset values. One Hot Encoding (OHE) transforms all add-on values into math characteristics. Eight ML methods—AB, RF, DT, KNN, GNB, LR, SVM, and LDA—sort the scaled dataset. Classifier characterisation determines the best assembly and FS strategy for each component scale ASD dataset. After these evaluations, the four FSTs (IGAE, GRAE, RFAE, and CAE) are used to prioritize relevant characteristics because ASD risk is unclear (see Table 4). Figure 1 recommends a study strategy for assessing ASD datasets and identifying the key ASD identification risk factors.

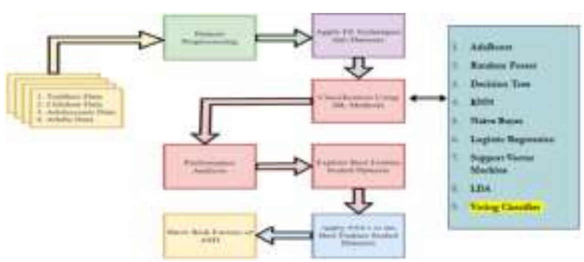

Fig 1 Proposed architecture

iii) Dataset collection:

This program stacks and examines ASD screening datasets for various age gatherings. It might include information structure checks, variable understanding, and dataset experiences.

1. Adult Screening Data: - Adult screening datasets are believed to be designed to assess ASD in adults [3], [4], [5]. A complete ASD screen may include adult-specific behaviors, communication skills, and other characteristics.

_Score	A10_Score	...	gender	ethnicity	jundice	austim	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: - The Toddler Dataset collects and analyzes information from children between the ages of 1 and 3. The dataset focuses on educational achievement, social bonding, and interpersonal skills, which are well-defined for this age group to detect autism syndrome at an early stage.

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: - The Adolescent dataset allows for the investigation of ASD in 12-18 year olds, which may represent issues in adolescent ASD, including social behavior, communication skills, etc.

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: - Child datasets from early infancy to preadolescence enable for the study of ASD-related aspects such developmental milestones, social interactions, and communication abilities [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 transforms raw information into business-helpful data. Information researchers accumulate, sort out, clean, check, break down, and orchestrate information into diagrams or papers. Data can be handled physically, precisely, or electronically. Data ought to be more significant and decision-production simpler. Organizations might upgrade activities and settle on basic decisions quicker. PC programming improvement and other mechanized information handling innovations add to this. Big data can be transformed into significant bits of knowledge for quality administration and independent direction.

v) Feature selection:

Feature selection chooses the most steady, non-repetitive, and pertinent elements for model turn of events. As data sets extend in amount and assortment, purposefully bringing down their size is significant. The fundamental reason for feature selection is to increment prescient model execution and limit processing cost.

One of the vital pieces of feature engineering is picking the main attributes for machine learning algorithms. To reduce the input elements, feature selection methods remove duplicated or redundant elements and limit the selection to those that are important to the ML model in general. Rather than permitting the ML model pick the main qualities, feature selection ahead of time enjoys a few benefits.

vi) Algorithms:

AdaBoost, or Adaptive Boosting, An ML approach that joins essential models further develops grouping exactness. It starts with a single-stage selection tree and iteratively develops new models, focusing more on information that was misclassified in previous models. AdaBoost can further develop credit card fraud detection by learning from earlier models and expanding execution by coordinating these models into serious areas of strength for a that can produce right expectations.[54]

```
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

Random Forest ensemble learning predicts utilizing a few choice trees. Training decision trees on irregular information subsets and averaging their forecasts works. This ensemble technique further develops characterization and relapse accuracy, takes out overfitting, and performs well [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

A Decision Tree is a tree-like model with center hubs addressing quality tests, branches addressing test results, and leaf hubs addressing class marks. Decision Trees show decision-production plainly. They can help recognize critical ASD prediction factors and are interpretable.

