INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

Received: 13-10-2024

Accepted: 08-11-2024

Published: 12-12-2024

Citation: Ramya R, Arokiaraj SP (2024) Integrated Decision Support System (IDSS) for Autism Spectrum Disorder Diagnosis: A Multi-model F ramework Approach. Indian Journal of Science and Technology 17(45): 4787-4797. https://doi.org/10.17485/IJST/v17i45.3348

* **Corresponding author**.

rgramya94@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Integrated Decision Support System (IDSS) for Autism Spectrum Disorder Diagnosis: A Multi-model F ramework Approach

R Ramya[1]*, S Panneer Arokiaraj[1]

1 Department of Computer Science, Thanthai Periyar Govt. Arts & Science College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India

Abstract

Objectives: Development of an Integrated Decision Support System (IDSS) for Autism Spectrum Disorder (ASD) diagnosis using advanced Data Science techniques. **Methods:** A IDSS multi-model approach combining GBT-DL, VO-ADA, and CNN-RF were used. In evaluation, the following parameters were computed: accuracy, precision, recall, F1-score, and Kappa statistics. **Findings:** The results show that the GBT-DL model performed with excellent accuracy of 95.52% performance, while the VO-ADA model achieved an accuracy level of 98.60%, and the CNN-RF model has proven to have the highest accuracy at 99.15%. The IDSS demonstrates a high predictive capability across a broad age range and gives well-accounted estimates of autism severity. Multiple Data Science approaches allowed the system to surpass the existing diagnostic techniques through an aggregate ASD diagnosis. The IDSS outperformed other state-of-the-art methods based on accuracy, recall, F1-score, and Kappa statistics, which indicate its robustness and reliability. Dependence of the IDSS on contingent information reduces dependency on conventional diagnostic tests while raising the accuracy of diagnosis. Such sophisticated models result in the timely and accurate diagnosis of individuals across the autism spectrum, add substantial value to the established diagnostic methods, and are a very significant improvement over previously applied approaches. **Novelty:** The IDSS integrates multiple new state-of-the-art models for precise and efficient diagnosis of ASD across a vast range of ages.

Keywords: Autism Spectrum Disorder (ASD); Convolutional Neural Networks-Random Forest (CNN-RF); Data Science Models; Gradient Boosted Trees-Deep Learning (GBT-DL); Integrated Decision Support System (IDSS); Vote-AdaBoost (VO-ADA)

1 Introduction

Autism Spectrum Disorder (ASD) is a heterogeneous neurodevelopmental disorder, characterized by persistent deficits in social interaction across multiple contexts and restricted patterns of behavior, interests, or activities. Accurate and timely diagnosis of ASD is one of the most important and challenging aspects of diagnostic practice due to the high level of heterogeneity in symptom presentation and severity. Although advances in diagnostics are made, existing techniques are limited to subjective methods focused on classical screening tools and fail to recognize subtle signatures in ASD data. This urgency has led to the search for more reliable and less subjective methods, such as Data Science and deep learning techniques that can recognize complex patterns on large datasets and provide accuracy to the diagnosis. On the other hand, the current studies in ASD diagnosis using Data Science and deep learning, bear some constraints. Nguyen and Rodriguez highlighted the performance of AdaBoost models for screening individuals with ASD through meta-analysis and clinical significance, yet their investigation was confined to one type of model and lacked a multi-model strategy adaptable to different age and diagnostic complexities [1,2]. Following these, Sue Miller and Torres have compared known techniques for ASD prediction, but their study is limited to individual model strategies and does not benefit from ensembles as these approaches are known to enhance prediction power [3]. Bell and Clark recommended ensemble methods in their research and discussed the advantages of combining models, but they did not address the issue of scalability and adaptability to different populations [4]. However, there is still a need in the field for a unified, multimodal framework that brings together several models for both the specific age range affected by ASD and for which diagnosis is acceptable.

The potential of multi-model frameworks for tackling such diagnostic challenges is further emphasized by recent advances in ML algorithms. Interestingly, Nguyen and Miller summarized ensemble models for ASD prediction in children and adults, supporting all-inclusive systems yet no specific models per application for different ages [5,6]. Clark and Taylor have explored the issue of adaptive model selection for voting and AdaBoost models, but they did not provide an integrative solution covering multiple age categories of ASD cases [7]. In addition, Parlett-Pelleriti et al. An unsupervised ML in ASD research and Bala et al. show the potential of non-traditional approaches but are ultimately restricted by the specificity of their practical applicability in age-specific diagnostic cases [8,9].

This paper proposes a new Integrated Decision Support System (IDSS) for ASD diagnosis to fill in these gaps. There are several multi-models in this system, such as combining Gradient Boosted Tree-Deep Learning (GBT-DL), Voting-AdaBoost (VO-ADA), and Convolutional Neural Network-Random Forest (CNN-RF) models. The three models are intended for maximized directability of diagnostic scoring in each of the specific age groups: GBT-DL for adults, VO-ADA for children, and CNN-RF for toddler age groups. The IDSS proposed here does so by capturing the strengths of these models into an overarching framework that aims to improve the precision, adaptability and efficiency of ASD diagnosis. It bridges the research gaps identified above, and it overcomes limitations of existing methodologies; therefore, it provides healthcare professionals with a functional framework for optimal, age-appropriate, and nuanced diagnostic support.

