Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 6 (2023)

Design of Diagnostic Framework for Detecting Autism Spectrum Disorder using Conditional Mutual Information Maximization-Random Forest (CMIM-RF) Approach

[1] Keerthi Guttikonda, [2] Dr. G. Ramachandran, [3] Dr. G. V. S. N. R. V. Prasad

[1] Research Scholar , Dept. of Computer Science and Engineering
Annamalai University
Annamalainagar – 608002
Tamilnadu, India
[2] Associate Professor, Dept. of Computer Science and Engineering
Annamalai University
Annamalainagar – 608002
Tamilnadu, India
[3] Professor, Dept. of Computer Science and Engineering
Gudlavalleru Engineering College
Gudlavalleru - 521356
Andhra Pradesh, India

Email: [1] keerthi.guttikonda@gmail.com, [2] gmrama1975@gmail.com,
[3] gutta.prasad1@gmail.com

Abstract: Autism Spectrum Disorder (ASD) has detrimental effects on social interaction, communication, and repetitive behaviors. Standardized diagnosis approaches for ASD often include extensive, subjective clinical observations and behaviour assessments. Due to recent advances in Machine Learning (ML), more precise and time-efficient ASD diagnosis may soon be achievable. Using machine learning approaches such as data pre-processing, feature extraction, and classification, this study proposes a diagnostic paradigm for ASD. In this work, it was analysed that the most advanced machine learning algorithms used in each component and compared their performance on a benchmark dataset. In terms of accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve, proposed Conditional Mutual Information Maximization-Random Forest (CMIM—RF) method was found to be superior to conventional diagnostic methods. According to the findings of this study, ML-based diagnostic frameworks may one day become an indispensable resource for clinicians and researchers in the area of ASD. However, more study is required to validate these findings in larger samples and evaluate the potential for bias. Physiological data, genetic information, and brain imaging are examples of additional aspects that might be used in an attempt to improve diagnosis accuracy.
Keywords: Autism Spectrum Disorder, Machine Learning, Random Forest, Diagnostic Framework

1. INTRODUCTION

Autism Spectrum Disorder (ASD) is an ailment that changes the way a person behaves, talks, and interacts with others. The diagnosis of autism can be complex because symptoms can differ greatly among individuals and there is no singular test to diagnose it [1]. Instead, the diagnosis is usually determined through a combination of observations of the individual's behaviour and development, as well as input from parents, caregivers, and other professionals. Some common challenges in diagnosing autism include difficulty identifying and interpreting social cues, difficulty with communication, and repetitive behaviours that may not be apparent to others. Additionally, there is still a lack of understanding and awareness of autism in some communities, which can make it difficult for individuals with autism to receive a proper diagnosis. Early detection and intervention for autism can improve outcomes for individuals with the disorder [14]. Some ways to detect autism in early include the following:

1. **Watch for developmental milestones**: Parents and other caretakers should know the typical developmental milestones for children and watch for any delays or changes in those milestones, such as a child who doesn't make eye contact, doesn't respond to his or her name, or doesn't point to things.
2. **Look for behavioral signs**: Some common signs of autism in terms of behaviour are trouble making friends, having trouble communicating, and doing the same thing over and over, like rocking or flapping.
3. **Consult with a pediatrician:** Parents and caregivers should consult with a pediatrician if they have any concerns about their child's development. Pediatricians can perform developmental screenings and refer families to specialists for further evaluation if necessary.
4. **Seek an evaluation from a specialist:** An evaluation from a specialist such as a developmental pediatrician, psychologist, or neuropsychologist with experience in autism can provide a more detailed and accurate assessment.

It is important to note that early intervention is key, and the earlier the diagnosis, the better. Also, it's important to note that autism can't be diagnosed only through a blood test or a scan, it requires a multi-disciplinary evaluation. Recently, Machine Learning (ML) techniques have been proposed as a potential solution to improve the diagnostic accuracy and efficiency of ASD. ML algorithms can analyze large amounts of data, such as behavioral and physiological measures, and identify patterns that are not easily discernible by human experts.

This study presents ML based diagnostic framework for identifying individuals with ASD. This proposed framework combines behavioral, physiological and demographic data to make predictions about the presence of ASD [10]. The effectiveness of the approach is tested using a dataset of children that includes both individuals with ASD and controls with generally developing cognitive abilities. Experimental results demonstrate the potential of framework to improve the diagnostic accuracy and efficiency of ASD. This research aligns with previous studies that have applied ML techniques to the diagnosis of ASD [9] [19]. This proposed framework adds to the existing literature by incorporating a wider range of data sources and evaluating the performance on a larger dataset. Overall, this study highlights the potential of ML-based diagnostic frameworks to improve the diagnosis of ASD and suggests the need for further research in this area. The major contributions of this research work are:

1. This study proposes a ML based diagnostic framework for ASD which comprises of three key components: data pre-processing, feature extraction, and classification.
2. The state-of-the-art machine learning techniques are reviewed and employed in each component of the proposed framework.
3. The effectiveness of the concept is exhaustively assessed and compared to conventional diagnostic techniques.
4. Results demonstrate that the proposed Conditional Mutual Information Maximization-Random Forest (CMIM-RF) method outperforms traditional diagnostic methods in terms of accuracy, sensitivity and specificity.
5. The research suggests that ML-based diagnostic frameworks have the potential to enhance the diagnostic process for ASD and serve as a valuable tool for clinicians and researchers.

