ENHANCED AUTISM SEVERITY PREDICTION: A FUSION OF CONVOLUTIONAL NEURAL NETWORKS AND RANDOM FOREST MODEL

R. Ramya, S. Panneer Arokiaraj*

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

* Corresponding Author, e-mail: *rgramya94@gmail.com*

Abstract: A neurological condition affecting both the brain and behavior is identified as autism spectrum disorder (ASD). Due to the absence of a reliable medical test for detecting autism, diagnoses rely on historical evidence. Essential in assessing the degree of autism are models like Convolutional Neural Networks (CNNs) and Random Forest (RF). In order to reduce the number of diagnostic tests required for autism diagnosis, this research work presents a new hybrid model that combines the strengths of RF and CNNs, providing healthcare solutions. It is noteworthy that this model properly predicted the severity of autism with an astounding accuracy rate of 99.15% when applied to historical data gathered from the Kaggle Repository.

Key words: autism spectrum disorder (ASD), data Science models, random forest and convolutional neural networks (CNN-RF) model, accuracy, precision, recall, F1-score and Kappa statistic.

1. INTRODUCTION

Autism Spectrum Disorder (ASD), commonly referred to as autism, is a multifaceted neurodevelopmental condition influencing communication, social engagement, behavior, and sensory processing. Termed a "spectrum" disorder, it encompasses a broad range of symptoms and levels of impairment among individuals with autism. Individuals with autism spectrum disorder may experience varying degrees of challenges, with some facing mild obstacles and others encountering more substantial difficulties that necessitate ongoing support and care throughout their lives. The precise origins of autism remain not entirely clear, yet it is thought to stem from a blend of genetic and environmental influences. Timely identification and intervention play a pivotal role in enabling those with autism to lead fulfilling lives and achieve their maximum capabilities.

This research work aims to predicting autism in a historical dataset involves using hybrid data science techniques to identify patterns and relationships that may indicate the likelihood of a person autism level. In essence, the proposed hybrid models, which integrate Convolutional Neural Network and Random Forest models, have the potential to propel the field of autism diagnosis and assessment forward. By harnessing the strengths of both models, these hybrid approaches enhance accuracy, interpretability, and feature selection capabilities, ultimately contributing to more effective decision-making and personalized care for individuals with autism.

The article is organized in a manner that first delves into a comprehensive literature review, covering a broad spectrum of current methods and data science models related to autism detection. The workflow of the suggested work is outlined in the section detailing the working model. The section on the suggested system provides an overview of autism, outlines the dataset characteristics, explores CNN-RF, and elucidates on its hybridization. The results and discussion section offers insights into the effectiveness and error rate of the proposed model in predicting autism. The article concludes by discussing the significance of this research project and outlining potential avenues for its future expansion.

2. LITERATURE REVIEW

The diagnosis and understanding of ASD have witnessed significant advancements through the integration of Data Science Models. This literature review provides a comprehensive overview of key studies that contribute to optimizing ASD diagnosis, particularly focusing on the development and application of a hybrid RF & CNN model for accurate severity prediction.

Kaushik Vakadkar et al. explored ASD detection in children through machine learning techniques, emphasizing the role of ML in early diagnosis [1]. Rabbi et al. proposed a CNN model for early-stage ASD detection, demonstrating the potential of deep learning in improving early diagnostic capabilities [2]. Shorten et al. delved into deep learning applications across various domains, providing insights into potential applications of deep learning techniques, including those related to ASD [3].

Ahmed et al. utilized eye tracking-based techniques in ASD diagnosis, showcasing the integration of novel technologies with ML and deep learning [4]. Mujeeb Rahman KK and Subashini MM employed static facial features and deep neural networks for identifying ASD in children, emphasizing the significance of facial features in the diagnostic process [5].

Minissi et al. conducted a systematic review assessing ASD based on ML and social visual attention, contributing to a better understanding of spectrum-level deficits [6]. Parlett-Pelleriti et al. explored unsupervised ML applications in ASD research, providing valuable insights into the potential of unsupervised techniques [7].