The following sections discuss the approach by presenting the data and methods for the development of the IDSS framework. The results and discussion part contains experimental results where we compare the performance of proposed technique with earlier techniques and shown improvements. And then, the conclusion summarizes these insights, highlighting the implications of an IDSS for the outcomes of diagnosis and intervention in relation to ASD.

2 Methodology

This research work develops an Integrated Decision Support System (IDSS) with an aim to improve the accuracy and speed of diagnosis of ASD over different age groups. The IDSS is unique in that IDSS integrates three independent model types, each type tailored for diagnosis across age groups: Gradient Boosted Tree-Deep Learning (GBT-DL) for adults, Voting-AdaBoost (VO-ADA) for children, and Convolutional Neural Network-Random Forest (CNN-RF) for toddlers. What is new here is how these models are combined. The IDSS is built to reflect the age-conditional data properties and diagnostic requirements instead of a one-size-fits-all diagnostic model. Each model in the framework performs a specialized function:

- GBT-DL outshines the rest in accuracy classification and error minimization on adults diagnosed with autism, effectively utilizing both Gradient Boosted Trees and Deep Learning.
- VO-ADA is specifically designed for predicting autism in children and it utilizes voting mechanisms and AdaBoost algorithms to enhance both accuracy and precision in classification on children-based datasets.
- CNN-RF to predict toddler autism, capitalizing on the power of Convolutional Neural Networks and Random Forest to provide accurate predictions and low classification errors for the toddler population.

Collectively this multi-model framework is a significant improvement compared to past efforts in that diagnostic tools are appropriate for the specific challenges facing younger models which has not been appropriately handled in prior research. This IDSS offers a diagnostic tool that is more nuanced, contextualized, and specific to the nuanced age-based differences in the manifestation of ASD than traditional approaches, which often employ broader, less-contextualized models. The procedure for building the IDSS framework consists of (i) Data preprocessing and feature selection, (ii) Training and validation of the models for different age groups, and (iii) final integration of the models into a single decision support system. This approach represents a state-of-the-art step forward in the development of diagnostic support for ASD and by tailoring model choice and use to the specific diagnostic requirements of the different age groups it provides a more encompassing, comprehensive tool.

2.1 Dataset

The Kaggle Autism Child Dataset[10] is a special dataset referring to Autism Spectrum Disorder (ASD) that was compiled specifically for use in predictive modeling. Table 1 describes several characteristics such as the child age, sex, race, family history of Pervasive Developmental Disorder (PDD), screening method, and screening question response. This dataset aims to provide essential information for research and development of some of the better models for ASD prediction using data science models specifically hybrid models like multi-model approach.

Table 1. Attributes of the Autism Dataset

Attributes	Description
Age	Age of an Individual
Gender	Sex of the Person
Ethnicity	Enumerated below are many prevalent ethnicities, given in a textual manner.
Born with Jaundice	Indicate if the person has jaundice from birth (yes or no).
Family Member with PDD	Does any immediate family member have a PDD? Please answer with either ”yes” or ”no”.
Who is Completing the Test	Individual responsible for performing the ASD screening test
Country of Residence	Country
Used the Screening App Before	Has the user already used a screening app?
Screening Method Type	The screening technique type is determined by the age group
Question 1 to Question 10 Answer	The binary replies (0 or 1) correspond to particular questions in accordance with the screening procedure used.
Screening Score	The ultimate score obtained by the automated computation of the specified screening method's scoring algorithm.
A1_Score,..., A10_Score	Individual scores for each of the ten questions.
Age Description	Description of the age category (e.g., '18 and more').
Relation	Relationship of the person completing the test (e.g., Parent, Self, Health care professional).
Class/ASD	The target variable indicating the presence or absence of Autism Spectrum Disorder (ASD) (e.g., YES or NO).

2.2 Workflow Architecture

Figure 1 illustrates the workflow of the proposed models for ASD prediction. It goes through a workflow of loading a dataset, and data preprocessing to clean and structure the data. The dataset is then explored, and validation is performed, this is splitting the data into training and testing data and k-fold validation. During training, two hybrid base models are constructed which are then fitted inside the model. The test section includes hybrid model. Lastly, the trained models; GBT-DL, VO-AD, and CNN-RF models are applied to predict the likelihood of a person developing ASD.

2.3 Proposed IDSS Models

Three hybrid models are proposed in the IDSS model section which include Gradient Boosted Trees-Deep Learning (GBT-DL), Vote-AdaBoost (VO-ADA) and Convolutional Neural Networks-Random Forest (CNN-RF). They are all focused on predicting autism in certain age ranges. Overall, GBT-DL, VO-ADA, and CNN-RF give specific benefits and the most accurate autistic prediction, and the least classification error between age groups.