The rest of this paper's sections are set up in the following way: In Section 2 study of the literature is given. Section 3 includes the methods used. The results and discussion are presented in Section 4. Section 5 outlines the conclusion and suggestions for future enhancements.

2. LITERATURE SURVEY

This section talks about the different state-of-the-art ML approaches that have been published for detecting ASD.

The work [3] discussed the use of ML techniques for diagnosing ASD using Magnetic Resonance Imaging (MRI) data. Combining Recursive Feature Elimination (RFE) and a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel, the authors provide a unique method for the diagnosis of ASD. The study utilised a dataset of functional and structural MRI data from individuals both with and without ASD to train and evaluate the method. SVM with an RBF kernel was used to distinguish between those with and

without ASD after RFE was performed to choose the most relevant characteristics from the MRI data. The study's results showed that the method could correctly divide people into those with and those without ASD, with an overall accuracy of about 86%. The performance of the method was also compared to other machine learning methods like Random Forest (RF) and k-Nearest Neighbours (k-NN) and found to be better.

The research work in [7] explores the application of machine learning techniques in diagnosing ASD and Attention - Deficit/Hyperactivity Disorder (ADHD) using functional and structural MRI data. The authors review a range of machine learning methods that have been utilised in the past for this purpose. They highlight that the combination of functional and structural MRI data with machine learning methods has demonstrated the potential to accurately identify individuals with ASD and ADHD. The paper explains that functional MRI (fMRI) can be used to pinpoint brain regions that are activated during different cognitive tasks, while structural MRI (sMRI) can be used to identify brain regions that are abnormal in size or shape. The authors mention that ML techniques such as SVM, RF, and Convolutional Neural Network (CNN) have been employed to classify individuals with ASD and ADHD based on their functional and structural MRI data. They also mention the challenges and limitations of utilising machine learning for diagnosing ASD and ADHD, which include the lack of large and diverse datasets, the variability of the functional and structural MRI data, and the complexity of the disorders.

In this work [16] proposed an algorithm that is based on feature correlation and ranking, and is applied to a clinical dataset of ASD patients. The method is used to identify the most relevant characteristics from the dataset, and then different ML algorithms such as Decision Tree (DT), Gradient Boosting Classifier (GBC), AdaBoost, and Logistic Regression (LR) are applied to the selected features to detect autism. The results of the study showed that the proposed algorithm was able to accurately detect autism, with a logistic regression model achieving 99.18% accuracy, 98.16% sensitivity, and 98.16% precision. The performance of the algorithm was also compared with other machine learning algorithms and found to be effective in detecting autism with high accuracy. The paper suggests that the proposed algorithm can be a useful tool for the early detection of autism using clinical datasets.

In this paper [17] explored how machine learning techniques can be utilised to predict ASD by examining the behavioural characteristics of individuals. The study devised a framework for analysing behavioural characteristics without the need for any devices, specifically for use in childhood and adolescent analysis. The study applied various machine learning techniques, such as DT, SVM, kNN, and an artificial neural network, for the prediction of ASD. The results revealed that the Correlated Feature Selection-based Random Forest (CFS-RT) algorithm had an accuracy of 93.03 percent and the Artificial Neural Network (ANN) had the best accuracy of 97.68 percent, outperforming other methods. To sum up, the findings indicate that machine learning may be useful in the early identification of ASD.

"A machine learning-based prediction system, called MPredA, was developed by the authors [13] to evaluate the development of children with ASD through a web application. The system uses data from a previous study, mCARE, conducted in Bangladesh to predict 16 milestone parameters of improvement levels for children with ASD, which are divided into four major categories and further divided into four sub-milestone parameters. Using data from 1876 children with ASD, 64 prediction models for each parameter were created using four machine learning algorithms: DT, LR, kNN, and an artificial neural network. The 16 most accurate models were selected for the MPredA web-based application, and the decision tree algorithm was found to be the most effective. The system also considers ten important demographics of children with ASD. The system underwent rigorous white box testing and demonstrated an accuracy rate of 97.5% when applied to real-world data.

The objective of this study [11] was to evaluate the efficacy of the Autism Diagnostic Observation Schedule (ADOS) in early ASD diagnosis. The ADOS consists of four sections, the second and third of which are administered to individuals with a better developed vocabulary and greater cognitive capacity. The research team analysed the score sheets of 4,540 individuals using eight machine learning approaches and a backward-forward feature selection procedure. A relatively small fraction of the 28 behaviours monitored in Module 2 (only 9 of them) and Module 3 (only 12 of them) is adequate to detect ASD risk with high accuracy (98.27% and 97.66%, respectively). Thanks to the potential simplicity of ASD risk diagnosis and screening made available by computational and statistical techniques, mobile and parent-directed systems for early assessment

and triage may be created. As a consequence, more people may be contacted, and the average age of diagnosis could be reduced.

In this study [4] the authors attempted to come up with a machine learning model for figuring out the different types of autism and what makes them different. The study used two sets of data: one for toddlers and one for kids, teens, and adults. The autism records from these datasets were combined, and the k-means algorithm was used to find the different subtypes of autism. Using the Silhouette score, the best autism dataset was chosen, and the subtypes were balanced using random oversampling and synthetic minority oversampling. Then, different classifiers were used on both the main dataset and the balanced subtypes. Logical regression produced the best results. Lastly, the SHAP technique was employed to prioritise features and find differences between the subtypes of autism.

