ASSESSING THE COMBINED DRUG EFFICACY FOR ADHD TREATMENT IN ADULTS AND CHILDREN USING A NOVEL BIG DATA ANALYTICAL MODEL: A DOSAGE-DEPENDENT STUDY

Mr. Ramakrishnan Varadharajan[1], Dr. P. Anbalagan[2], Dr.M.S.Saravanan[3]

[1] Research Scholar, Department of Computer Science and Engineering,
Annamalai University, Annamalai Nagar, Tamil Nadu. India. Pincode:608002

[2] Assistant Professor, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Tamil Nadu. India. Pincode:608002.

[3] Professor, Department of Computer Science and Engineering, Saveetha School of Engineering, Chennai, Tamil Nadu, India. Pincode:602107

Abstract

A common neurodevelopment condition known as attention deficit hyperactivity disorder (ADHD) is defined by age-inappropriate symptoms of inattention and hyperactivity that continue into adulthood in the majority of children who are diagnosed with the condition as a child. The purpose of this study was to find out how combo medications work for treating ADHD patients. Both children and adults were included in the research. For this investigation, a total of 1500 patient records were collected. The Map Reduce method was used to eliminate duplicate data missing values and to create effective data partitioning. Recursive Feature Elimination (RFE) was employed to remove less significant features from a dataset and improve machine learning model performance. To find the efficacy of the drug used to treat ADHD Brain disorder in adults & children in this study, we proposed a time-dependent flexible deep recurrent network (TDF-DRN) technique. According to preliminary research, combining medication therapy can be a successful strategy for treating the symptoms of ADHD in both adults and children. The best medicine combinations and dosages seem to differ depending on the age group. Compared to adults, children need fewer dosages and benefit from various combinations. Additionally, the safety concerns and potential adverse effects of different drug combinations are assessed.

Keywords: Attention Deficit/Hyperactivity Disorder (ADHD), Combinations of Drugs, Recursive Feature Elimination (RFE), Time-Dependent Flexible Deep Recurrent Network (TDF-DRN)

1. Background of the study

The symptoms and signs of Attention Deficit Hyperactivity Disease (ADHD), a developmental disorder, consist of hyperactivity, daydreaming, as well as distractibility, either accompanied by or without impulsivity. The main characteristic of ADHD is an accumulation of symptoms that lasts longer than is typical for children of the same age, including impulsivity and hyperactivity. Children of Mexican descent consistently have a lower prevalence of ADHD than children of other racial or ethnic groups. At the pediatric age, boys have a majority of ADHD, which is three to five times higher than girls. A lack of data on the proportions in each category, although at least fifteen to twenty percent of children diagnosed with ADHD continue to have the complete diagnosis as adults [1].

Stress, sadness, resistance to change, and drug use disorders are the most common comorbidities with ADHD, which affects around 70% of instances. To diagnose ADHD, signs of inattention begin before the age of twelve and produce moderate to severe impairments in psychological functioning, social interactions, academic performance, and vocational performance as determined through interviews and observations in various settings [2]. Figure 1 shows the symptoms of ADHD in children.

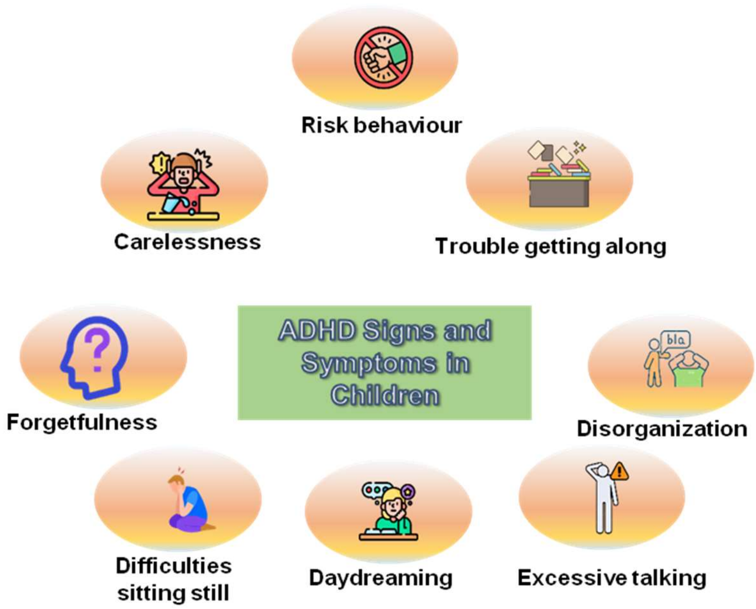


Figure 1: ADHD signs in Children

1.2 Identifying Young Children with ADHD

It is challenging to diagnose ADHD in very young children despite the reality that symptoms of the illness are occurring in adolescents or children. Accordingly, children who are preschool age or younger and have been identified as possessing ADHD are probably to require assessment by a specialized professional, such as a child's developmental pediatrician, psychiatrist, or psychologist [3].

1.3 Drug Efficacy for ADHD

To treat children with ADHD who are between the ages of 7 and 12 who do not currently use prescription ADHD medication. The gadget produces minimal stimulation via electricity, which is delivered through the child's forehead along with stimulating using microscopic areas. When considering eTNS, it's critical to talk about expectations, safety measures, and any negative impacts. The effectiveness of drugs is an essential aspect of the all-encompassing treatment strategy for people with this neurological disease [4]. ADHD symptoms are mainly managed with the help of pharmaceutical therapies, including stimulant and non-stimulant drugs. To increase the activity of neurotransmitters, including norepinephrine and dopamine, which improve cognitive function and impulse control, healthcare professionals frequently provide stimulants like a drug called amphetamine derivatives. A primary focus is comprehending the effects that depend on dosage, acknowledging that the ideal balance is different for adults and children. When creating effective treatment programs, variables such as the developmental stage, metabolic variations, and possible long-term consequences on cognitive function must be taken into account [5].