```
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

K-Nearest Neighbors is a non-parametric method that orders pieces of information by their k-nearest neighbors' element space greater part class. KNN helps find information designs without expecting a capability. It can find neighborhood ASD dataset relationships that may not be worldwide [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

Naive Bayes Bayes' hypothesis based probabilistic classifier accepts include autonomy. Naive Bayes is quick and great with high-layered datasets. Its straightforwardness and snappiness make it ideal for ASD information examination.

```
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

Logistic Regression: Logistic-based direct model for twofold order predicts class participation likelihood. Logistic Regression is interpretable and shows what qualities mean for ASD risk. It is a binary classification baseline.

```
# 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 Vector Machine Supervised learning methods find an ideal hyperplane to split the classes into a space with many layers. SVM handles convoluted choice limits well. It might further develop ASD arrangement by catching non-linear accuracy [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

Linear Discriminate Analysis dimensionality decrease and grouping technique that distinguishes straight component blends that ideally partition classes. [23,26] LDA diminishes dimensionality and features recognizing attributes. It further develops interpretability and may assist with distinguishing ASD discovery factors.

```
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

A Voting Classifier, Numerous classifiers are educated and coordinated to make a last forecast in joining ensemble learning. This venture involves AdaBoost and Random Forest as basis classifiers.

```
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)
eclf1 = VotingClassifier(estimators=[('ab', clf1), ('rf', clf2), ('dt', clf3)], voting='soft')
eclf1.fit(X_train, y_train)
y_pred = eclf1.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: Precision measures the fraction of a specific occurrence or test that is well-characterized. Precision is attained using the following formula:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

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

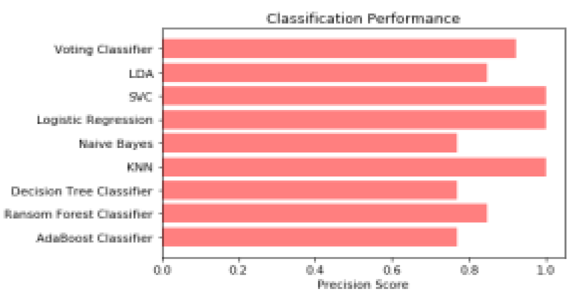

Fig 15 Precision comparison graph

Recall: ML recall evaluates a model's ability to capture all class events. The proportion of precisely predicted positive predictions that result in real benefits measures a model's ability to detect a certain type of occurrence.

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

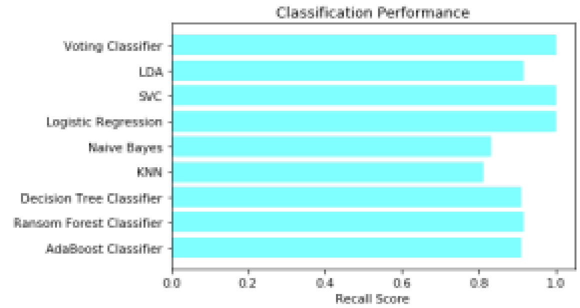

Fig 16 Recall comparison graph

Accuracy: A model's accuracy is defined as the proportion of correct predictions in grouping positions.

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

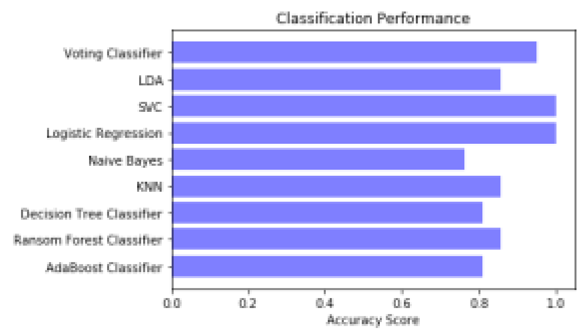

Fig 17 Accuracy graph

F1 Score: The F1 score detects both false positives and false negatives, giving it a consistent accuracy and validation approach for unbalanced datasets.