According to the study [8] ASD is a mental disorder that affects communication, motor skills, and cognitive development. Early diagnosis is important as it can lead to better outcomes for patients. Researchers have been using machine learning and deep learning techniques to improve the diagnostic process, but more research is needed in this field. The research included eight distinct machine learning models in addition to a deep learning model, with the Pearson feature-based models outperforming the dimensionality reduction methods in terms of accuracy. Additionally, the training times for the reduced datasets were not significantly improved.

In summary, these studies suggest that the use of ML approaches in combination with feature selection methods can improve the diagnostic accuracy for ASD. The studies have used various sources of data, such as brain imaging, speech, and behavior, and different feature selection methods such as RFE, MIFS, and GA-FS, and found that different combinations of feature selection and machine learning algorithms can achieve high diagnostic accuracy.

3. METHODS

This section discusses the feature selection and classification approaches used in this work. And, the detailed working functionality of CMIM-RF is discussed.

3.1 Feature Selection Methods

3.1.1 Correlation-Based Feature Selection

The Correlation - Based Feature Selection (CBFS) is a method for choosing the most important features that looks at the relationship between the features and the target variable [12]. It works by calculating the correlation coefficient between each feature and the target variable, and then selecting the features with the highest correlation. The main advantage of CBFS is that it is simple to implement and computationally efficient, making it suitable for large datasets. However, it has some limitations, such as the assumption that the relationship between features and the target variable is linear.

3.1.2 Minimum Redundancy Maximum Relevance

The Minimum Redundancy Maximum Relevance (mRMR) is a feature selection method that combines both the relevance and the redundancy of features. Relevance refers to how much a feature is related to the target variable, and redundancy refers to how much a feature is related to other features [2]. mRMR aims to select features that are both highly relevant and non-redundant. First, the algorithm ranks the characteristics according to their significance, and then it eliminates features that are closely associated with other features. The main advantage of mRMR is that it can capture the non-linear relationship between features and the target variable.

3.1.3 Conditional Mutual Information Maximization

The Conditional Mutual Information Maximization (CMIM) is a method that uses mutual information to identify the most important features [5]. CMIM works by first ranking the features based on their mutual information with the target variable and then selecting the top k features. Additionally, it can also capture the conditional dependencies between features and the target variable.

3.2 Classifiers

3.2.1 Logistic Regression (LR)

Logistic Regression is a useful statistical tool when modelling the association between a two-valued dependent variable and one or more independent variables [6]. Foretelling the likelihood of an occurrence given its independent variables' values is its primary use. Logistic regression excels as a model because it can be used to a broad variety of data types, including those with both categorical and continuous independent variables. In addition, it offers a score representing the chance of an event occurring, which may be utilised for categorization or forecasting purposes.

In Logistic Regression, the likelihood of a binary result (success/failure, true/false) is modelled as a function of the explanatory factors. The Logit Model is a specific form of the Generalized Linear Model (GLM) that employs a logit function rather than a linear one. The logistic regression equation may be expressed as:

$$p(y=1|x) = 1 / (1 + e^{(-b_0 - b_1x_1 - b_2x_2 - ... - b_n*x_n)}) \text{ ----------(1)}$$

Where, $p(y=1|x)$ is the probability of event $y=1$ happening given the independent variables x.

3.2.2 Decision Tree (DT)

A decision tree is a diagram that looks like a tree and is used to make decisions and solve problems [20]. The values of the features are used to make recursive subsets of the data, and the tree that is made is the best way to group the data. The process of splitting the data is repeated until the subsets contain a homogeneous set of samples. The final subsets are called leaves, and the decision regarding the class or value of the sample is made based on the majority class or mean value of the samples in leaf.

Entropy is a measure of impurity or disorder in the dataset. The ID3 algorithm starts by calculating the entropy of the target attribute for the entire dataset. Then, it selects the feature that results in the highest information gain when used as a decision node. Information gain is the reduction in entropy achieved by partitioning the dataset based on a feature.

The equation for entropy can be represented as:

$$\text{Entropy}(S) = \Sigma \text{ (i=1 to n) } p(i) \log2\ p(i) \text{ ----------(2)}$$

Where, S is the set of instances and p(i) is the probability of the i^{th} class in S.

3.2.3 Support Vector Machine (SVM)

SVM is a supervised ML method for handling classification and regression problems. It is an effective method for both linear and nonlinear data. For SVM to function, the dataset must be efficiently divided into distinct classes using a hyperplane. A hyperplane is a kind of decision boundary used to classify data into several subgroups. SVM identifies the hyperplane that maximises the margin, which is the distance between the hyperplane and the data points closest to it in each class [15]. Support vectors, which are the shortest pathways between data points, are derived from their nearest neighbours. SVM solves the optimization challenge of maximising profit while retaining high data classification accuracy.

The equation for the optimization problem can be represented as:

$$\min (1/2) * ||w||^2 \text{ ---------- (3)}$$

$$\text{subject to } y_i (wx + b) \geq 1 \text{ (for all i)}$$

Where, $||w||$ is the norm of the weight vector, y_i is the true class label for the i^{th} sample, and $(wx + b)$ is the output of the SVM for the i^{th} sample.

