International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

A Linear Regularized Normalized Model for Dyslexia and ADHD Prediction Using Learning Approaches

Pavan Kumar Varma Kothapalli[1*], V Rathikarani[2], Gopala Krishna Murthy Nookala[3]

[1*]Research Scholar, Department of Computer Science and Engineering, Annamalai University , Chidambaram, Tamilnadu, India.
Email:- kdvpkvarma@gmail.com
[2]Assistant Professor, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu, India.
Email:- rathikavijayakumar82@gmail.com
[3]Professor,SRKR Engineering College. *Email:-* gopinukala@gmail.com

***Corresponding Author**: Pavan Kumar Varma Kothapalli
*Research Scholar, Department of Computer Science and Engineering, Annamalai University , Chidambaram, Tamilnadu, India.
Email:- kdvpkvarma@gmail.com

Abstract-A learning disability called dyslexia typically affects school-age kids. Children have trouble spelling, reading, and writing words. Children who experience this problem often struggle with negative emotions, rage, frustration, and low self-esteem. Consequently, a dyslexia predictor system is required to assist children in overcoming the risk. There are many current ways of predicting dyslexia. However, they need to provide higher prediction accuracy. Also, this work concentrates on another disorder known as Attention-Deficit Hyperactivity Disorder (ADHD). The prediction process is more challenging as the prediction process shows some negative consequences. The data is typically gathered from online resources for the prediction process. This study examines how the predictor model predicts dyslexia and ADHD using learning strategies. Here, the most important features for accurately classifying dyslexia, non-dyslexia and ADHD are extracted using a new Support Vector Machine (SVM) for feature selection based on the $L1-$norm and $L2-$norm. Based on the weighted values, the predicted model provides improved subset features from the internet-accessible dataset. The accuracy, precision, F1-score, specificity, sensitivity, and execution time are all examined here using 10-fold cross-validation. The maximum accuracy reached with this feature subset during the prediction process is carefully reviewed. The experiment results imply that the anticipated model is used to accurately predict the defect and as a tool for CDSS. Recently, dyslexia and ADHD prediction has been greatly aided by computer-based predictor systems. The expected model also effectively fits the experimental design, bridging the gap between feature selection and classification.

Keywords- Dyslexia, low self-esteem, support vector machine, $l1$-norm and $l2$-norm, clinical decision support system

1. Introduction

Despite having a normal IQ, reading and spelling difficulties are signs of the neurocognitive learning disease dyslexia. It is one of the most prevalent learning disorder that affect 5% to 12% population [1]. Dyslexia frequently has detrimental effects on a person's ability to succeed in school or at work, also on their emotional and social development, as well as on their self-esteem. According to studies [2], the sooner dyslexia is identified and treated in the classroom, the less detrimental consequences it has.To establish a user's gaze position, ‘eye-tracking’ analyses pupil dilation and corneal reflection on the eye caused by infrared footage,for example, on a computer screen. This is a major innovation in educational technologies [3]. Fixations, which are stay-put moments of gaze that last 200 to 300 milliseconds, and saccades, which are quick ballistic movements that last 15–80 milliseconds, are the primary eye movements that take place during reading.

Readers with dyslexia display eye movement patterns considerably different from typical a reader [4], which includes noticeably longer fixations, shorter saccade durations and more saccades directed backwards. According to certain theories, the individual's difficulty decoding and understanding printed words are the root cause of the atypical eye movements [5]. Eye tracking has long held the promise of being used with computational techniques to provide detailed information on a person's cognitive processes. Developing technologies to detect reading problems from eye movements accurately is an essential first step in fulfilling this promise. Author in [6] has successfully used machine learning approaches to detect dyslexia using eye movements.

ADHD is a mental disorder seen in children worldwide, with a 7% prevalence. It remains as a disorder in life and seems as a risk factor for mental health issues like self-harm, emotional problems, anti-social behaviors, disruption and defiance [7]. The foremost ADHD symptoms encountered in childhood are impulsivity, hyperactivity and inappropriate attention, which greatly impact performance and behaviour at home and school. ADHD prediction is generally based on the comprehensive evaluation made by the psychiatrist, paediatrician or psychologist; however, clinical manifestations are not informal to predict [8] – [10]. Some methodologies like Magnetic Resonance Imaging

(MRI), Positron Emission Tomography (PET) and Computed Tomography (CT) are not yet competent for the reasonable prediction of ADHD in developing countries where the disorder prediction is controversial among researchers. The advancements in learning approaches help to offer better decisions to be taken during times of complexity. Here, an expert system is designed to enhance the prediction accuracy with L1- and L2-norm-based regularization SVM with diverse functions. The model has the competency to predict the irrelevant features and diminishes the corresponding coefficients to zero. With various hyper-parameter values, various features are chosen. The optimal value has to be searched to offer an optimal feature subset.