1.4 Impulsivity and Hyperactivity

Young children have to demonstrate a minimum of 5 hyperactivity-impulsivity signs, whereas adults and teenagers who are seventeen or older should show a minimum of six. The signs are required to have lasted for at least a month and be disruptive and unsuitable for the individual's stage of development. They obtain standing with a heartbeat while it is expected to remain sitting; they move about or peek when they don't seem suitable [6]. Figure 2 shows the inflammation in ADHD.

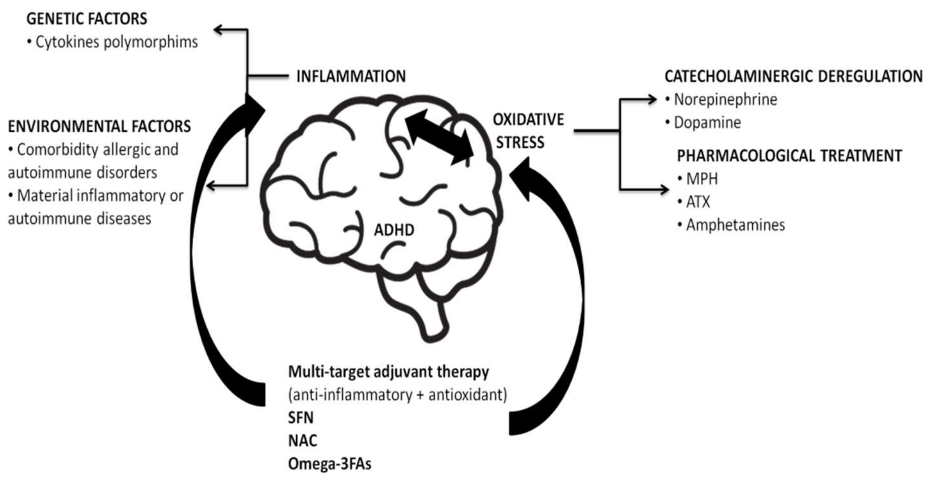


Figure 2: Function of inflammation in ADHD

Adolescent and adult use of drugs, risky sexual behavior, and criminal behavior are among the problematic effects associated with the main symptoms of ADHD. There is a strong correlation between conduct difficulties and hyperactivity-impulsivity-inattention, which places people with those signs at a higher risk of engaging in persistent criminal activity, especially when it comes to antisocial and transferring behaviors. However, distinctions become apparent with closer inspection of individual symptoms and subtypes, especially regarding impulsivity and co-occurring psychopathology [7].

The goal of the study is to guide medical professionals in developing individualized, extremely successful pharmacological treatments for ADHD. Seeking a nuanced knowledge of combined medication efficacy in the treatment of ADHD ultimately contributes to the larger objective of improving the level of life for individuals impacted by this complicated neurodevelopment illness, regardless of developmental stage.

Key contributions

- An essential addition is the preparation of data using the Map Reduce methodology.
- This assists in providing a targeted and pertinent set of characteristics for estimating the effectiveness of ADHD treatment.
- According to the study, treating ADHD symptoms in both adults and children can be accomplished by combining drugs and other therapies.
- The study employs a comprehensive set of evaluation metrics for assessing the efficacy of the model, including Sensitivity, Specificity, Precision, Accuracy, AUC Curve, F1 Score, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).

In section 2 of the paper, we combine a literature review for background and insight. Section 3 provides a deeper dive into the approach. In Section 4, we present an in-depth evaluation and discussion of the results. In section 5, the relevance of the conclusion is discussed in depth.

2. Related works

The study [8] assessed the tolerability, long-term safety, and effectiveness of cannabinoids in pediatric epileptic patients. Concomitant antiepileptic medication trough concentrations were evaluated at baseline, one, two, and three months into therapy, and as further clinically warranted. The paper [9] considered meta-analyses of biofeedback applied to children with ADHD. A new meta-analysis of randomized controlled trials (RCTs) was carried out, updating earlier findings and including methodological changes. The study [10] suggested a novel deep learning-based automated diagnosis methodology to distinguish between children who have two distinct types of ADHD, namely Inattentive ADHD and Combined ADHD, utilizing the EEG signals of healthy children. A common neurodevelopment issue impacting many children is ADHD. Accordingly, computerized ADHD diagnosis provides a lot of potential benefits. The paper [11] focused on identifying the most essential characteristics and automating the diagnosis process with existing classification approaches for better diagnosis. ASD is a neurodevelopment disease that is accompanied by sensory problems, such as excessive or insufficient Sensitivity to touch, noises, or odors. The study [12] involved the development of a virtual reality classroom that was integrated with persistent and focused attention activities. The attention tasks had to be carried out in the presence of visual-audio hybrid distractions. The paper [13] focused on meta-analyses that have indicated that transcranial electrical stimulation can reduce clinical signs and enhance mental skills like memory retention and attention, which are compromised in ADHD. While pharmaceuticals account for the majority of treatments for ADHD, tES is becoming more and more popular as a substitute strategy.

The study [14] used the power spectrum, bicoherence, complexity, and biomarker candidates for detecting ADHD youngsters in a machine learning technique. Characterizing the EEG of children with ADHD is one of the critical objectives of this informative neuroimaging technology. One primary objective of the EEG, a proper neuroimaging technique for researching ADHD, is to describe the EEG characteristics of children with ADHD. The study [15] proposed to perform an auditory oddball task, and Event-Related Potentials from EEG signals were analyzed over time and frequency. According to the information gathered, a machine-learning model was created to distinguish between patients with ADHD and healthy control subjects. The paper [16] discussed the neurodevelopment illness known as ADHD, which is characterized by impulsivity, hyperactivity, and inattentiveness. The study [17] examined whether a comparable tendency would hold in natural home and school environments by following study participants into the next academic year. According to a study done in an identical summer therapy environment, many children with ADHD who also received behavioral intervention did not require medication or very low doses of it to enhance Sensitivity.