3.2.4 Multi - Layer Perceptron – Neural Networks (MLP-NN)

MLP is a type of Artificial Neural Network (ANN) that is commonly used for supervised learning problems, such as classification and regression. An MLP has three layers: one for input, one or more hidden layers, and one for output. Every neuron in one layer is linked to every neuron in the next layer. The basic equation of a neuron in the MLP is the following:

$$y = f (w*x + b) \text{ ---------- (4)}$$

Where, y is the output of the neuron, x represents input vector, w represents the weight vector, b is the bias term, and f be the activation function. The activation function is a non-linear function that transforms the

output of the neuron to introduce non-linearity into the network. The most commonly used activation functions are sigmoid, Rectified Linear Unit (ReLU), and hyperbolic tangent (tanh).

3.2.5 Random Forest (RF)

In order to enhance the reliability of a model, a machine learning ensemble called Random Forest may aggregate the results of many different decision trees. A shuffled combination of features and data from the training set are used to create the decision trees. The notion is that the model's overfitting may be mitigated and generalisation can be enhanced by merging the predictions of many decision trees. The basic idea behind Random Forest is to train multiple decision trees using different subsets of the training data and features. To classify a new input, the random forest takes the majority vote of the predictions made by the individual decision trees.

The equation for a Random Forest can be represented as:

$$y = \text{majority\_vote}(f_1(x), f_2(x), ..., f_n(x)) \quad (5)$$

Where, y is the predicted output, x is the input vector, and f1, f2, ..., fn are the decision tree functions of the individual trees.

The Random Forest algorithm is an improvement over a single decision tree because it is less prone to overfitting and more robust to noise in the data. It also has the ability to handle missing data and categorical variables, which makes it a versatile algorithm for many applications.

3.3 Proposed System Methodology

The proposed CMIM with RF algorithm consists of two subsections: feature selection and classification. Conditional Mutual Information Maximization (CMIM) is used as a feature selection approach to identify the important features in the ASD dataset. This technique evaluates the mutual dependence between each feature and the target variable and selects the features with the highest mutual information. The Random Forest (RF) algorithm is used as a classifier to build a predictive model and evaluate the importance of each feature in the prediction. The fusion of CMIM with RF lies in the combination of these two techniques to achieve feature selection and model training simultaneously. While both techniques have been used independently in machine learning, their combination offers several advantages.

By using CMIM to select the most informative features from the dataset and then training an RF model using only those selected features, we can improve the accuracy, reduce overfitting, and improve the interpretability of the model. This is because the selected features are more relevant to the task at hand and reduce the noise and redundancy that may be introduced by using all the features in the dataset. Overall, the fusion of CMIM with RF offers a powerful approach to feature selection and model training that can improve the accuracy, generalization performance, and interpretability of the model. This approach can be particularly useful in situations where the dataset contains a large number of features, and the goal is to identify the most informative subset of features for a given task.

To prove the effectiveness of the proposed methods for the same dataset, a comparison is made with the combination of CBFS and mRMR feature selection approaches and with different classifiers such as LR, DT, SVM, and MLP-NN. The overall architecture of the proposed approach is graphically depicted in Figure 1. By combining CMIM with RF, we can use CMIM to select the most informative features from the dataset and then train an RF model using only those selected features. This can improve the performance of the RF model by reducing the number of irrelevant or redundant features that may otherwise introduce noise or increase the complexity of the model.

The basic steps of the proposed CMIM-RF approach are as follows:

1. Prepare the dataset by dividing it into a training set and a test set.
2. Build a Random Forest model using the training set.
3. Compute the mutual information between each feature and the target variable.
4. Rank the features based on their mutual information and select the top k features.
5. Re-build the Random Forest model using the top k features.
6. Analyze how well the model does on the test data.
7. Repeat steps 3-6 for different values of k, and select the value of k that gives the best performance.