Bala et al. proposed efficient ML models for early-stage ASD detection, contributing to the development of effective diagnostic tools [8]. Jacob et al. discussed algorithmic approaches for classifying ASD, providing a research perspective on ML applications in ASD diagnosis [9]. Kashef introduced an enhanced convolutional neural network for efficient ASD diagnosis, highlighting the continual improvement of deep learning models [10].

Kanchana and Khilar predicted ASD in adults using a random forest classifier, showcasing the versatility of ensemble techniques in ASD diagnosis [11]. Savyanand et al. presented the findings regarding the efficacy of different machine learning and deep learning approaches in classifying stars and galaxies. This could include insights into the accuracy, precision, recall, and other performance metrics achieved by the proposed methods[12].

This literature review highlights significant progress in diagnosing ASD through Data Science Models. These studies emphasize the diverse applications of machine learning and deep learning in identifying ASD patterns, predicting severity, and enhancing our understanding of spectrum-level deficits. Now, focusing on the proposed CNN RF model workflow, our approach builds on literature insights. The hybrid model merges Convolutional Neural Network (CNN) and Random Forest (RF) methods, harnessing their respective strengths. The structured workflow includes a literature review, data collection, and preprocessing for relevant ASD features.

3. WORKING MODEL

Figure 1 illustrates the overall workflow of the proposed CNN-RF Model for predicting autism using data science techniques. The workflow begins by data preprocessing the dataset such as to handle missing values, outliers, and categorical attributes. Unwanted disturbances are eliminated, and methods for extracting features are employed to choose the most pertinent ones, diminishing data dimensionality for improved training efficiency. Following the preprocessing of the dataset, hybrid CNN-RF data science models are then utilized to forecast the output label. The accuracy of hybrid data science techniques is observed and compared. Accuracy alone may not be sufficient, so additional metrics like accuracy, precision, recall, F1-Score and kappa statistic values are computed to provide a more comprehensive evaluation of hybrid technique performance.

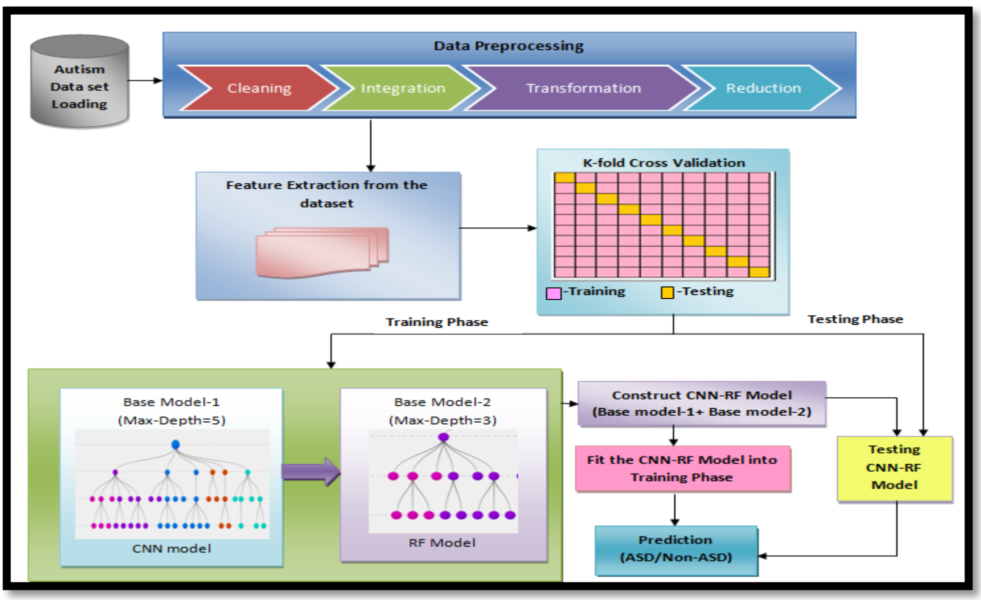


Figure 1. Workflow of the Proposed CNN-RF Model

Creating a new CNN-RF model for enhancing precision in autism prediction involves building a hybrid model that leverages the strengths of both approaches. The proposed CNN-RF model is an implemented using Python 3 version, with historical dataset and relevant features for autism prediction.