The study [18] assessed the interpolated dataset's capacity to discern between young people suffering from and without ADHD. It accomplished this by using a deep learning technique to fill in missing data in ADHD rating scales. The paper [19] focused on the diagnosis of ADHD, a neurodevelopment illness with a wide range of manifestations, based on objective accounts of symptoms. The establishment of neuroimaging-based diagnostics for the diagnosis of ADHD has benefited from the use of machine learning classifications. The paper [20] presented a comprehensive approach to diagnosing the mixed form of ADHD. In youths, ADHD occurs more than any other neurobehavioral disorder. Unfortunately, there are no established diagnostic approaches that are dependable and affordable because their cause is unclear. The study [21] found that stimulant drugs are the gold standard for treating ADHD. However, these medications do not work for every patient, and they can have adverse effects and be misused for reasons other than medical. Adults with ADHD endure personal anguish, and the economy suffers as a whole. The study [22] focused on an ANN-based Clinical decision support system (CDSS) to study the effectiveness of neurofeedback (NF) in the treatment of ADHD. To aid physicians in making more informed decisions regarding their patients, CDSS analyzes raw data and turns it into useful information.

Methods and materials

The data collection, data preprocessing, and feature extraction phases make up the first part of the method's multiple sections. We introduced a novel approach called the time-dependent flexible deep recurrent network (TDF-DRN) technique in this study, aiming to evaluate the effectiveness of the drug for treating ADHD in both adults and children. Figure 3 demonstrates the proposed method flow.

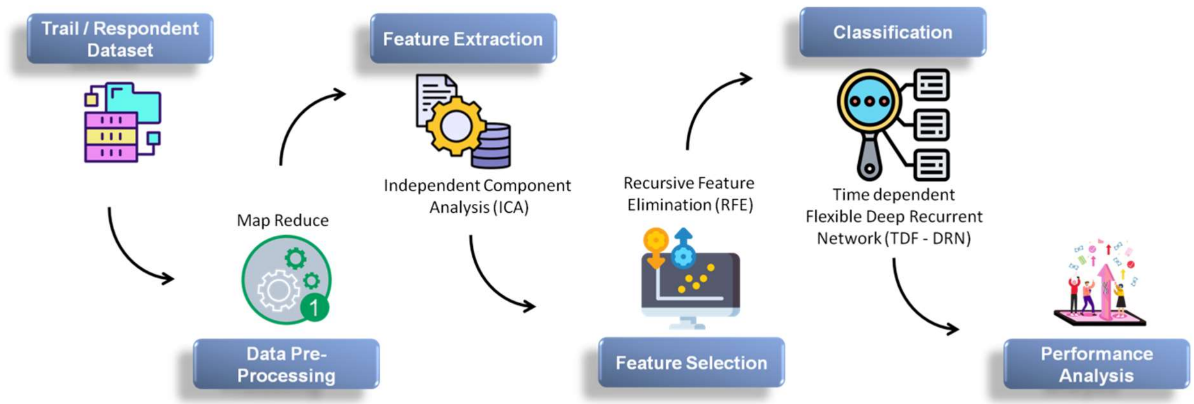


Figure 3: Block diagram for the proposed method

3.1 Data collection

This data is organized into 1500 rows and 728 columns. An assessment status is usually considered high or excellent. The digit 839,999 seconds could indicate an examination or assessment that lasted around 839,999 seconds if they represent a period in seconds. Based on a 24-hour day, this time is translated into approximately 9.72 days. There will be a significant amount of time spent on the assessment. They could stand for a particular category or kind in a system of classifications. It denotes any topic that is up for review, such as a specific class, group, assessment type, technique, product, etc. Lower-level outcomes or measurements might be represented by numbers like 0, 1, 2, and 3. Results or measures with values such as 10, 11, 12, 13, 14, and 15 can be moderate to high. A common occurrence in datasets is the use of the placeholder value "-1" to denote unavailable or missing data or a specific condition code. It might stand for instances where the data was not collected, the response was irrelevant, or neither of those things.

In most cases, the positive numbers stand for recorded data, answers, or measurements. The study or experiment that yielded these values has influenced. These positive integers and the unit of measure would have different meanings depending on the domain of the study. An entire dataset called Last Trial has its values set to 360.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

3.2 Data Preprocessing using Map Reduce

Data preparation is a crucial step in the process chain allowing the distributed processing during massive databases using the Map-Reduce computing style. Data cleansing, feature extraction, duplicate item removal from datasets, format conversion, and many other procedures are examples of data preprocessing steps. When processing large datasets, Hadoop Map Reduce offers the appropriate framework for carrying out various operations in parallel. Processing several heterogeneous datasets is made possible by the Map-Reduce paradigm. Where α and β stand for dataset lineages, r for variables, and w for variables-entity as shown in Equation (1-2).

$$\text{Map: } (r_1, w_1)_\alpha \rightarrow [(r_2, w_2)]_\alpha \quad (1)$$

$$\text{Reduce: } (r_2, [w_2])_\alpha \rightarrow (r_2, [w_3])_\alpha \quad (2)$$

The map function transforms an input key and variables couple $(r1, w1)$ into a list of intermediate key/variables pairs $[(r2, w2)]$. The compilation of items is aggregated using the reduction method $[r2]$connected to $r2$ and generates an amount collection $[w_3]$, resulting in a connection to $r2$. It ought to remain mentioned that each of the operations' inputs and results belong to the equivalent category α. For mathematical connections, the merging process yields a self-merge if $\alpha = \beta$, a procedure analogous to a self-adoption. Observe how nearly identical the Map and Reduce signatures in the new model are to those in the old Map Reduce. The dataset's ancestries and the fact that reduce produces key/variables list rather than simple integers are significant variations. Since the decreased outcome in Google Map Reduce is final, consumers can include items they need in $[w_3]$, and sending r_2 for the next step is not necessary. This is to ensure that before the data can be correctly combined, partitioning it first and sorting it to implement duplicate keys are required. However, it provided the data that the altered variables correspond to remains structured identically, as the data mapping variables represented expressed in r_2, values continue to be converted across phases, and they even have been considered identical according to the concept.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