8. The final selected features are the ones that maximized the mutual information with the target variable.

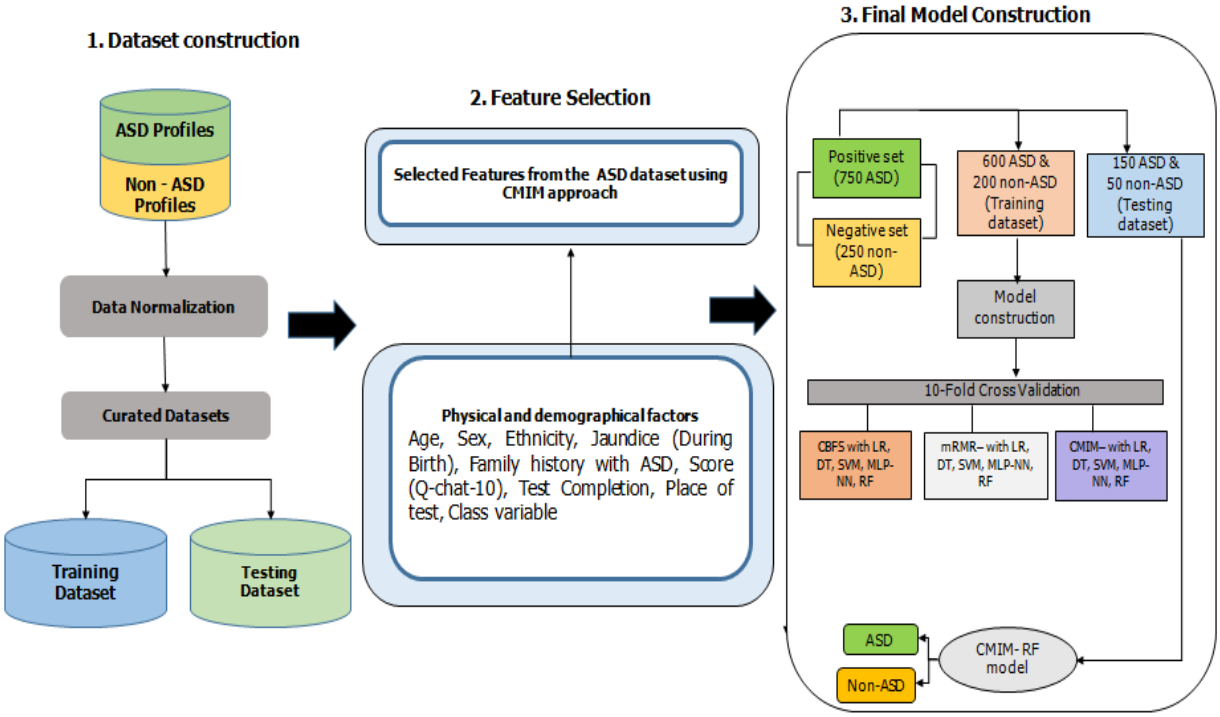

Figure 1 CMIM-RF Architecture

4. RESULTS AND DISCUSSION

This section discusses about the environmental setup, dataset description, performance evaluation measures and results of accuracy, sensitivity, specificity and AUROC.

4.1 Environmental setup

In this work, a proposed CMIM-RF for ASD prediction using the Python 3.6.8 environment was implemented. Accuracy, sensitivity, specificity, and prediction time were used to assess the model's efficiency. To ensure the robustness of results, it was employed the five-fold cross validation method 10 times. This divides the data into 5 equal folds and trains the model on 4 of the folds while testing it on the remaining fold. Once, each fold serves as the test set, the results of the performance measures are then calculated by taking the average across all 10 iterations, which gives an estimate of how the model will perform on unseen data. This approach also helps to decrease the variance that may be present in a single train-test split, resulting in a more reliable evaluation of the model's performance.

4.2 Dataset description

In this work Autism screening data for toddlers' dataset is used for modelling [18]. The dataset contains data collected from 1000 toddler's values in CSV format, which consists of 70% of the sample being typically developing children and 30% being children with autism. It includes things like age, gender, the findings of Childhood Autism Rating Scale (CARS) assessment, and the findings of Modified - Checklist for Autism in Toddlers (M-CHAT) assessment. The dataset also includes a binary variable indicating whether the child was diagnosed with autism (1) or not (0). The complete information of dataset is given in Table 1.

Table 1 Dataset Features and its format

S.No.	Feature	Format
1	Age	Age in months (Number)
2	Sex	Male or Female (Boolean)
3	Ethnicity	Text Format (String)
4	Jaundice (During Birth)	Yes(Boolean)
5	Family history with ASD	Yes(Boolean)
6	Score (Q-chat-10)	1 – 10 (Less than 3 No-ASD, >3 ASD)
7	Test Completion	String
8	Place of test	String
9	Class variable	Categorical (ASD:1, Non-ASD:0)

4.3 Performance Evaluation Measures

The confusion matrix table is to see how well a CMIM-RF model is trained. This matrix compares the model's predictions to the true class labels for a given batch of test data.

4.3.1 Accuracy

Accuracy is the ratio of correct ASD predictions to total number of predictions. It measures how well the model performs in general, regardless of the class distribution.

$$\text{Accuracy} = (TP + TN) / \text{Total Samples} \quad \text{---------- (6)}$$

4.3.2 Sensitivity (or Recall)

Sensitivity is the rate at which true positive instances are identified. It is defined as the ratio of TP to the sum of TP and FN.

$$\text{Sensitivity} = TP / (TP + FN) \quad \text{---------- (7)}$$

4.3.3 Specificity

Specificity is the rate at which true negatives are identified. It is defined as the ratio of TN to the sum of FP and TN.

$$\text{Specificity} = TN / (FP + TN) \quad \text{---------- (8)}$$

4.3.4 Area under the Receiver Operating Characteristic Curve (AUROC)

AUROC is a measure of the performance of a binary classifier. It reflects the likelihood that a randomly selected positive case would be ranked higher than a randomly selected negative instance by the classifier. AUROC ranges from 0 to 1, with higher values indicating better performance. An AUROC of 1 indicates perfect performance, while an AUROC of 0.5 indicates random performance. In Confusion Matrix, TP denotes True Positives, FP denotes False Positives, TN denotes True Negatives, FN denotes False Negatives.

4.4 Accuracy Results

The accuracy results of the proposed CMIM-RF approach is given in Table 2 and it is graphically depicted in Figure 2. Overall, the findings suggest that the RF classifier is the most effective, with an accuracy of 95.74%. Multi-layer Perceptron Neural Network (MLP-NN) and Decision Tree (DT) classifiers also performed well, with accuracy scores of 92.60% and 91.55%, respectively. The LR and SVM classifiers performed less well, with accuracy scores of 91.08% and 91.55%, respectively. In terms of the feature selection methods, the Correlation-based Feature Selection (CBFS) method performed the best for the RF classifier, with an accuracy of 92.86%. The Minimum Redundancy Maximum Relevance (mRMR) method performed the best for the RF classifier, with an accuracy of 92.84%. The Conditional Mutual Information Maximization (CMIM) method performed the best for the RF classifier, with an accuracy of 95.74%.