The description of this approach is presented in the "Methodology" section, providing a concise overview of the steps taken to develop and evaluate the predictive model for autism detection using historical data. It is important to note that the methodology should be thoroughly explained, specific implementation details, hyper parameter tuning, and cross-validation techniques should be included to ensure the study's reliability and efficiency.

4. METHODOLOGY

4.1. Dataset Analysis

Table 1, presented below, delineates details regarding the autism adult dataset accessible in the Kaggle repository [13]. This dataset is employed to assess the efficacy of the chosen hybrid data science model within a distinct domain, specifically concerning autism prediction.

Table 1. Description of the Autism Dataset

S. No.	*Attributes*	*Description*
1	*Case_No*	*Case number for each entry.*
2	*A1_Score to A10_Score*	*Binary features indicating the presence (1) or absence (0) of certain characteristics.*
3	*Age*	*Age in months*
4	*Qchat-10-Score*	*Qchat-10 questionnaire score*
5	*Sex*	*Gender of the individual (m: male, f: female).*
6	*Ethnicity*	*Ethnicity of the individual.*
7	*Jaundice*	*Whether the individual had jaundice (yes/no).*
8	*Family_mem_with_ASD*	*Whether there is a family member with ASD (yes/no).*
9	*Who completed the test*	*Person completing the test (family member, Health Care Professional, etc.).*
10	*Labels:Class/ASD Traits*	*Indicates the presence (Yes) or absence (No) of ASD traits.*

The Figure 2 represents graphical analysis of autism dataset includes both categorical (Sex, Ethnicity, Jaundice, Family_mem_with_ASD, Who completed the test) and numerical features (Age_Mons, Qchat-10-Score). The binary features A1-A10 seem to represent certain characteristics or behaviors. The dataset contains a mix of demographic information (age, gender, ethnicity) and potential risk factors for ASD (family history, jaundice, Qchat-10 score).

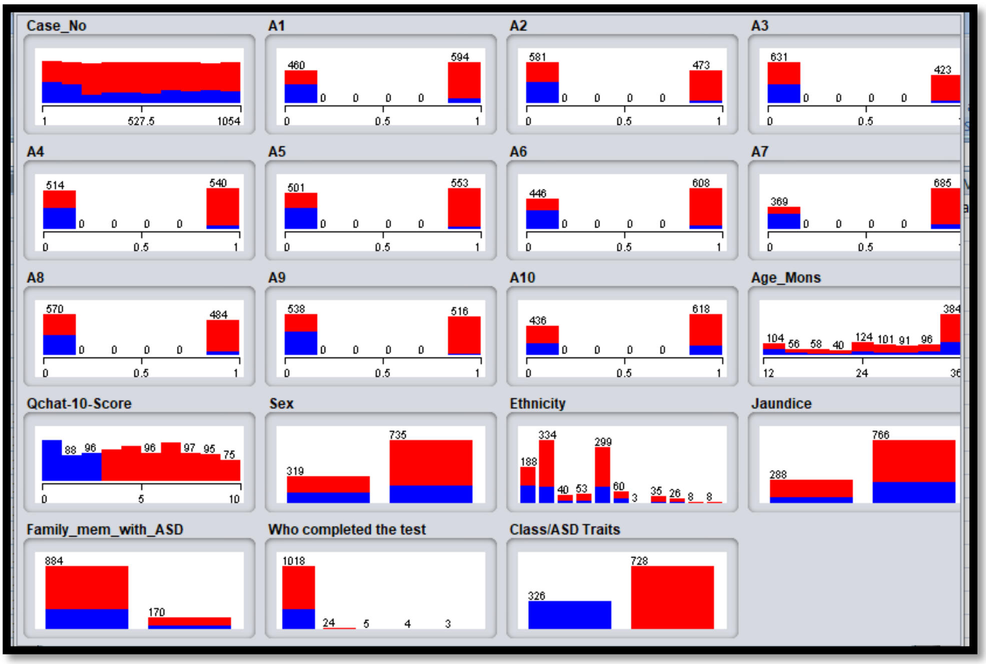


Figure 2. Autism Dataset Analysis

4.2. Preprocessing

The dataset comprises 1054 instances with 18 attributes, inclusive of a class variable. However, before training the model and conducting analysis, the data underwent preprocessing due to the presence of non-contributing and categorical attributes. The following steps were executed during preprocessing:

- **Removal of Non-Contributing Attributes:** Three attributes, specifically 'Case_No,' 'Who completed the test,' and 'Qchat-10-Score,' were identified as non-contributing and subsequently eliminated from the dataset.
- **Label Encoding for Binary Features:** Four features, each with two classes – 'Sex,' 'Jaundice,' 'Family_mem_with_ASD,' and 'Class/ASD_Traits' – underwent label encoding. Label encoding involves converting categorical labels into a numeric format for machine readability. In this process, labels were transformed into numeric values, with repeated labels assigned the same numeric value.
- **One-Hot Encoding for Multiclass Features:** The 'Ethnicity' feature, presenting 11 classes, underwent one-hot encoding. This technique is employed for categorical features with more than two classes to prevent introducing hierarchical ordering by the model. It represents each class as a binary vector, where each class corresponds to a unique binary value (1), while other classes are denoted by 0s.

Through these preprocessing steps, the dataset was refined for model training, eliminating irrelevant attributes, and converting categorical attributes into a format conducive to analysis and modeling.

4.3. Proposed Model

The proposed system, the CNN-RF Algorithm, offers a novel hybrid approach by integrating Convolutional Neural Networks (CNNs) and Random Forest (RF) classifiers. This innovative model combines the strengths of CNNs in feature extraction with the robustness of RF in classification tasks. By leveraging the complementary capabilities of both algorithms, the CNN-RF Algorithm aims to enhance predictive accuracy and generalization performance across diverse datasets. This hybrid system presents a promising avenue for addressing complex classification challenges effectively and efficiently.

Utilizing Convolutional Neural Networks (CNNs) for autism prediction involves processing input data, such as behavioral data, and subsequently making predictions based on patterns learned from the data. For an input of size W x W x D, with Dout number of kernels having a spatial size of F, a stride of S, and padding of P, the formula to compute the size of the output volume is expressed in Equation (1) as follows:

$$\text{Output size (Hout)} = \frac{(\mathrm{W} - \mathrm{F} + 2 * \mathrm{P})}{\mathrm{S} + 1} \tag{1}$$

The formula is applied to autism historical dataset. If the input volume is square (W x W), then the output volume will also be square (Hout x Hout). The depth (number of channels) of the output volume will be equal to Dout, as each kernel produces one output channel.

In the Random Forest framework, the feature importance for a specific feature, denoted as RFfi sub(i), is determined by averaging its importance across all trees in the forest. The normalized feature importance for feature i in a particular tree j, indicated as normfi sub(ij), is calculated by dividing the feature's importance value in that tree by the sum of the importance values of all features in that tree. The formulas are articulated as follows:

Feature Importance in Random Forest (RFfi sub(i)):

$$\text{RFfi sub(i)} = \frac{1}{\mathrm{T}} * \text{sum}\big(\text{normfi sub(ij)}\big) \text{for j} = \text{1 to T} \tag{2}$$

Here, T represents the total number of trees in the Random Forest model.

Normalized Feature Importance in Tree (normfi sub(ij)):

$$\text{normfi sub(ij)} = \frac{\text{Importance of feature i in tree j}}{\text{sum(Importance of all features in tree j)}} \tag{3}$$

In numerous Random Forest implementations, the determination of feature importance values relies on how a specific feature in the historical autism dataset contributes to diminishing impurity within the tree nodes. The importance values associated with a feature in the historical autism dataset are derived from the node splits involving that particular feature across all trees in the forest. Features that recurrently participate in node splits and result in substantial reductions in impurity generally exhibit higher importance values.

The hybrid CNN-RF Model, which merges the capabilities of Convolutional Neural Network (CNN) and Random Forest, stands as a robust data science approach for forecasting autism levels. The accompanying Figure 3 illustrates the pseudo code for this innovative methodology.