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

ISSN 2454 - 5015

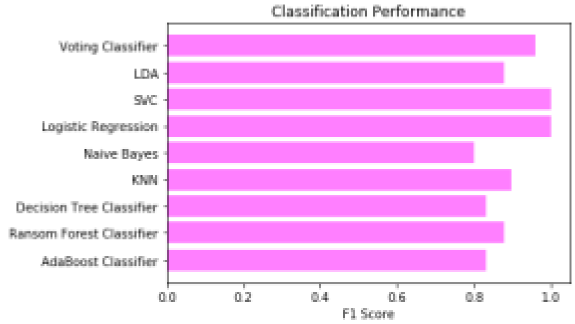

Fig 18 F1Score

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	0.945	0.94	0.984	0.922
Naive Bayes	0.979	0.96	0.980	0.978
Logistic Regression	0.993	0.98	1.000	0.998
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

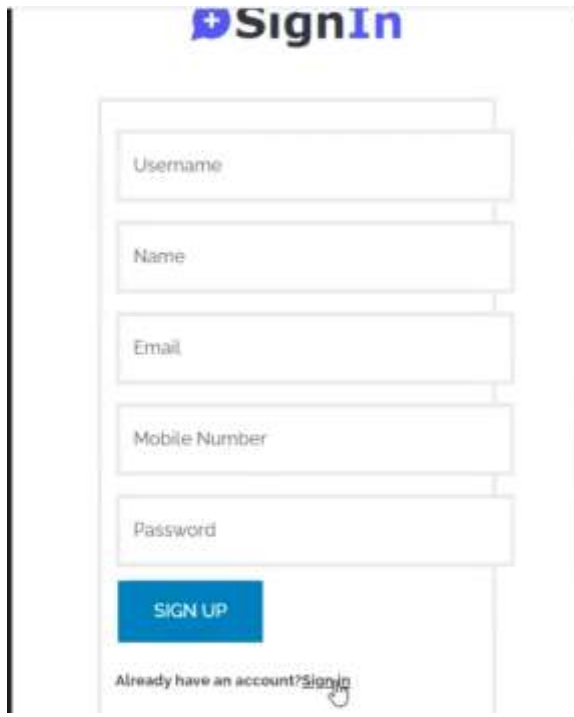

Fig 21 Signin page

Fig 22 Login page

ISSN 2454 - 5015

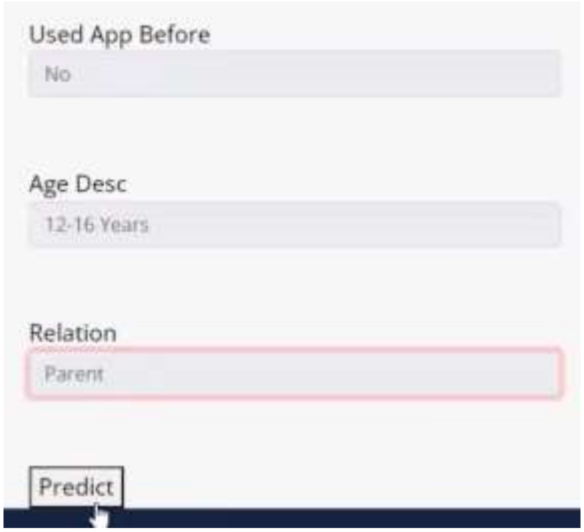

Fig 23 User input

Fig 24 Predict result for given input

5. CONCLUSION

An original ML framework has been produced for early identification of ASD utilizing complex calculations and component scaling strategies. The approach performs well across all age groups in relevant ASD datasets of toddlers, adolescents, children, and adults, demonstrating its adaptability and therapeutic relevance [12,13]. Ideal sequences and factor scaling algorithms for early ASD identification and quick treatment are identified by the framework. The accuracy of ASD detection was further improved by outfitting approach using RF and AdaBoost. The user-friendly front-end Flask framework that allows variable properties to be deployed and evaluated demonstrates its suitability and convenience in real-world applications. The test involves using judgement to identify important risk factors and develop persuasiveness to support accurate diagnosis of ASD.

6. FUTURE SCOPE

The task plans to gather further information on ASD and foster a more complete prognostic model for all age gatherings to work on the proof for ASD and other neuroplastic messes [18]. This study can involve more ASD patients and diversity. The research proposes using ML to create more extensive prediction models and update the system to improve ASD detection accuracy and reliability. The study may examine how the suggested approach might detect and forecast additional neuroplastic diseases. More data will be collected in this study, the model will be refined, and it may be expanded to other neuroplastic disorders.