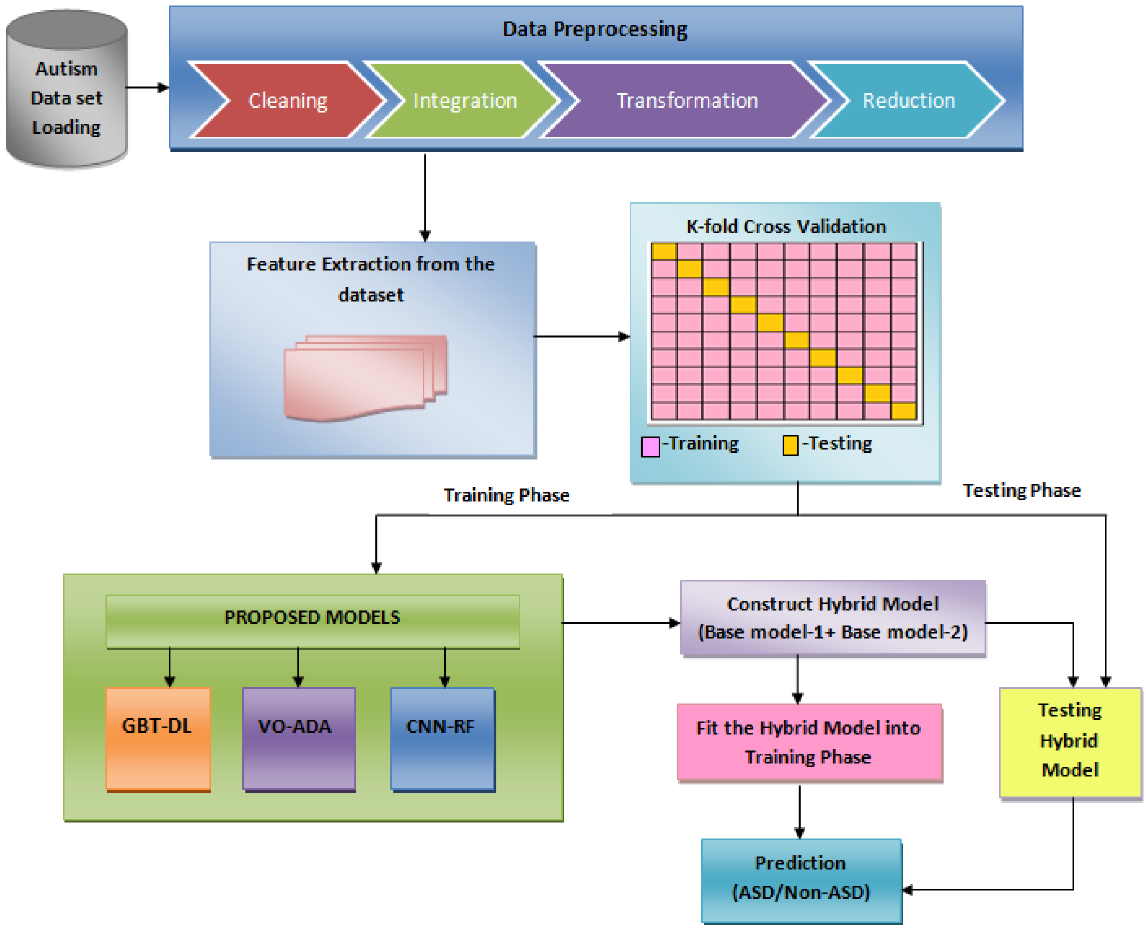

Fig 1. General Workflow Architecture of the Proposed IDSS Models

2.3.1 GBT-DL Model

The algorithm below describes how the hybrid model using GBT and NNs is built: first step the training dataset is divided into a training set and a validation set; Second step is GBT and NN are constructed on the training set based on GBT's best hyperparameters to minimize the cost function and the NN's initial hyperparameters, thus creating two different models; third step for GBT and NN predetermine the label on the validation set; A set of training features is restored from these predictions by stacking. Fourth step is a meta-model is built based on these new training features of the label using the GBT and NN design on the entire training dataset; next step is each model generated the label over the entire training dataset on the test dataset; the features will be generated again by stacking; The final label generated by the final meta-model over the entire test dataset is a hybrid model[11].

Algorithm: GBT-DL Model (Gradient Boosted Tree and Deep Learning)

Input:

- Training dataset (X_train, y_train)
- Testing dataset (X_test)
- GBT hyperparameters (gbt_params)
- Deep Learning hyperparameters (DL_params)

Output:

- Autism predictions on the dataset

1. Load the ASD Dataset
2. Preprocess the Autism Spectrum Disorder Dataset
• Normalize the data: Let X represent the dataset, and Xnorm be the normalized dataset.
• Split the dataset into training and testing sets: let Xtrain and Xtest represent the training and testing sets respectively.
3. Split the training dataset:$(Xtrain, ytrain) \rightarrow (Xval, yval)$.
4. Train the GBT model:$fGBT = TrainGBT(Xtrain, ytrain, gbt\_params)$
5. Train the NN model :$fNN = TrainNN(Xtrain, ytrain, nn\_params)$
6. Generate predictions on the validation set:
$\hat{y}val\_GBT = fGBT(Xval)$
$\hat{y}val\_NN = fNN(Xval)$
7. Stack predictions :$Xval\_stacked = [\hat{y}val\_GBT, \hat{y}val\_NN]$
8. Train a meta-model: $fmeta = TrainMetaModel(Xval\_stacked, yval)$
9. Train the GBT and NN models:
$fGBT\_final = TrainGBT(Xtrain, ytrain, gbt\_params)$
$fNN\_final = TrainNN(Xtrain, ytrain, nn\_params)$
10. Generate predictions on the testing dataset:
$\hat{y}test\_GBT = fGBT\_final(Xtest)$
$\hat{y}test\_NN = fNN\_final(Xtest)$
11. Stack the predictions for the testing dataset:
$Xtest\_stacked = [\hat{y}test\_GBT, \hat{y}test\_NN]$
12. final predictions using meta-model:
$\hat{y}test\_final = fmeta(Xtest\_stacked)$
13. Output the final predictions from the hybrid model $\hat{y}test\_final$

2.3.2 VO-ADA Model

The VO-ADA algorithm blends Voting and AdaBoost models to forecast Autism Spectrum Disorder (ASD) outcomes from the dataset. It begins by loading the ASD dataset, segregating features (X) and the target variable (y). Data is then split into training and testing sets using train_test_split, crucial for unseen data evaluation. Two DecisionTree Models form base Models for ensemble diversity. An AdaBoost model is established with base_Model_1 and 50 estimators, emphasizing hard-to-classify instances. A Voting Model merges predictions from base Models and AdaBoost using soft voting, considering class probability estimates. The ensemble model, including the voting Model, is trained on the training data. In below algorithm explains overall, VO-ADA harnesses both voting and AdaBoost techniques to create a robust ASD status predictor from the dataset.

Algorithm: VO-ADA (Hybrid model for Vote and AdaBoost)
Input data: AutismSpectrumDisorder Dataset
Output: Prediction (ASD/NON-ASD)
1. Load the ASD Dataset
2. Preprocess the Autism Spectrum Disorder Dataset
• Normalize the data: Let X represent the dataset, and Xnorm be the normalized dataset.
• Split the dataset into training and testing sets: let Xtrain and Xtest represent the training and testing sets respectively.
3. Define X & y are features and target variable
4. Split data into training & testing sets
5. X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
6. Define base classifiers for the ensemble
7. base_classifier_1 = DecisionTreeClassifier(max_depth=3)
8. base_classifier_2 = DecisionTreeClassifier(max_depth=5)
9. Create an AdaBoost classifier with a decision tree as base estimator
10. adaboost_classifier = AdaBoostClassifier(base_classifier_1, n_estimators=50)
11. Create a voting classifier
12. voting_classifier = VotingClassifier(estimators=[('dt1', base_classifier_1), ('dt2', base_classifier_2), ('adaboost', adaboost_classifier)], voting='soft')
13. Fit the ensemble model on the training data
14. voting_classifier.fit(X_train, y_train)
15. Prediction on the test set

16. y_pred = voting_classifier.predict(X_test)
17. Evaluate accuracy of the combined model
18. accuracy = accuracy_score(y_test, y_pred)

2.3.3 CNN-RF Model

In this regard, CNN-RF is a hybrid model so as to well analyze and predict the results on the Autism Spectrum Disorder (ASD) dataset. The process in the below algorithm starts with loading the ASD dataset, then the preprocessing steps are undertaken such as normalizing the data and splitting into train and test sets. Then a CNN is trained on the training data, having its architecture defined and a model compiled and its parameters optimal to minimize loss. The CNN is now assessed using the testing data set, which offers predicted outputs for examination. Then, we train a Random Forest Model using these features by tuning hyperparameters, if required. We evaluate the hybrid model performance by classifying the extracted features using the RF Model, and then the evaluation parameters are calculated. The evaluation results are printed, which completes the execution of the algorithm. This algorithm is expected to provide accurate prediction results for ASD analysis by combining the feature extraction functionality of CNNs with the stability that RF Models possess to potentially provide effective predictions from the extracted features [12].