3.3 Feature extraction using Independent Component Analysis (ICA)

One technique for extracting features from multivariate random signals that can be used to transform these features into greater significance features is Independent Component Analysis (ICA). Each observable signal is considered to be a linear combination of an equal number of independently floating indicators that have been measured or identified. Assuming $y = [y_1, y_2, \dots, y_i, \dots,$ [redacted]$]$ ym consists of m linear permutations, which are linear combinations. The combinations are the result of the m-way synthesis of separate linear components in Equation (3).

$$x_j = a_{j1}s_1 + a_{i2}s_2 + \cdots + a_{jn}s_n \tag{3}$$

Applying the symbol a_k, the n-way split could be broken down into its parts. Assuming a zero mean for the mixing variables and the independent components does not reduce generalizability. We can express the combinations using vector notation

$y = [y_1, y_2, \dots, y_i, \dots, y_m]$as y. The independent variables are denoted by the g in this context. The g is used to represent the independent variables in this situation. This paradigm is known as ICA. It is possible to estimate the inverse of matrix E and find its independent components. The ICA procedure is as follows: We can assume that the separate parts and mixing variables have mean values of 0 without limiting our generalizability, where ak is the total number of elements we desire. If we use vector notation to describe these combinations, we can refer to them as x. The following are the measures to take:

Step 1: Data normalization involves removing the Mean from the data to make it more consistent. So, ensures that the sum of the parts is never less than zero.

Step 2: Streamlining Data handling the data preparation differently. To achieve whiteness, one must manipulate the mixture in that the components no longer interact with one another, and their disparities cancel each other out.

Step 3: A certain amount of independence is necessary for the analysis of the data. Many more instances exist, such as the Fast ICA method, ICA with non-Gaussianity maximization, ICA with variance maximization or reduction, ICA with negative entropy, and numerous additional instances. Here, the Negentropy approach was employed.

Step 4: Reconstruct the information. To get the result, we increase the whitening output by the result we got after applying the independence requirement. To get the output, multiply the supplied data by the transpose.

3.4 Feature selection using Recursive Feature Elimination (RFE)

Our feature selection approach consists of two steps. RFE is a feature selection strategy that uses machine learning performance to eliminate features based on their relevance. To decrease the quantity of components utilized for model training, RFE is applied. We choose the number of features to use from the original parts which is why RFE is chosen. As a wrapper approach, RFE fits a learning model and eliminates less essential elements. This process is known as RFE. The model's coefficients, or feature importance properties, are used for evaluating the features. RFE attempts to remove potential correlations and collinearity in the framework by removing restricted amounts of characteristics for every repetition. RFE stipulates how many components must be retained; however, the initial number of acceptable characteristics is unspecified. To find the optimal number of features, cross-validation with RFE is employed to evaluate different feature subgroups and select the collection of characteristics with the greatest score.

3.5 Time-Dependent Flexible Deep Recurrent Network (TDF-DRN)

The concept of time-dependent flexibility refers to the algorithm's ability to modify and react in an adaptable manner to alterations over time when used in relation to the Time-Dependent Flexible Deep Recurrent Network (TDF-DRN).

Time-dependent flex describes an analytical approach or modeling strategy that permits flexibility in the handling of time-dependent data. The following suggests that the framework takes the data's time component. When treating ADHD, a patient's symptoms and medication response change over time. As a result, a time-dependent model considers the evolution and change of these variables over various time intervals. The flexibility implies that the model can change or adapt to different circumstances. It suggests that the model's structure is relatively flexible, allowing it to capture various patterns and dynamics in the data instead of remaining imprecise. According to TDF-DRN, the model is designed to manage time-dependent data flexibly and adaptively. Such flexibility is essential for identifying the complex and evolving nature of symptoms and drug responses in the context of treating ADHD. According to the temporal patterns found in the data, the model might be able to modify its predictions, giving an improved and dynamic depiction of the effectiveness of the treatment.

The desired designation is the treatment result, while the ideal sets of attributes have been collected to constitute the analyzing input. This study used a deep learning method, the Deep Recurrent Network (DRN) to achieve this goal. The major goal of this approach is to develop the algorithm to solve the specified problem more quickly and accurately. Additionally, it instructs the system multiple times to get a high-performance frequency. A modified version of the used LSTM, RNN, is typically used to overcome problems with mood forecasting. The method for the construction of the DRN model of structure is shown in Figure 4. As a result, the suggested method predicts young individuals' anxiety levels using the DRN technique. Multiple layers of RNN variables make up the framework.

Compared to traditional LSTM and RNN methods, DRN technology improves accuracy while reducing calculation time. Moreover, it gathers high-level data to forecast the label classification accurately. The optimal feature vectors are used to generate outcomes by fine-tuning the parameters of this approach. The total amount of filtering in this framework, which convolves the

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

input using a 1D Convolutional layer, is used to construct the function of activation. Next, the dimensionality of the feature map is computed using the framework (Equation (4)) as follows:

$$FM_{Z2} = \frac{J_{Z1} - CK_Z + 2B_G}{G_{Gj}} + 1 \quad (4)$$

Where B_g.G. indicates the padding size, G indicates the step size, FM_{Z2} indicates the dimension of the attribute mapping, IH_1 shows the width of the input signal before convoluted, and CK_Z shows the combination kernel's gradient that will be obtained. Equations (5) through (6-7) are used to obtain the regional characteristics of the input employing the following convolution kernel:

$$IH_{1E}(j) = l(WM_E.Y(j:j + E - 1) + P) \quad (5)$$

$$IH_{1E} = [IH_{1E}(1); IH_{1E}(2) \ldots .IH_{1E}(FM_{Z2})] \quad (6)$$

$$IHk'_{1E} = relu(IH_{1E}) \quad (7)$$

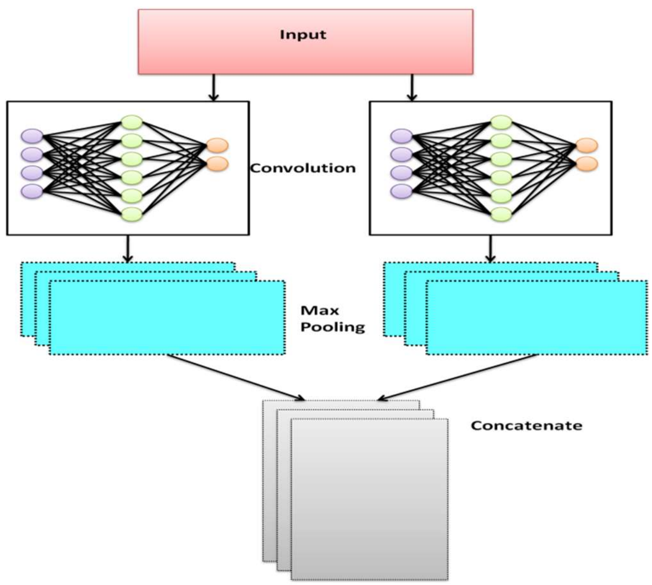


Figure 4: Architecture of a DRN

Where the result following convolution is denoted by IH_1, and WM_E means the kernel of the convolution operation has height E. As a consequence, all organized findings pertaining to the language collection that have been obtained are pooled using the maximum pooling layer, as shown by Equations (8) and (9).

$$IHk'_{1E} = max\ (IHk'_{1E}) \quad (8)$$

$$IHk' = Concatenate\ (IHk'_{1E_1}, IHk'_{1E_2},) \quad (9)$$

4. Result and discussion

A Windows 8 operating system, a 2.33 GHz CPU, and 6 GB of RAM were employed in

the investigation. Python was utilized throughout the testing process. The accuracy, precession,f1-score, recall, specificity, Sensitivity, Area under the Average Accuracy Curve (AUC Curve), Root Mean Squared Error (RMSE), and Mean squared Error (MSE)of each categorization procedure classifier were used to gauge the performance. The outcomes show that our recommended model performance is much better than other traditional methods like Random Forest (RF), and Classification Tree (CT) [23, 24, 25], Support vector machines (SVM) [25].

Accuracy

Accuracy is an outcome metric that measures a classifier's overall effectiveness. The overall number of reliable forecasts divided by the entire quantity of specimens is used to compute value. It demonstrates the accuracy with data samples are categorized using classification techniques. It is computed in accordance with Equation (10).

$$Accuracy = \frac{t_p+t_n}{t_p+t_n+f_p+f_n} \tag{10}$$

The relative effectiveness of Random Forest (RF) 83.3% and Classification Tree (CT) 88.9%) demonstrate how such techniques are capable of predicting how new medications will behave in clinical trials in addition to identifying important elements of the distinct cellular physiological processes that these drug classes produce. Figure 5 depicts the values that correspond to the accuracy measures. The suggested TDF-DRN outperforms other approaches with greater accuracy. The accuracy of the proposed system is shown in Table 1.

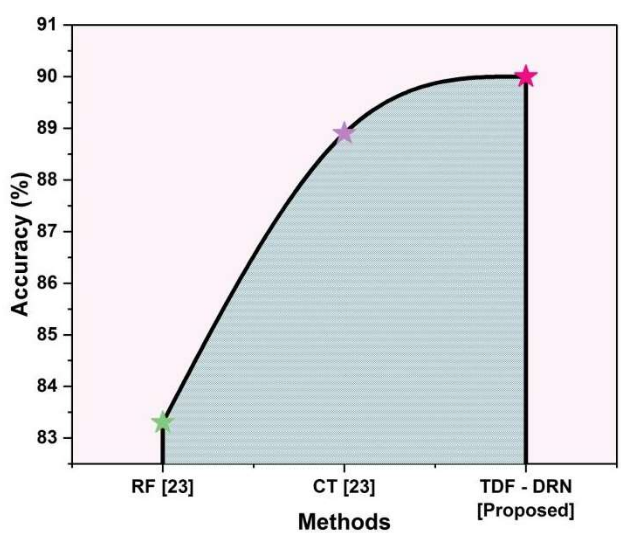


Figure 5: Output for Accuracy

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Table 1: Accuracy comparison

Methods	Accuracy (%)
R.F. [23]	83.3
CT [23]	88.9
TDF - DRN [Proposed]	90

Precision

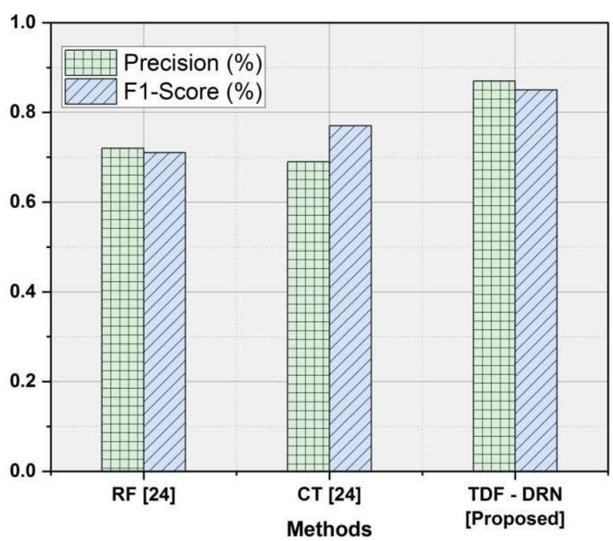
Precision, or the proportion of truthfully correct positive estimates is calculated by dividing the amount of tp forecasts delivered by the simulation using the overall amount of positive forecasts estimated by the algorithm as shown in Equation (11).

$$Precision = \frac{TP}{TP+} \tag{11}$$

The total accuracy that takes into consideration both features of the model is shown by the F1-score as shown in Equation (12), resulting in representing the harmonics means of the simulation's accuracy and recall.

$$F1 - score \ = \frac{2 \times precision \times recall}{precision + recal} \tag{12}$$

The table illustrates the performance metrics for a variety of classifiers, with the proposed TDF-DRN beating the others in terms of precision and F1-Score, obtaining percentages of 92.13%, 93.42%, 90.14%, and 95.95%, respectively. The table provides the performance metrics for the other classifiers, as shown in Table 2 and Figure 6.



Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Figure 6: Output for Precision and F1-score

Table 2: Precision and f1-score comparison with existing method

Methods	Precision (%)	F1-Score (%)
RF [24]	0.72	0.71
CT [24]	0.69	0.77
TDF - DRN [Proposed]	0.87	0.85

Sensitivity

The percentage of actual positive instances that are accurately classified as positive is known as Sensitivity. It suggests that a different percentage of real positive cases are mistakenly forecasted as negative. The definition of Sensitivity is in Equation (13):

$$Sensitivity \ = \frac{T_P}{T_P+F_N} 7 \tag{13}$$

Specificity

The percentage of actual negatives that were expected to be negatives is known as specificity. It suggests that there would be an additional percentage of accurate negative data that were misinterpreted as positive and referred to as false positives. It has the following definition in Equation (14):

$$Specificity = \frac{T_N}{T_N+F_P} \tag{14}$$

Sensitivity and specificity provide a thorough assessment of a model's performance by taking its accuracy in identifying both positive and negative examples, thereby offering valuable data about its overall categorization abilities. Figure 7 depicts the values that correspond to the Sensitivity and specificity measures. The suggested TDF-DRN outperforms existing approaches with greater output. The Sensitivity and specificity of the proposed system are shown in Table 3.

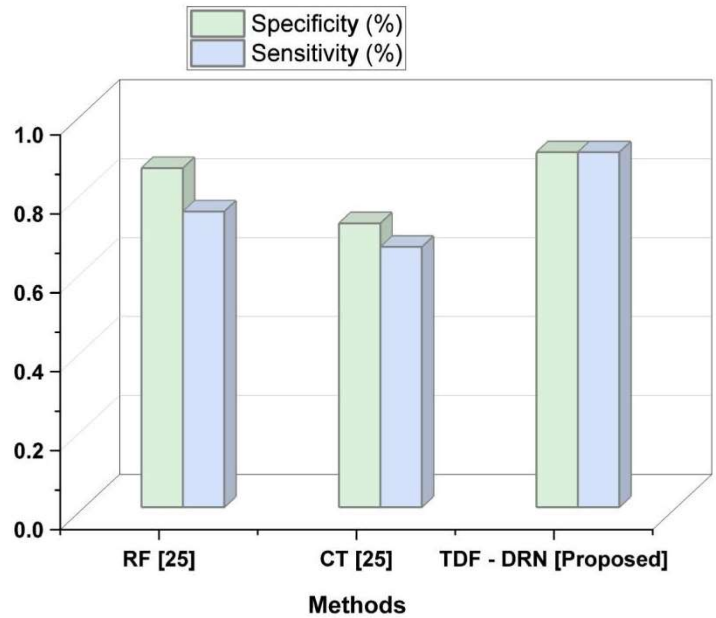


Figure 7: Output for Sensitivity and specificity

Table 3: Sensitivity and specificity comparison between existing methods

Methods	Specificity (%)	Sensitivity (%)
RF [25]	0.86	0.75
CT [25]	0.72	0.66
TDF - DRN [Proposed]	0.90	0.90

AUC Curve

An indicator used to assess a binary classification model's performance is the AUC. The AUC curve, often referred to as the Receiver Operating Characteristic (ROC) curve, visually represents the percentage of true positives as well as the rate of false positives at various levels. Shown in Figure 8.

True Positive Rate: The proportion of actual positive instances that the simulation accurately identifies in Equation (15).

$$TP\ Rate = \frac{T_P}{T_P + F_N} \tag{15}$$

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

The proportion of actual negative instances that the model incorrectly labels as positive is known as the false positive rate as shown in Equation (16).

$$FP\ Rate = \frac{F_P}{F_P + T_N} \quad (16)$$

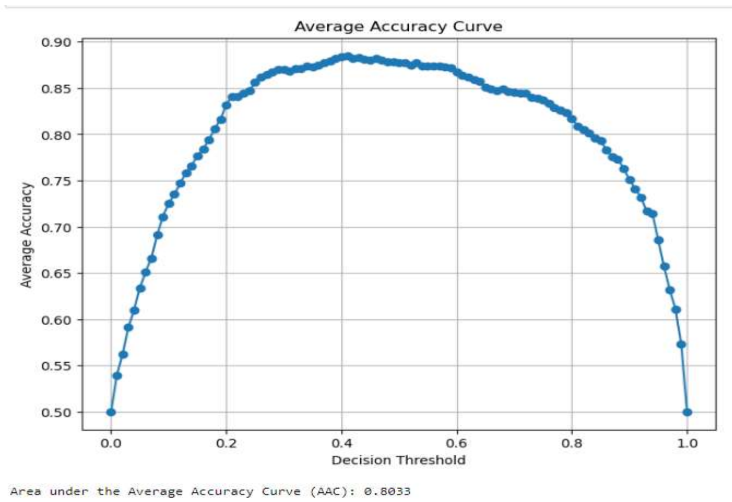


Figure 8: Output for AUC curve

The actual positive rate versus the false positive rate at different threshold values is plotted to form the ROC curve; we compute the Area under this ROC curve or AUC.