REFERENCES

[1] M. Bala, M. H. Ali, M. S. Satu, K. F. Hasan, and M. A. Moni, ‘‘Efficient machine learning models for 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, and G. Chillemi, ‘‘Machine learning data analysis highlights the role of parasutterella and alloprevotella in autism spectrum disorders,’’ Biomedicines, vol. 10, no. 8, p. 2028, Aug. 2022.

[3] R. Sreedasyam, A. Rao, N. Sachidanandan, N. Sampath, and S. K. Vasudevan, ‘‘Aarya—A kinesthetic companion for children with autism spectrum disorder,’’ J. Intell. Fuzzy Syst., vol. 32, no. 4, pp. 2971–2976, Mar. 2017.

[4] J. Amudha and H. Nandakumar, ‘‘A fuzzy based eye gaze point estimation approach to study the task behavior in autism spectrum disorder,’’ J. Intell. Fuzzy Syst., vol. 35, no. 2, pp. 1459–1469, Aug. 2018.

ISSN 2454 - 5015

[5] H. Chahkandi Nejad, O. Khayat, and J. Razjouyan, ‘‘Software development of an intelligent spirography test system for neurological disorder detection and quantification,’’ J. Intell. Fuzzy Syst., vol. 28, no. 5, pp. 2149–2157, Jun. 2015.

[6] F. Z. Subah, K. Deb, P. K. Dhar, and T. Koshiba, ‘‘A deep learning approach to 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, and G. F. Fragulis, ‘‘The contribution of machine learning and 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, and M. Alshahrani, ‘‘Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques,’’ Electronics, vol. 11, no. 4, p. 530, Feb. 2022.

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

[10] S. P. Abirami, G. Kousalya, and R. Karthick, ‘‘Identification and exploration of facial expression in children with 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, and 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, 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.

[13] 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.

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

[15] E. Dritsas and M. Trigka, ‘‘Stroke risk prediction with machine learning techniques,’’ Sensors, vol. 22, no. 13, p. 4670, Jun. 2022

[16] V. Chang, J. Bailey, Q. A. Xu, and 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 and 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, 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.

[19] 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.

[20] 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.

[21] 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.

[22] B. Tyagi, R. Mishra, and N. Bajpai, ‘‘Machine learning techniques to predict autism spectrum disorder,’’ in Proc. IEEE Punecon, Jun. 2019, pp. 1–5.

[23] F. Thabtah and D. Peebles, ‘‘A new machine learning model based on induction of rules for autism detection,’’ Health Informat. J., vol. 26, no. 1, pp. 264–286, Mar. 2020.

[24] 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, pp. e732–e732, Feb. 2016.

[25] S. B. Shuvo, J. Ghosh, and A. S. Oyshi, ‘‘A data mining based approach to 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 and M. Ulas, ‘‘Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children,’’ in Proc. 6th Int. Symp. Digit. Forensic Secur. (ISDFS), Mar. 2018, pp. 1–4.

[27] F. N. Buyukoflaz and A. Ozturk, ‘‘Early autism diagnosis of children with 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, and M. F. Harun, ‘‘Classification of adults with 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, and 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 and T. Padma, ‘‘An ensemble classification and binomial cumulative based PCA for diagnosis of Parkinson’s disease and autism spectrum disorder,’’ Int. J. Syst. Assurance Eng. Manage., early access, pp. 1–16, Jul. 2022.

[31] R. Abitha, S. M. Vennila, and I. M. Zaheer, ‘‘Evolutionary multiobjective optimization of artificial neural network for classification of autism spectrum disorder screening,’’ J. Supercomput., vol. 78, no. 9, pp. 11640–11656, Jun. 2022.