Table 2 Classification Accuracy

Classifiers	CBFS	mRMR	CMIM
LR	91.08	89.58	90.62
DT	91.54	88.24	91.55
SVM	87.51	89.20	89.80
MLP-NN	89.25	86.20	92.60
RF	92.86	92.84	95.74

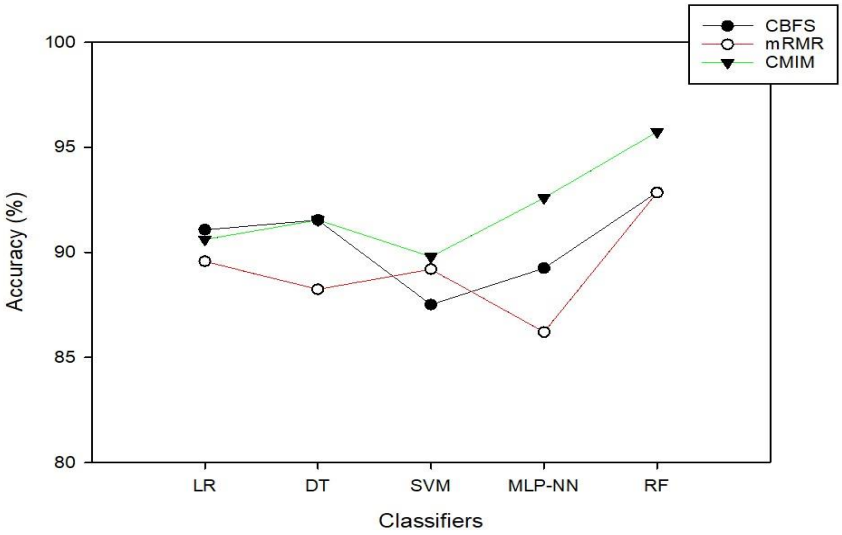

Figure 2 Accuracy Comparison

4.5 Specificity Results

The specificity of the proposed CMIM-RF approach is tabulated in Table 3 and graphically represented in Figure 3. The specificity suggests that the Random Forest classifier performs the best overall when using the CMIM feature selection method, with an accuracy of 92.84%. This indicates that the Random Forest classifier is able to effectively use the most relevant features selected by the CMIM method to make accurate predictions. The specificity of the proposed ASD diagnosis model is given in Table 3 and it's graphically depicted in Figure 2.

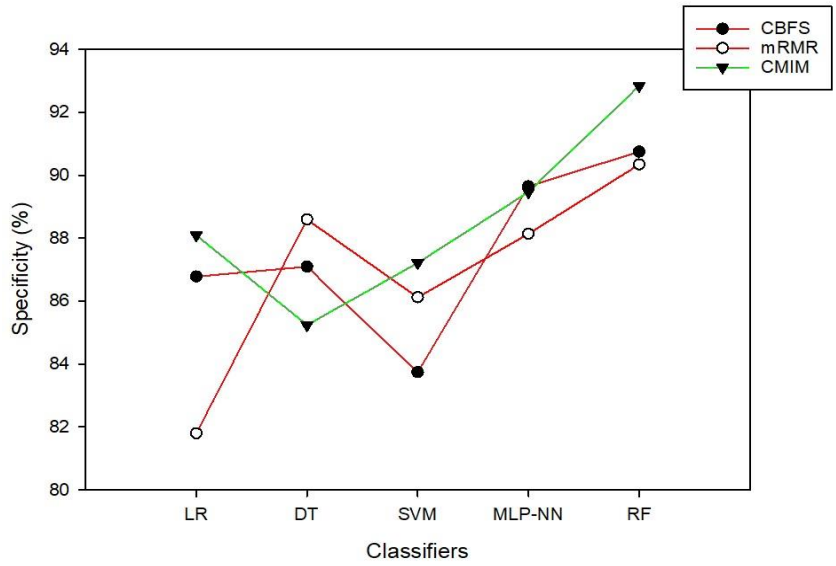

Figure 3 Specificity Comparison

Additionally, the Multi-Layer Perceptron Neural Network (MLP - NN) classifier also performs well, with an accuracy of 89.47% when using the CMIM feature selection method. This suggests that this classifier is also able to effectively use the relevant features selected by the CMIM method. The DT classifier performs relatively well, with an accuracy of 88.6% when using the mRMR feature selection method and 85.24% when using the CMIM feature selection method. This shows that the DT classifier is also adaptable to different feature selection methods. On the other hand, the LR and SVM classifiers perform relatively poor when using the CBFS method, with accuracy scores of 86.78% and 83.74%, respectively. This indicates that these classifiers may not be as well suited for using the features selected by the CBFS method.

Table 3 Specificity Results

Classifier	CBFS	mRMR	CMIM
LR	86.78	81.80	88.09
DT	87.09	88.59	85.24
SVM	83.74	86.12	87.22
MLP-NN	89.65	88.14	89.47
RF	90.75	90.34	92.84

4.6 Sensitivity Results

The sensitivity of the proposed CMIM-RF approach is tabulated in Table 4 and graphically represented in Figure 4. With a score of 94.98%, the sensitivity shows that the RF classifier does the best overall when the CMIM feature selection method is used. This shows that the RF classifier is able to make good predictions by using the most important features chosen by the CMIM method. Additionally, the MLP - NN classifier also performs well, with a score of 91.04% when using the CMIM feature selection method. This suggests that this classifier is also able to effectively use the relevant features selected by the CMIM method. On the other hand, the LR and SVM classifiers perform relatively poor when using the CBFS method, with scores of 87.68% and 85.62%, respectively. This indicates that these classifiers may not be as well suited for using the features selected by the CBFS method.