Mean squared Error (MSE)

The MSE provides a statistic that sums together the ratio of the squared variance across the actual results and projected values, a popular statistic for assessing a regression algorithm's effectiveness. It provides a gauge of the actual values that match the model's predictions. The MSE efficiently computes the mean square variance between the actual and expected values. Considering a reduced MSE suggests predictions made by the model have been correlated compared to the actual information, it is indicative of higher accuracy. It's essential that the MSE is greatly affected by the presence of significant errors in a small number of data points, as the MSE is vulnerable to outliers in Equation (17).

$$MSE = \frac{1}{m} j = 1m\left(X_j - \hat{X}_j\right)^2 \quad (17)$$

M - quantifies the amount of information, X_j - is the value that has been observed for the j-th data point, $\hat{X}_j$ - represents the anticipated value for the j-th component of data.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Root Mean Squared Error (RMSE)

As an additional statistic, RMSE is utilized to assess the efficacy of a regression model. It takes the square root to make the metric more understandable and is derived from the MSE. It measures the average magnitude of the errors between expected and actual values in Equation (18).

$$RMSE = \sqrt{\frac{1}{m} j = 1m\left(X_j - \hat{X}_j\right)^2} \quad (18)$$

m - Quantifies the amount of information, X_j - is the value that has been observed for the i-th data point, $\hat{X}_j$ - represents the anticipated value for the ith component of data

To find the most effective model for making accurate predictions, compare their RMSE values or use it as a benchmark. It is critical, though the details of the issue and the data.

MSE and RMSE, two very comparable metrics, shift on considerations like punishing more significant mistakes and the statistic can be understood in the original data units. Both have their place in practice and picking one over the other usually comes down to the needs of the current issue. Figure 9 and Table 4 show the output of the existing method with the TDF-DRN method.

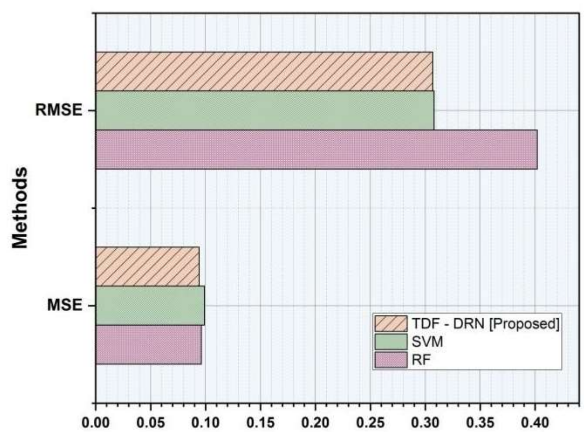


Figure 9: Output for MSE and RMSE

Table 4: MSE and RMSE comparison between existing methods

Methods	RF	SVM	TDF - DRN [Proposed]
MSE	0.0963	0.099	0.0942
RMSE	0.402	0.308	0.3069

4.1. Discussion

Strong machine learning methods called Random Forests (RF) and classification trees (CT)[23, 24, 25] are used to predict the effectiveness of medications for treating ADHD. However, they have drawbacks, such as overfitting, which can result in highly variable forecasts and decreased generalization. Classification trees are prone to overfitting, which impairs predictive accuracy and leads to poor conception by collecting disturbance in training data instead of underlying patterns. Support vector machines (SVMs) [25] are employed in investigations on the effectiveness of ADHD medications; nevertheless, they have drawbacks, including interpretability problems, scalability problems, and Sensitivity to kernel function and parameter selection. Large datasets, heterogeneous data sources, and a variety of patient groups present challenges for them. Since it can learn hierarchical representations of characteristics, capture dynamic trends in patient data across time, and simulate temporal dependencies, a time-dependent flexible deep recurrent network is a potent tool for predicting drug efficacy in ADHD treatment. In the context of ADHD, this flexibility is critical for capturing complex interactions between factors influencing treatment success.

5. Conclusion

The study was designed to discover in the event that combined drugs work well for the treatment of ADHD in both adults and children. Data preprocessing operations, such as removing duplicate data and filling in missing values through efficient data partitioning, were handled using the Map Reduce technique. A TDF-DRN technique was devised to assess the effectiveness of the medication used to treat ADHD in both adults and children. According to preliminary research, combining drugs with other therapies can be an effective way to reduce ADHD symptoms in both age groups. The study also stresses that there could represent differences in the best medication combination and dosage for adults and children. In comparison to adults, children could need less medication and advantages of different combinations. When evaluating a model's efficacy in various categorization performance elements, it is frequently taken in conjunction with Sensitivity (0.90%), specificity (0.90%), precision (0.87%), accuracy(0.90%), AUC curve(0.803%) and the F1 score (0.85%). The MSE (0.0942) and RMSE (0.3069) rates are lower than those of existing techniques. To improve the TDF-DRN's adaptability to the changing nature of ADHD symptoms, investigate integrating real-time monitoring data. Children require lower dosages than adults perform well and can gain advantages from different combinations.

Reference

1. Dobrosavljevic, M., Larsson, H. and Cortese, S., 2023. The diagnosis and treatment of attention-deficit hyperactivity disorder (ADHD) in older adults. Expert Review of Neurotherapeutics, 23(10), pp.883-893.
2. Koutsoklenis, A. and Honkasilta, J., 2023. ADHD in the DSM-5-TR: what has changed and what has not. Frontiers in psychiatry, 13, p.1064141.