Algorithm: CNN-RF Algorithm (Hybrid model for Convolutional Neural Network and Random Forest)

Input: data: Autism Spectrum Disorder Dataset

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:* $F_{CNN} = \{f1, f2, \dots, fm\}$ *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

Figure 3. Pseudo code of the proposed CNN-RF Model

In Figure 3, the provided pseudo code outlines a systematic process for handling a dataset, involving its division into training and testing sets. Initially, a conventional data science model is trained on the designated training data. Subsequently, features are extracted from the training data using the trained traditional model. These extracted features serve as the basis for training a deep learning model. The same sequential process is applied to the testing data, where features are extracted utilizing the trained traditional model. The hybrid Convolutional Neural Network (CNN) and Random Forest model then make predictions based on these features. The final step involves evaluating the performance of the hybrid model by assessing the accuracy of its predictions in comparison to the actual labels.

By harnessing the strengths of both CNN and Random Forest models, the CNN-RF Model aspires to elevate prediction accuracy and interpretability in determining autism levels. Beyond accurate predictions, the hybrid models offer valuable insights into the factors influencing autism diagnosis. The interpretability inherent in the Random Forest component enables the identification of the most crucial visual and non-visual features contributing to the prediction. Such insights contribute to an enhanced understanding of autism spectrum disorders, paving the way for more targeted interventions and treatments.

5. ANALYSIS AND RESULTS

5.1. Evaluation Matrices

This section describes an analysis conducted on experimental results using hybrid data science model on the Autism Dataset. The purpose of this analysis is to assess the performance of hybrid model based on statistical matrices, such as accuracy, precision, recall, F1-Score and kappa statistic were determined through the application of the following mathematical expressions:

$$\text{Accuracy} = \frac{(\text{Number of Correct Predictions})}{(\text{Total Number of Predictions})} \tag{4}$$

$$\text{Recall} = \frac{(\text{True Positives})}{(\text{True Positives} + \text{False Negatives})} \tag{5}$$

$$\text{Precision} = \frac{(\text{True Positives})}{(\text{True Positives} + \text{False Positives})} \tag{6}$$

$$\text{F1} - \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

$$\text{Kappa} = \frac{(\text{Po} - \text{Pe})}{(1 - \text{Pe})} \tag{8}$$

Where Po (Observed Agreement) is the proportion of agreement between raters or the model's observed accuracy and Pse (Expected Agreement) is the proportion of agreement expected by chance.

Table 2. Result Analysis of the proposed CNN-RF Model

K-fold Cross Validation	*CNN-RF Model - Performance metrics*							
	Training Phase (80%)				***Testing Phase (20%)***			
	K=3	***K=5***	***K=10***	***K=15***	***K=3***	***K=5***	***K=10***	***K=15***
Accuracy	*89.45*	*95.56*	***99.15***	*97.52*	*89.25*	*95.36*	***99.4***	*97.98*
Precision	*88.00*	*94.26*	***98.89***	*96.32*	*87.87*	*94.04*	***98.73***	*96.89*
Recall	*88.09*	*94.48*	***97.18***	*96.78*	*87.85*	*94.42*	***97.11***	*96.69*
F1-Score	*11.54*	*5.38*	***1.98***	*3.2*	*12.54*	*5.54*	***2.04***	*2.98*
Kappa statistic	*0.74*	*0.91*	***0.96***	*0.94*	*0.74*	*0.91*	***0.96***	*0.95*

Table 2, presented below, furnishes performance metrics for the hybrid model under various K-fold cross-validation values (where K denotes the number of folds). K-fold cross-validation is a methodology employed to evaluate the effectiveness of a hybrid data science model. This approach involves partitioning the dataset into K subsets, or folds. The model is then trained on K-1 folds and tested on the remaining fold, repeating this process K times to ensure each fold serves as a test set precisely once.

The values in the Table 2 represent the performance metrics for each combination of K and the training and testing phases. At K=10, during both the training and testing phases, the hybrid model achieved an accuracy of 99.15%, precision of 98.89%, recall of 97.18%, F1-Score of 5.68%, and a Kappa statistic of 0.96. In figure 4 also explains graphical representation of performance analysis in CNN-RF model.