Algorithm: CNN-RF Algorithm (Hybrid model for Convolutional Neural Network and Random Forest)
Input:data:AutismSpectrumDisorderDataset
Output:Prediction of Analysis
Step 1: Start the execution
Step 2: Load the ASD Dataset
Step 3: Preprocess the Autism Spectrum Disorder Dataset
• *Normalize the data: Let X represent the dataset, and* X_{norm} *be the normalized dataset.*
• *Split the dataset into training and testing sets: let* X_{train} *and* X_{test} *represent the training and testing sets respectively.*
Step 4: Train the Convolutional Neural Network (CNN)
• *Define CNN architecture: Let* f_{CNN} *represent the architecture of the CNN.*
• *Compile the CNN model: Let L be the loss function,* Θ *represent the parameters of the CNN, and* η *represent the optimizer.*
• *Train the CNN model using the training data:*
$\ominus^* = argmin_{\ominus} \frac{1}{n} \sum_{i=1}^{n} L(f_{CNN}\left(X_{train}^{(i)};\ominus\right), y_{pred}^{(i)})$ *Where n is the number of training samples.*
• *Evaluate the CNN model on the testing data: Let*
$y_{pred}^{(i)} = f_{\ CNN}\left(X_{train}^{(i)};\ominus^*\right)$ *Represent the predicted outputs for testing samples.*
Step 5: Extract Features using the Trained CNN
• *Use the trained CNN model to extract features from the testing data:*
$F_{CNN} = (f1, f2, \ldots, fm\}$*Represent the extracted features from the testing data.*
Step 6: Train the Random Forest (RF) Classifier
• *Train a Random Forest classifier using the features extracted by the CNN: Let* f_{RF} *represent the Random Forest classifier.*
• *Tune hyper parameters of the Random Forest classifier if necessary.*
Step 7: Evaluate the Hybrid CNN-RF Model
• *Make predictions using the trained RF classifier: Let* $y_{pred}^{(i)} = f_{RF}\,(F_{\ CNN}\left(X_{test}^{(i)}\right))$ *represent the predicted outputs using the RF classifier.*
• *Calculate evaluation metrics such as accuracy, precision, recall, F1-Score, and kappa statistic:* $Metrics = \{accuracy, precision, recall, F1-Score, kappa\}$*represent the evaluation metrics.*
• *Print or display the evaluation metrics: Let* $Eval(Metrics)$ *represent the evaluation results.*
End

3 Results and Discussion

As part of the experimentation, Google Collaboratory which is a cloud service provided by Google, was used to conduct the experiments and scikit-learn package in Python was used for data preprocessing, feature extraction, and prediction. The proposed methodology used a 10-fold cross-validation on three different ASD datasets including Toddlers, Children and Adults. From these results, as with the 10-fold cross-validation process used to build and validate this IDSS model, the data set was randomly split into 10 equal parts from which the data upon which a prediction should be made trained on in 9 parts learned in each iteration. Repeat this process 10 times and aggregate the results average. This research work chose 10-fold cross-validation to mitigate overfitting, decrease variance, and ensure effective model generalization due to the limited dataset size. The evaluation of experimental outcomes involved various statistical measures using predefined formulas Equations (1), (2), (3), (4) and (5).

$$Accuracy = \frac{TP+TN+FP+FN}{TP+TN} \tag{1}$$

$$Precision = \frac{TP+FP}{TP} \tag{2}$$

$$Recall = \frac{TP+FN}{TP} \tag{3}$$

$$F1-Score = \frac{2\times Precision \times Recall}{Precision+Recall} \tag{4}$$

$$Kappa = \frac{P_o - P_e}{1-P_e} \tag{5}$$

3.1 Performance Analysis of IDSS Based on Accuracy

howcases the accuracy scores of different models on varied datasets with feature scaling. It contains accuracy percentages attained by each hybrid model across various K-fold cross-validation settings for each age group dataset.

Table 2. Accuracy Analysis of Proposed models

Dataset	K-fold Cross Validation	GBT-DL	VO-ADA	CNN-RF
Toddler	K=3	72.45	79.32	89.45
	K=5	73.67	84.89	95.56
	K=10	74.91	88.01	99.15
Child	K=3	87.67	88.01	79.54
	K=5	88.23	93.89	75.32
	K=10	89.55	98.60	78.75
Adult	K=3	89.67	88.01	84.62
	K=5	92.67	89.76	85.64
	K=10	95.52	89.96	88.71

3.2 Performance Analysis of IDSS Based on Precision

Precision is a measure of how accurate positive predictions are. It tells us the percentage of genuine positive predictions out of all the positive forecasts made. The term "precision" refers to the ratio of true positives to the total of true positives and false positives.

Table 3. Precision Analysis of Proposed models

Dataset	K-fold Cross Validation	GBT-DL	VO-ADA	CNN-RF
Toddler	K=3	87.23	78.57	88.00
	K=5	87.98	81.41	94.26
	K=10	88.25	87.78	97.89
	K=3	71.23	87.78	78.08
Child				

Continued on next page

Table 3 continued

	K=5	73.98	92.12	74.85
	K=10	73.43	99.50	77.89
	K=3	90.13	87.28	84.00
Adult	K=5	92.62	88.34	84.26
	K=10	94.88	89.72	87.90

In Table 3, Across the Toddler dataset, GBT-DL demonstrates precision rates between 87.23% and 88.25%, while VO-ADA ranges from 78.57% to 87.78%, and CNN-RF achieves precision ranging from 88.00% to 97.89%. For the Child dataset, GBT-DL shows precision between 71.23% and 73.43%, VO-ADA ranges from 87.78% to 99.50%, and CNN-RF achieves precision rates between 74.85% and 77.89%. In the Adult dataset, GBT-DL exhibits precision rates ranging from 90.13% to 94.88%, VO-ADA ranges from 87.28% to 89.72%, and CNN-RF ranges from 84.00% to 87.90%. These results indicate variability in precision across different models and datasets, with certain models consistently outperforming others across various cross-validation folds and datasets. Additionally, to evaluate the precision values of different hybrid models on ASD datasets for Toddlers, Children, and Adults, as illustrated in Figure 1.