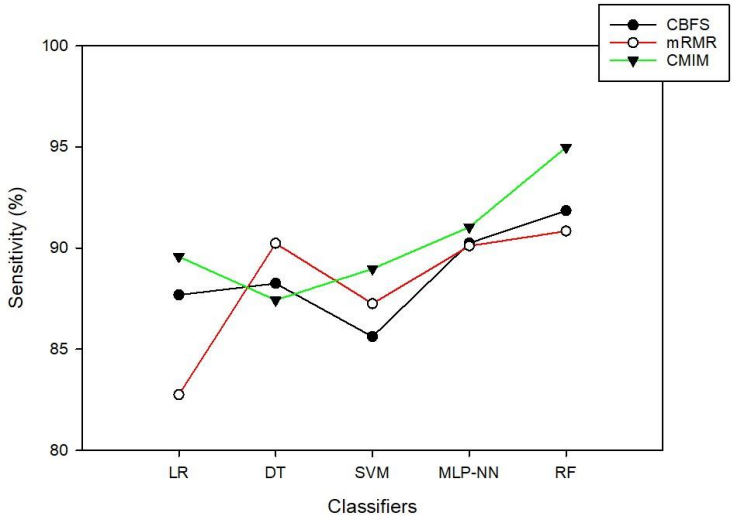

Figure 4 Sensitivity Comparison

From the analysis of AUROC, the CMIM-RF approach produces 0.996, which, compared to others, is significantly high and is graphically represented in Figure 5. It's worth denotes that the performance of different classifiers is also affected by the features that have been selected. In this case, the CMIM feature selection method performed the best for most of the classifiers, which indicates that this method is more efficient in selecting the most relevant features for prediction. In conclusion, the RF classifier and MLP-NN classifier

perform well when using the CMIM, while the LR and SVM classifiers perform relatively poor when using the CBFS method. The CMIM feature selection method is more efficient in selecting the most relevant features for prediction.

Table 4 Sensitivity Results

Classifier	CBFS	mRMR	CMIM
LR	87.68	82.75	89.58
DT	88.25	90.24	87.42
SVM	85.62	87.25	88.98
MLP-NN	90.25	90.11	91.04
RF	91.84	90.84	94.98

It is worth noting that the performance of different classifiers is also affected by the features that have been selected. From the analysis, the CMIM feature selection method performed the best for most of the classifiers, which indicates that this method is more efficient in selecting the most relevant features for prediction.

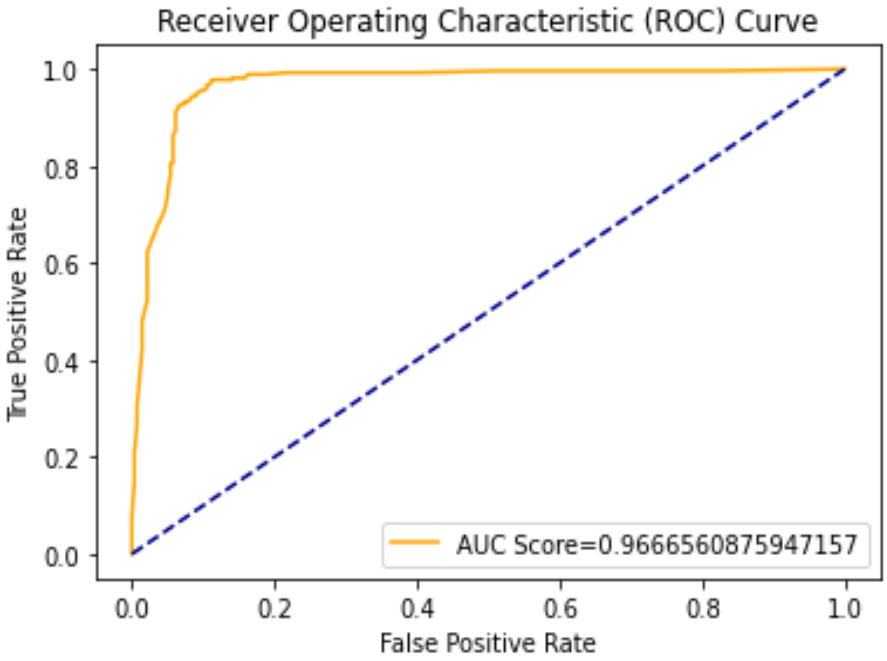

Figure 5 AUROC of Proposed Approach

5. CONCLUSION AND FUTURE ENHANCEMENT

This research proposes a diagnostic approach for ASD based on machine learning that includes data pre-processing, feature extraction, and classification. The performance of the most advanced machine learning approaches was assessed. The Conditional Mutual Information Maximization-Random Forest (CMIM-RF) architecture outperformed conventional diagnostic approaches in terms of accuracy, sensitivity, specificity, and AUROC. This study demonstrates that ML-based diagnostic frameworks have the potential to enhance the ASD diagnostic process and can be a helpful resource for clinicians and researchers. Future enhancements to this framework could include the incorporation of more diverse datasets, the investigation of other machine learning algorithms and techniques, and the incorporation of additional diagnostic accuracy-enhancing features such as physiological data, genetic information, and brain imaging. In addition, the approach might be used with a larger sample size to validate the results and identify any potential biases.