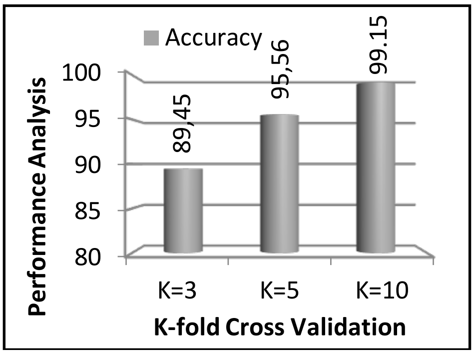


Figure 4(i) Accuracy

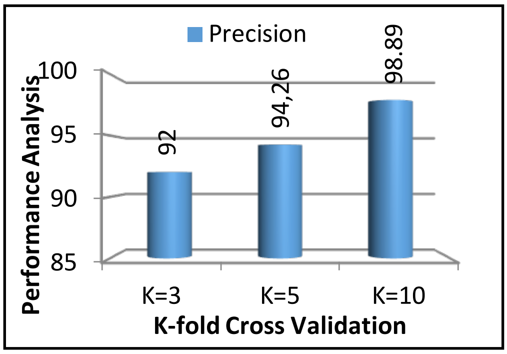


Figure 4(ii) Precision

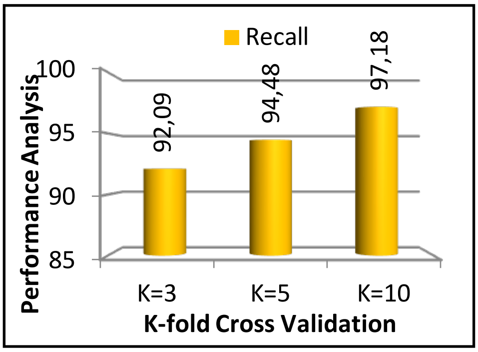


Figure 4(iii) Recall

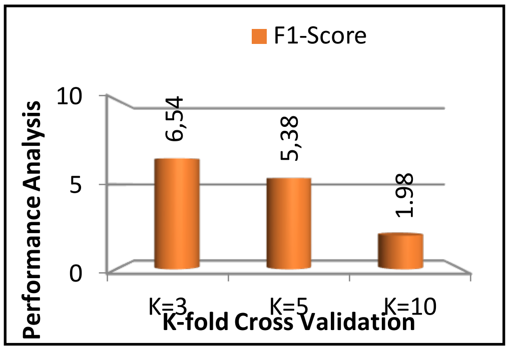
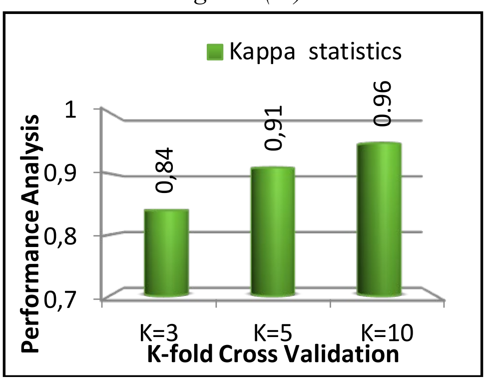


Figure 4(iv) F1-Score



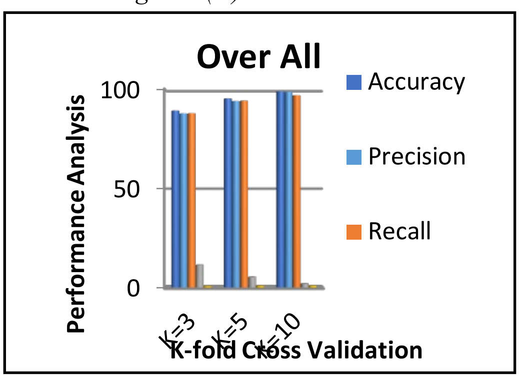


Figure 4. Performance Analysis of Proposed CNN-RF Model

5.2. Comparison of Data Science Models

The combination of Convolutional Neural Network and Random Forest addresses limitations observed in standalone models, presenting a more comprehensive and accurate prediction. This finding suggests that the combination of the Random Forest and CNN models is leading to improved results in terms of autism prediction precision compared to when each model is used individually. It's common in data science to observe that combining different models can lead to better performance than using them in isolation.