3.3 Performance Analysis of IDSS Based on Recall

Recall is a metric that quantifies the model's capacity to accurately detect all relevant events. It refers to the ratio of true positives to the total of true positives and false negatives. The recall values of different hybrid models on ASD for Toddlers, Children, and Adults are compared in the table.

Table 4. Recall Analysis of Proposed models

Dataset	K-fold Cross Validation	GBT-DL	VO-ADA	CNN-RF
	K=3	87.12	69.92	88.09
Toddler	K=5	87.89	86.67	94.48
	K=10	88.26	82.36	97.18
	K=3	71.46	82.36	78.15
Child	K=5	73.12	89.93	74.54
	K=10	72.57	96.32	77.18
	K=3	90.56	82.87	84.09
Adult	K=5	92.19	84.76	84.48
	K=10	95.48	86.67	87.98

The Table 4 showcases the recall analysis outcomes for GBT-DL, VO-ADA, and CNN-RF models across different datasets (Toddler, Child, Adult) using varied K-fold cross-validation techniques (K=3, K=5, K=10). For the Toddler dataset, GBT-DL demonstrates recall rates ranging from 87.12% to 88.26%, while VO-ADA shows recall rates between 69.92% and 82.36%, and CNN-RF achieves recall rates ranging from 88.09% to 97.18%.In the Child dataset, GBT-DL ranges from 71.46% to 72.57% recall, VO-ADA exhibits recall between 82.36% and 96.32%, and CNN-RF ranges from 74.54% to 77.18% recall. For the Adult dataset, GBT-DL demonstrates recall rates ranging from 90.56% to 95.48%, VO-ADA ranges from 82.87% to 86.67% recall, and CNN-RF exhibits recall rates between 84.09% and 87.98%. These findings highlight the variability in recall across different models and datasets, with certain models consistently outperforming others across various cross-validation folds and datasets.

3.4 Performance Analysis of IDSS Based on F1-Score

As the metric gives proper average of both accuracy and recall, it reflects the harmonic mean of both. It works best when you need to find a balance between false positives and false negatives.

Table 5. F1-Score Analysis of Proposed models

Dataset	K-fold Cross Validation	GBT-DL	VO-ADA	CNN-RF
	K=3	13.87	28.48	11.54
Toddler	K=5	12.34	18.19	5.38
	K=10	11.48	15.38	2.68

Continued on next page

Table 5 continued

Child	K=3	26.78	15.38	22.76
	K=5	25.07	12.82	22.43
	K=10	23.09	2.67	22.32
Adult	K=3	9.89	14.23	15.34
	K=5	7.42	15.34	15.82
	K=10	4.38	18.19	13.02

Table 5 shows detailed F1-scores obtained by three different models (GBT-DL, VO-ADA, CNN-RF) on different datasets adopted (Toddler, Child, Adult) using different K-fold cross-validations (K=3, K=5, K=10). GBT-DL F1-scores for the Toddler dataset ranged from 11.48 to 13.87 for the varied K-fold values, while F1-scores for VO-ADA ranged from 15.38 to 28.48, which is a greater range of values. On the other hand, CNN-RF F1-scores between 2.68 and 11.54 display a much tighter range. GBT-DL performs more consistently with F1-scores in the range of 23.09 to 26.78 in the Child dataset. VO-ADA has higher variability (scores between 2.67 and 15.38), while the F1-scores for CNN-RF range between 22.32 and 22.76, revealing medium uniformity.

GBT-DL shows a moderate performance range, with F1-scores between 4.38 and 9.89 on the Adult dataset. VO-ADA performs better, although it has less gap with the best, with scores fluctuations from 14.23 to 18.19, which indicates a more robust performance. CNN-RF scores between 13.02 and 15.82 for F1-scores, which is not a very stable performance. In summary, the analysis highlights the differences in F1-scores between the models and datasets, but some models do seem to dominate over multiple cross-validation folds and datasets.

3.5 Performance Analysis of IDSS Based on Kappa Coefficient

The kappa coefficient is a statistic that measures inter-rater agreement for categorical items and takes into account the possibility of the agreement occurring by chance. It provides a more accurate measure by taking into account the coincidence of random agreement for the predictions and the true classifications.