3. Halperin, J.M. and Marks, D.J., 2019. Practitioner review: Assessment and treatment of preschool children with attention-deficit/hyperactivity disorder. Journal of Child Psychology and Psychiatry, 60(9), pp.930-943.
4. Heal, D.J., Gosden, J. and Smith, S.L., 2022. New drugs to treat ADHD: opportunities and challenges in research and development. New Discoveries in the Behavioral Neuroscience of Attention-Deficit Hyperactivity Disorder, pp.79-126.
5. Higgins, G.A. and Silenieks, L.B., 2022. The effects of drug treatments for ADHD in measures of cognitive performance. In New Discoveries in the Behavioral Neuroscience of Attention-Deficit Hyperactivity Disorder (pp. 321-362). Cham: Springer International Publishing.
6. Dekkers, T.J., de Water, E. and Scheres, A., 2022. Impulsive and risky decision-making in adolescents with attention-deficit/hyperactivity disorder (ADHD): The need for a developmental perspective. Current opinion in psychology, 44, pp.330-336.
7. Crepaldi, M., Colombo, V., Mottura, S., Baldassini, D., Sacco, M., Cancer, A. and Antonietti, A., 2020. Antonyms: a computer game to improve inhibitory control of impulsivity in children with attention-deficit/hyperactivity disorder (ADHD). Information, 11(4), p.230.
8. Sands, T.T., Rahdari, S., Oldham, M.S., CaminhaNunes, E., Tilton, N. and Cilio, M.R., 2019. Long-term safety, tolerability, and efficacy of cannabidiol in children with refractory epilepsy: results from an expanded access program in the U.S. Cns Drugs, 33, pp.47-60.
9. Riesco-Matías, P., Yela-Bernabé, J.R., Crego, A. and Sánchez-Zaballos, E., 2021. What do meta-analyses have to say about the efficacy of neurofeedback applied to children with ADHD? Review of previous meta-analyses and a new meta-analysis. Journal of attention disorders, 25(4), pp.473-485.
10. Ahmadi, A., Kashefi, M., Shahrokhi, H. and Nazari, M.A., 2021. Computer-aided diagnosis system using deep convolutional neural networks for ADHD subtypes. Biomedical Signal Processing and Control, 63, p.102227.
11. Hossain, M.D., Kabir, M.A., Anwar, A. and Islam, M.Z., 2021. Detecting autism spectrum disorder using machine learning techniques: An experimental analysis on toddler, child, adolescent and adult datasets. Health Information Science and Systems, 9, pp.1-13.
12. Yeh, S.C., Lin, S.Y., Wu, E.H.K., Zhang, K.F., Xiu, X., Rizzo, A. and Chung, C.R., 2020. A virtual-reality system integrated with neuro-behavior sensing for attention-deficit/hyperactivity disorder intelligent assessment. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(9), pp.1899-1907.
13. Lipka, R., Ahlers, E., Reed, T.L., Karstens, M.I., Nguyen, V., Bajbouj, M. and Kadosh, R.C., 2021. Resolving heterogeneity in transcranial electrical stimulation efficacy for attention deficit hyperactivity disorder. Experimental Neurology, 337, p.113586.

14. Altınkaynak, M., Dolu, N., Güven, A., Pektaş, F., Özmen, S., Demirci, E. and İzzetoğlu, M., 2020. Diagnosis of Attention Deficit Hyperactivity Disorder with combined time and frequency features. Biocybernetics and Biomedical Engineering, 40(3), pp.927-937.
15. Tachmazidis, I., Chen, T., Adamou, M. and Antoniou, G., 2021. A hybrid A.I. approach for supporting clinical diagnosis of attention deficit hyperactivity disorder (ADHD) in adults. Health Information Science and Systems, 9, pp.1-8.
16. Coles, E.K., Pelham, W.E., Fabiano, G.A., Gnagy, E.M., Burrows-MacLean, L., Wymbs, B.T., Chacko, A., Walker, K.S., Wymbs, F., Robb Mazzant, J. and Garefino, A., 2020. Randomized trial of a first-line behavioral intervention to reduce the need for medication in children with ADHD. Journal of Clinical Child & Adolescent Psychology, 49(5), pp.673-687.
17. Cheng, C.Y., Tseng, W.L., Chang, C.F., Chang, C.H. and Gau, S.S.F., 2020. A deep learning approach for missing data imputation of rating scales assessing attention-deficit hyperactivity disorder. Frontiers in psychiatry, 11, p.673.
18. Luo, Y., Alvarez, T.L., Halperin, J.M. and Li, X., 2020. Multimodal neuroimaging-based prediction of adult outcomes in childhood-onset ADHD using ensemble learning techniques. NeuroImage: Clinical, 26, p.102238.
19. Amado-Caballero, P., Casaseca-de-la-Higuera, P., Alberola-Lopez, S., Andres-de-Llano, J.M., Villalobos, J.A.L., Garmendia-Leiza, J.R. and Alberola-Lopez, C., 2020. Objective ADHD diagnosis using convolutional neural networks over daily-life activity records. IEEE Journal of Biomedical and Health Informatics, 24(9), pp.2690-2700.
20. Groeneveld, K.M., Mennenga, A.M., Heidelberg, R.C., Martin, R.E., Tittle, R.K., Meeuwsen, K.D., Walker, L.A. and White, E.K., 2019. Z-score neurofeedback and heart rate variability training for adults and children with symptoms of Attention-Deficit/Hyperactivity Disorder: A retrospective study. Applied psychophysiology and biofeedback, 44, pp.291-308.
21. Tajik, A., Nikfar, S., Elyasi, S., Rajabi, O. and Varmaghani, M., 2023. Cost-effectiveness and budget impact analysis of lisdexamfetamine versus methylphenidate for patients under 18 with attention-deficit/hyperactivity disorder in Iran. Child and Adolescent Psychiatry and Mental Health, 17(1), p.115.
22. Shahmoradi, L., Liraki, Z., Karami, M., Savareh, B.A. and Nosratabadi, M., 2019. Development of decision support system to predict neurofeedback response in ADHD: An artificial neural network approach. ActaInformaticaMedica, 27(3), p.186.
23. Gunther, E.C., Stone, D.J., Gerwien, R.W., Bento, P. and Heyes, M.P., 2003. Prediction of clinical drug efficacy by classification of drug-induced genomic expression profiles in vitro. Proceedings of the National Academy of Sciences, 100(16), pp.9608-9613.
24. Chen, T., Tachmazidis, I., Batsakis, S., Adamou, M., Papadakis, E. and Antoniou, G., 2023. Diagnosing attention-deficit hyperactivity disorder (ADHD) using artificial intelligence: a clinical study in the U.K. Frontiers in Psychiatry, 14, p.1164433.

25. Luo, Y., Alvarez, T.L., Halperin, J.M. and Li, X., 2020. Multimodal neuroimaging-based prediction of adult outcomes in childhood-onset ADHD using ensemble learning techniques. NeuroImage: Clinical, 26, p.102238.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930