Table 3 Shows CNN-RF model generally outperforms both CNN and RF individually across all values of K-fold Cross Validation, with higher accuracy, precision, recall, and Kappa statistic. The CNN model has a slight decrease in performance as K-fold Cross Validation increases, especially in terms of accuracy and Kappa statistic. The RF model shows relatively stable performance across different values of K-fold Cross Validation. In figure 5 also explains graphical representation of comparative analysis in various models.

Table 3: Comparison of different Data Science Models

Performance analysis	*CNN*			*RF*			*CNN-RF*		
	K=3	*K=5*	*K=10*	*K=3*	*K=5*	*K=10*	***K=3***	***K=5***	***K=10***
Accuracy	*79.54*	*75.32*	*78.75*	*84.62*	*85.64*	*88.71*	***89.45***	***95.56***	***98.75***
Precision	*78.08*	*74.85*	*77.89*	*84.00*	*84.26*	*87.90*	***88.00***	***94.26***	***97.89***
Recall	*78.15*	*74.54*	*77.18*	*84.09*	*84.48*	*87.98*	***88.09***	***94.48***	***97.18***
F1-Score	*22.76*	*22.43*	*22.32*	*15.34*	*15.82*	*13.02*	***11.54***	***5.38***	***2.68***
Kappa statistic	*0.62*	*0.61*	*0.67*	*0.69*	*0.70*	*0.71*	***0.74***	***0.91***	***0.95***

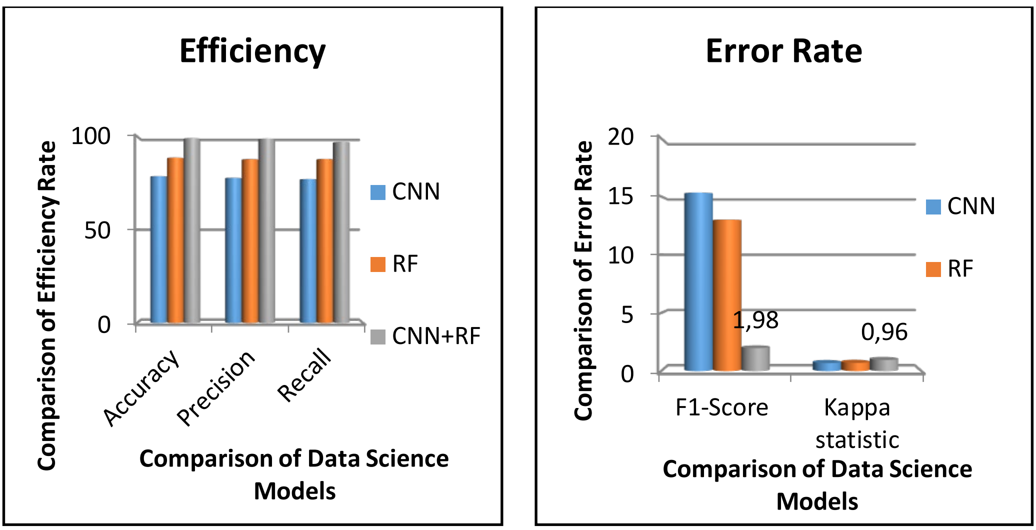


Figure 5.Comparison of Data Science Models with Efficiency and Error Rate

5. CONCLUSION

This research emphasizes the pivotal role of advanced data science models in addressing challenges in Autism Spectrum Disorder (ASD) diagnosis. With the absence of a reliable medical test for autism detection, reliance on historical data is essential. This research introduces a Hybrid Model that merges Random Forest (RF) and Convolutional Neural Networks (CNNs) to improve the precision of autism severity prediction. The integration of CNN and RF not only tackles limitations identified in standalone models but also showcases a more holistic and accurate prediction. The consistent superiority of the CNN-RF model over individual CNN and RF models in diverse K-fold Cross Validation scenarios, demonstrating elevated accuracy, precision, recall, and Kappa statistic, aligns with the common observation in data science that amalgamating different models often yields superior performance compared to using them independently. The results of the study make a substantial contribution to the progress of autism diagnosis, emphasizing the capacity of hybrid data science models to improve predictive accuracy and overall healthcare solutions.