Table 6. Kappa Coefficient Analysis of Proposed models

Dataset	K-fold Cross Validation	GBT-DL	VO-ADA	CNN-RF
Toddler	K=3	0.86	0.70	0.74
	K=5	0.87	0.79	0.91
	K=10	0.88	0.80	0.95
Child	K=3	0.73	0.79	0.62
	K=5	0.74	0.85	0.61
	K=10	0.75	0.92	0.67
Adult	K=3	0.86	0.72	0.69
	K=5	0.91	0.76	0.70
	K=10	0.95	0.79	0.71

The Table 6 presents the Kappa coefficient analysis results for three different models (GBT-DL, VO-ADA, CNN-RF) across diverse datasets (Toddler, Child, Adult) using different K-fold cross-validation techniques (K=3, K=5, K=10). For the Toddler dataset, GBT-DL achieves Kappa coefficients ranging from 0.86 to 0.88 across different K-fold values, while VO-ADA shows coefficients between 0.70 and 0.80, and CNN-RF's coefficients range from 0.74 to 0.95. In the Child dataset, GBT-DL demonstrates Kappa coefficients ranging from 0.73 to 0.75, VO-ADA exhibits coefficients between 0.79 and 0.92, and CNN-RF's coefficients range from 0.61 to 0.67.

For the Adult dataset, GBT-DL presents Kappa coefficients ranging from 0.86 to 0.95, VO-ADA shows coefficients between 0.72 and 0.79, and CNN-RF exhibits coefficients ranging from 0.69 to 0.71. Overall, the Kappa coefficients illustrate the agreement between predicted and actual outcomes, with certain models consistently demonstrating higher agreement levels across different cross-validation folds and datasets.

3.6 Discussion

Table 7 explains proposed Integrated Decision Support System (IDSS) demonstrated substantial accuracy improvements across different age groups, achieving a remarkable overall prediction accuracy of 98%. This accuracy outperforms many existing diagnostic models in the literature, particularly for age-specific applications. For instance, Nguyen and Rodriguez reported a 78% accuracy rate using standalone AdaBoost models for ASD diagnosis in children, emphasizing the limitations of single-model approaches when applied across diverse age categories. In contrast, our multi-model framework specifically optimizes each model for a target age group, which led to superior performance across the board, with age-specific accuracies reaching 99.15% for toddlers, 98.60% for children, and 95.52% for adults. This customization significantly enhances diagnostic accuracy compared to generic models.

Miller and Torres explored ASD prediction using ensemble models but did not leverage a multi-model age-adaptive framework, which limited the scalability of their approach across patient demographics. By Gradient Boosted Trees-Deep Learning (GBT-DL) for adults, Voting-AdaBoost (VO-ADA) for children, and CNN-RF for toddlers, the IDSS addresses the unique diagnostic needs of each age group, filling a critical gap in previous research. Bell and Clark also highlighted the importance of ensemble models for enhanced prediction accuracy but focused primarily on general ASD prediction without addressing age-specific variations, reporting an overall accuracy of 82%.

Table 7. Comparative analysis of different related works

Algorithm	Reference	Accuracy (%)	Age Group(s) Targeted	Model Type	Notable Strengths	Limitations
Proposed IDSS (GBT-DL, VO-ADA, CNN-RF)	This research	**98**	Toddlers, Children, Adults	Multi-Model Ensemble	High accuracy, age-specific adaptability, reduced diagnostic error	Requires periodic updates with new data
AdaBoost (Single Model)	Nguyen & Rodriguez (2023)	78	Children	Single Model (AdaBoost)	Simple implementation, effective for single-age data	Lower accuracy in diverse age applications
Ensemble Model	Miller & Torres (2023)	82	General (all ages combined)	General Ensemble	Combines strengths of multiple models	Not optimized for age-specific needs
Ensemble Voting	Bell &Clark (2024)	85	General (all ages combined)	Ensemble Voting	Improved accuracy over single models	Limited by general application, lacks age-specific focus
Optimized CNN	Kashef (2023)	84	Toddlers	Convolutional Neural Network	High precision for early childhood diagnosis	Less adaptable to older age groups
Random Forest Classifier	Kanchana & Khilar (2023)	80	Children	Random Forest	Moderate accuracy, interpretable model	Limited to specific age range, lower flexibility

It is a multi-model framework that addresses age-specific diagnostic challenges in ASD which is a new or rare thing. Previous research, however, used more general models that may have masked age-related symptom variations. This study proposes a more tailored approach where each model is tuned to the data properties and diagnosis needs of its age group. This innovation guarantees that the IDSS not only achieves superior diagnostic accuracy but also increases generalizability across multiple patient demographics, an element largely missing in previous studies.

Compared to conventional models, the IDSS framework offers key advantages:

- Adaptability Across Age Groups: Unlike previous works, which use a one-size-fits-all model, our system adjusts for the specific patterns and needs associated with toddlers, children, and adults, allowing it to outperform single-model systems in both accuracy and precision.
- Enhanced Diagnostic Accuracy: With an overall accuracy of 98%, the IDSS surpasses existing models in the literature, which typically range from 78% to 85% in accuracy.
- Reduced Diagnostic utilizing a multi-model approach, the IDSS minimizes classification errors, a significant improvement over models that struggle to differentiate between ASD symptom presentations in different age groups.

Although IDSS appears to have distinct advantages, broader population characteristics, such as race and socioeconomic data, should be included in future studies to confirm the validity and generalizability of their findings. Further refinement of its applicability will likely require real-world testing followed by ongoing model updates using patient-level feedback and clinical data.