[32] M. Alsuliman and H. H. Al-Baity, ‘‘Efficient diagnosis of autism with optimized machine learning models: An experimental analysis on genetic and 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, and D. Ghanta, ‘‘Cost-effective and 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, and U. Agarwal, ‘‘Analysis of randomization-based approaches for autism spectrum disorder,’’ in Pattern Recognition and Data Analysis with 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, and M. A. Moni, ‘‘Machine learning-based models for early stage detection of autism spectrum disorders,’’ IEEE Access, vol. 7, pp. 166509–166527, 2019.

[36] Kaggle. (2022). Autism Spectrum Disorder Detection Dataset for Toddlers. [Online]. Available: https://www.kaggle.com/fabdelja/autism-screeningfor-toddlers

[37] UCI. (2022). UCI Machine Learning Repository: Autistic Spectrum Disorder Screening Data for 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 for Children Data Set. [Online]. Available: https://shorturl.at/fiwLU

[40] D. Singh and B. Singh, ‘‘Investigating the impact of data normalization on classification performance,’’ Appl. Soft Comput., vol. 97, Dec. 2020, Art. no. 105524.

[41] D. Mease, A. J. Wyner, and A. Buja, ‘‘Boosted classification trees and 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, and D. N. Davis, ‘‘DMP_MI: An effective diabetes mellitus classification algorithm on imbalanced data with missing values,’’ IEEE Access, vol. 7, pp. 102232–102238, 2019.

[43] S. M. M. Hasan, M. A. Mamun, M. P. Uddin, and M. A. Hossain, ‘‘Comparative analysis of classification approaches for heart disease prediction,’’ in Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (ICME), Feb. 2018, pp. 1–4.

[44] D. Ramesh and Y. S. Katheria, ‘‘Ensemble method based predictive model for analyzing disease datasets: A predictive analysis approach,’’ Health Technol., vol. 9, no. 4, pp. 533–545, Aug. 2019.

[45] A. Arabameri and H. R. Pourghasemi, ‘‘Spatial modeling of gully erosion using linear and quadratic discriminant analyses in GIS and R,’’ in Spatial Modeling in GIS and R for Earth and Environmental Sciences. Amsterdam, The Netherlands: Elsevier, pp. 299–321, 2019.

[46] G.Viswanath, “Hybrid encryption framework for securing big data storage in multi-cloud environment”, Evolutionary intelligence, vol.14, 2021, pp.691-698.

[47] Viswanath Gudditi, “Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage”, Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol.12, 2021, pp.545-552.

[48] Viswanath Gudditi, “A Smart Recommendation System for Medicine using Intelligent NLP Techniques”, 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), 2022, pp.1081-1084.

[49] G.Viswanath, “Enhancing power unbiased cooperative media access control protocol in manets”, International Journal of Engineering Inventions, 2014, vol.4, pp.8-12.

[50] Viswanath G, “A Hybrid Particle Swarm Optimization and C4.5 for Network Intrusion Detection and Prevention System”, 2024, International Journal of Computing, DOI: https://doi.org/10.47839/ijc.23.1.3442, vol.23, 2024, pp.109-115.

[51] G.Viswanath, “A Real Time online Food Ording application based DJANGO Restfull Framework”, Juni Khyat, vol.13, 2023, pp.154-162.

[52] Gudditi Viswanath, “Distributed Utility-Based Energy Efficient Cooperative Medium Access Control in MANETS”, 2014, International Journal of Engineering Inventions, vol.4, pp.08-12.

[53] G.Viswanath,“ A Real-Time Video Based Vehicle Classification, Detection And Counting System”, 2023, Industrial Engineering Journal, vol.52, pp.474-480.

[54] G.Viswanath, “A Real- Time Case Scenario Based On Url Phishing Detection Through Login Urls ”, 2023, Material Science Technology, vol.22, pp.103-108.

[55] Manmohan Singh,Susheel Kumar Tiwari, G. Swapna, Kirti Verma, Vikas Prasad, Vinod Patidar, Dharmendra Sharma and Hemant Mewada, “A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification” published in Journal of Computer Science, Available at: https://pdfs.semanticscholar.org/69ac/f07f2e756b79181e4f1e75f9e0f275a56b8e.pdf