REFERENCES

[1] KaushikVakadkar, DiyaPurkayastha, Deepa Krishnan. Detection of Autism Spectrum Disorder in Children Using Machine Learning Techniques. *SN Computer Science*, vol.2, no.2&3, 2021, p.386. https://doi.org/10.1007/s42979-021-00776-5.

[2] M. F. Rabbi, S. M. M. Hasan, A. I. Champa, M. A. Zaman. A Convolutional Neural Network Model for Early-Stage Detection of Autism Spectrum Disorder. *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, Dhaka, Bangladesh, 2021, pp. 110-114, doi: 10.1109/ICICT4SD50815.2021.9397020.

[3] Shorten, C., Khoshgoftaar, T.M., Furht, B. Deep Learning applications for COVID-19. Journal Big Data, vol. 8, no.18, 2021. https://doi.org/10.1186/s40537-020-00392-9.

[4] Ahmed, I.A., Senan, E.M., Rassem, T.H., Ali, M.A.H., Shatnawi, H.S.A., Alwazer, S.M., Alshahrani, M. Eye Tracking-Based Diagnosis and Early Detection of Autism Spectrum Disorder Using Machine Learning and Deep Learning Techniques. *Electronics* vol.11, 2022, 530. https://doi.org/10.3390/electronics11040530.

[5] Mujeeb Rahman KK, Subashini MM. Identification of Autism in Children Using Static Facial Features and Deep Neural Networks. *Brain Sciences,* vol.12, no.1, 2022, p.94. https://doi.org/10.3390/brainsci12010094.

[6] Minissi, M.E., Chicchi Giglioli, I.A., Mantovani, F. et al. Assessment of the Autism Spectrum Disorder Based on Machine Learning and Social Visual Attention: A Systematic Review. J Autism Dev Disord 52, 2187–2202 (2022). https://doi.org/10.1007/s10803-021-05106-5

[7] Parlett-Pelleriti, C.M., Stevens, E., Dixon, D. et al. Applications of Unsupervised Machine Learning in Autism Spectrum Disorder Research: a Review. *Review Journal Autism Dev Disord,* 2022. https://doi.org/10.1007/s40489-021-00299-y.

[8] Bala, Mousumi, Mohammad Hanif Ali, Md. Shahriare Satu, Khondokar Fida Hasan, Mohammad Ali Moni. Efficient Machine Learning Models for Early Stage Detection of

Autism Spectrum Disorder. Algorithms, vol.15, no.5, 2022, p.166. https://doi.org/10.3390/a15050166.

[9] Shomona Gracia Jacob, Majdi Mohammed Bait Ali Sulaiman, Bensujin Bennet, Algorithmic Approaches to Classify Autism Spectrum Disorders: A Research Perspective. *Procedia Computer Science*, vol.201, 2022, pp. 470-477, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2022.03.061.

[10] Rasha Kashef. ECNN: Enhanced convolutional neural network for efficient diagnosis of autism spectrum disorder. *Cognitive Systems Research*, vol.71, 2022, pp.41-49, ISSN 1389-0417, https://doi.org/10.1016/j.cogsys.2021.10.002.

[11] Kanchana, R. Khilar. Prediction of Autism Spectrum Disorder using Random Forest Classifier in Adults. *2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA)*, Goa, India, 2022, pp. 242-249, doi: 10.1109/ICCCMLA56841.2022.9989304.

[12] Amit Sadanand Savyanavar, Nikhil Mhala, Shiv H. Sutar. Star-Galaxy classification using machine learning algorithms and deep learning. *International Journal on Information Technologies and Security*, vol.15, no.2, 2023, pp. 87-96. https://doi.org/10.59035/VVLR5284.

[13] Brown, J. Autism Spectrum Disorder Prediction Dataset. *Kaggle*, 2021. https://www.kaggle.com/brownj/asd-prediction-dataset

Information about the authors:

R. Ramya–Research Scholar, Department of Computer Science, ThanthaiPeriyar Govt. Arts & Science College(Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.

Dr. S. Panneer Arokiaraj – Associate Professor, Department of Computer Science, ThanthaiPeriyar Govt. Arts & Science College(Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India.

Manuscript received on 14 March 2024