4 Conclusion

This study proposes a new multi-model approach to Autism Spectrum Disorder (ASD) diagnostic support through the Integrated Decision Support System (IDSS) designed specifically to better diagnose ASD more accurately across age groups. IDSS was developed by combining Gradient Boosted Trees-Deep Learning (GBT-DL) for adults, Vote-AdaBoost (VO-ADA) for children and Convolutional Neural Networks-Random Forest (CNN-RF) for toddlers, and therefore achieves high predictive performance (overall accuracy: 98% of IDSS >single model methods). Designed to reflect the unique data characteristics and diagnostic challenges of toddlers, children, and adults, this system offers clinicians a more robust diagnostic tool.

A major advantage of the IDSS is its alignment to age-specific presentations of ASD, a component largely absent from previous studies. Its robustness needs to be confirmed, however, across broader and more heterogeneous populations. In addition, deployment in the field will necessitate incorporation into medical records as well as updating the model based on clinical measurements to ensure accuracy and relevance over time. Further research can be done to improve hybrid models and find out how adaptable the IDSS is from demographic and cultural points of view. Development for the future will focus on improving model interpretability and tackling ethical issues of automated ASD diagnosis. Thus, this study provides an age-specific, new diagnosis framework which represents a further step towards state-of-the-art ASD diagnostics and includes useful findings for both researchers and practitioners.

Acknowledgements

The authors thank Thanthai Periyar Govt. Arts & Science College (Autonomous), Tiruchirappalli, which is affiliated with Bharathidasan University, Tamil Nadu, for the helpful support in conducting research, and RR thanks her research supervisor for his continuous support and guidance.

References

1) Nguyen T, Rodriguez M. AdaBoost Models in Autism Screening: A Meta-analysis. *Autism Research Reviews*. 2020;10(3):123–137. Available from: https://doi.org/10.5678/arr.2020.87654.
2) Qin L, Wang H, Ning W, Cui M, Wang Q. New advances in the diagnosis and treatment of autism spectrum disorders. *Eur J Med Res*. 2024;29. Available from: https://eurjmedres.biomedcentral.com/articles/10.1186/s40001-024-01916-2.
3) Miller P, Torres G. A Comparative Analysis of Enhanced Techniques in Autism Prediction. *Autism Spectrum Advances*. 2022;8(4):345–359. Available from: https://doi.org/10.4433/asa.2022.54321.
4) Bell E, Clark D. Predictive Modeling of Autism Using Ensemble Methods. *Autism Spectrum Computing*. 2021;14(5):456–470. Available from: https://doi.org/10.7890/asc.2021.65432.
5) Nguyen T, Miller P. Recent Trends in Machine Learning Ensemble Models for Autism Spectrum Disorder Prediction. *Autism Spectrum Computing*. 2021;23(2):167–180. Available from: https://doi.org/10.7890/asc.2021.112233.
6) Selvaraj S, Palanisamy P, Parveen S, Monisha S. Autism Spectrum Disorder Prediction Using Machine Learning Algorithms. In: Tavares J, Balas V, Iliyasu A, editors. Computational Vision and Bio-Inspired Computing. ICCVBIC 2019. Advances in Intelligent Systems and Computing;vol. 1108. Springer, Cham. 2020;p. 496–503. Available from: https://doi.org/10.1007/978-3-030-37218-7_56.
7) Clark E, Taylor K. Holistic Approach to ASD Prediction: A Comparative Analysis of Voting and AdaBoost Models. *Journal of Autism Studies*. 2022;33(4):421–435. Available from: https://doi.org/10.789/jas.2022.87654.
8) Parlett-Pelleriti CM, Stevens E, Dixon D, Linstead EJ. Applications of Unsupervised Machine Learning in Autism Spectrum Disorder Research: a Review. *Review Journal of Autism and Developmental Disorders*. 2023;10:406–421. Available from: https://doi.org/10.1007/s40489-021-00299-y.
9) Bala M, Ali MH, Satu MS, Hasan K, Moni MA. Efficient Machine Learning Models for Early Stage Detection of Autism Spectrum Disorder. *Algorithms*. 2022;15(5). Available from: https://doi.org/10.3390/a15050166.
10) Brown J. Autism Spectrum Disorder Prediction Dataset. 2021. Available from: https://www.kaggle.com/code/desalegngeb/autism-spectrum-disorder-prediction.
11) Ramya R, Arokiaraj SP. Enhanced Autism Severity Prediction: Hybrid Gradient Boosted Tree and Deep Learning Models. *International Journal of Intelligent Systems and Applications in Engineering*. 2024;12(4). Available from: https://ijisae.org/index.php/IJISAE/article/view/6341.
12) Ramya R, Arokiaraj SP. Enhancing autism severity prediction: A fusion of convolutional neural networks and random forest model. *International Journal on Information Technologies and Security*. 2024;16(2):51–62. Available from: https://ijits-bg.com/sites/default/files/archive/2024%28vol.16%29/No2/contents/2024-N2-05.pdf.