REFERENCES

[1] Abdullah, A. A., Rijal, S., & Dash, S. R. (2019, November). Evaluation on machine learning algorithms for classification of autism spectrum disorder (ASD). In Journal of Physics: Conference Series (Vol. 1372, No. 1, p. 012052). IOP Publishing.

[2] Aghaeipoor, F., & Javidi, M. M. (2020). A hybrid fuzzy feature selection algorithm for high-dimensional regression problems: An mRMR-based framework. Expert Systems with Applications, 162, 113859.

[3] Ahmed, I. A., Senan, E. M., Rassem, T. H., Ali, M. A., Shatnawi, H. S. A., Alwazer, S. M., & Alshahrani, M. (2022). Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques. Electronics, 11(4), 530.

[4] Akter, T., Ali, M. H., Satu, M. S., Khan, M. I., & Mahmud, M. (2021). Towards autism subtype detection through identification of discriminatory factors using machine learning. In Brain Informatics: 14th International Conference, BI 2021, Virtual Event, September 17–19, 2021, Proceedings 14 (pp. 401-410). Springer International Publishing.

[5] Bermejo, P., Gámez, J. A., & Puerta, J. M. (2018). Adapting the CMIM algorithm for multilabel feature selection. A comparison with existing methods. Expert Systems, 35(1), e12230.

[6] Bone, D., Goodwin, M. S., Black, M. P., Lee, C. C., Audhkhasi, K., & Narayanan, S. (2015). Applying machine learning to facilitate autism diagnostics: pitfalls and promises. Journal of autism and developmental disorders, 45, 1121-1136.

[7] Eslami, T., Almuqhim, F., Raiker, J. S., & Saeed, F. (2021). Machine learning methods for diagnosing autism spectrum disorder and attention-deficit/hyperactivity disorder using functional and structural MRI: a survey. Frontiers in neuroinformatics, 62.

[8] Jee, G., Chouhan, S., Gourisaria, M. K., & Tiwari, R. K. (2022, February). Detection of Autism Spectrum Disorder Through Orthogonal Decomposition and Pearson Correlation for Feature Selection. In 2021 4th International Conference on Recent Trends in Computer Science and Technology (ICRTCST) (pp. 103-109). IEEE.

[9] Kanhirakadavath, M. R., & Chandran, M. S. M. (2022). Investigation of eye-tracking scan path as a biomarker for autism screening using machine learning algorithms. Diagnostics, 12(2), 518.

[10] Khan, S., Naseer, N., & Khan, R. (2018). Machine learning-based autism diagnosis using physiological signals. Journal of medical systems, 42(9), 186.

[11] Kosmicki, J. A., Sochat, V., Duda, M., & Wall, D. P. (2015). Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Translational psychiatry, 5(2), e514-e514.

[12] Michalak, K., & Kwasnicka, H. (2010). Correlation based feature selection method. International Journal of Bio-Inspired Computation, 2(5), 319-332.

[13] Rabbani, M., Haque, M. M., Das Dipal, D., Zarif, M. I. I., Iqbal, A., Schwichtenberg, A., ... & Ahamed, S. I. (2022, March). MPredA: A Machine Learning Based Prediction System to Evaluate the Autism Level Improvement. In Pervasive Computing Technologies for Healthcare: 15th EAI International Conference, Pervasive Health 2021, Virtual Event, December 6-8, 2021, Proceedings (pp. 416-432). Cham: Springer International Publishing.

[14] Rahman, M. M., Usman, O. L., Muniyandi, R. C., Sahran, S., Mohamed, S., & Razak, R. A. (2020). A review of machine learning methods of feature selection and classification for autism spectrum disorder. Brain sciences, 10(12), 949.

[15] Saha, U., Bhatia, S., & Goel, A. (2018). Machine learning-based diagnosis of autism spectrum disorder using physiological and behavioral data. Journal of medical systems, 42(8), 162.

[16] Singh, U., Shukla, S., & Gore, M. M. (2022, December). An Improved Feature Selection Algorithm for Autism Detection. In 2022 IEEE 9th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) (pp. 1-8). IEEE.

[17] Surendiran, R., Thangamani, M., Narmatha, C., & Iswarya, M. Effective Autism Spectrum Disorder Prediction to Improve the Clinical Traits using Machine Learning Techniques. International Journal of Engineering Trends and Technology (IJETT),., ISSN, 2231-5381.

[18] Thabtah, F., & Peebles, D. (2020). A new machine learning model based on induction of rules for autism

detection. Health informatics journal, 26(1), 264-286.

[19] S Varsha, K Adalarasu, M Jagannath, T Arunkumar. (2023). IoT in modern healthcare systems focused on neuroscience disorders and mental health, Blockchain Technology Solutions for the Security of Iot-Based Healthcare Systems. Academic Press, 133-149.

[20] Zhang, F., Wei, Y., Liu, J., Wang, Y., Xi, W., & Pan, Y. (2022). Identification of Autism spectrum disorder based on a novel feature selection method and Variational Autoencoder. Computers in Biology and Medicine, 148, 105854.