This work remainder is set up as follows: A review of pertinent literature is provided in Section 2. While Sections 3 and 4 provide more specifics on the experimental findings and comments, Section 5 presents the paper's conclusion.

2. Related works

Dyslexia affects 5%–12% of children; non-linear classifiers are widely used in recent studies to detect dyslexia and are trained with almost similar proportions. The performance of learning algorithms is significantly impacted by such a straightforward solution to the class-imbalance issue [11], which prevents the trained models from being re-analyzed to the level of the population. For instance in a 97-volunteer controlled trial (10-54) that was conducted [12], 48 were predicted to have dyslexia. Each participant read 12 Spanish passages totalling 60 words as part of the exam.The authors successfully identified dyslexia withA Support Vector Machine (SVM) model with the reading duration, mean fixation time, and participant age as variables produced an accuracy of 80.2% after 10-fold cross-validation. In a different study 185 Swedish youngsters were studied (ages 9–10), 97 of whom had common word decoding skills (5th percentile), and 88 were average readers (word decoding performance that is average or better) [13]. It was unsurprising that the SVM classifier achieved, given the substantial variance in reading proficiency among groups and the virtually same group sizes, 10-fold CV accuracy score of 95.6% was obtained. Various eye movement traits related to ante-grade and retrograde saccade lengths and fixation intervals were advantageous in the experiment. The dataset was analyzed in [14] utilizing hybrid SVM-Kernel technique based on Particle Swarm Optimization (PSO), resulting in 96% accuracy. The properties of eye movement were converted into essential components for this investigation [15].

Then, In a recent clinical diagnosis 32 (46%) of the 69 Greek youngsters (ages 8 to 12) studied in [16] had dyslexia. The youngsters in the control group also had their lack of reading difficulties which were clinically confirmed. Therefore there was a sign of a reading competence gap between the groups. Eye movements were recorded while two texts totalling 324 words were being read. Using SVM and LASSO, 97% accuracy was attained with saccade length, frequency of fast forward movements, and several saccades method of repeatedly focused phrases. Using Reading-eye movements($n = 61$), [17] an ambitious study in higher education was presented to distinguish between highly competent university students and underqualified pupils. Instead of using general eye movement metrics like mean fixation duration, the identification was predicated on subtle eye movement patterns connected to reading processes [18]. Reading time along the regression path, forward fixation, second-pass fixation and first-pass rereading time related to a sentence or paragraph, and these features all went towards making the SVM approach's classification accuracy of 80.3% possible [19].

Recent publications include two investigations that used significant test samples of juvenile dyslexics. The author in [20] found that the focal length had the strongest association with reading speed and accuracy among 2679 youngsters (ages 7-9). For 3644 respondents (aged 7–17) in a gamified online test with 32 Spanish language exercises, reading tasks were processed in [21] using a method that eliminated the need for eye movements. A diagnosis of dyslexia was given to 392 participants (10.8%) or people. The Random Forest (RF) classifier achieved 10-fold cross-validation scores of 79.7/79.1% precision and 80.4/78.4% recall in all 196 features; 4 representing 192 performance characteristics from the interaction during playing were used to separate people with dyslexia from those without it. The RF model's study revealed that gender and overall performance in Spanish lessons were the two most significant features. SVM is the most widely used technique. According to a recent review of techniques and uses of eye tracking, convolutional neural networks [22] and deep learning techniques [23] are becoming increasingly popular [24]. However, there are only three initial investigations utilizing the Random Forest method. The earlier investigation employed standard statistical approaches like discriminant analysis with an extremely small sample size. For investigation, discriminate that analysis performed by [25] includes 10 patients in three diverse diagnostic groups: control, ADHD, and developmental dyslexia. The author shows a discriminate analysis classifier that evaluated ADHD of 11 boys. and higher predictive accuracy was stated in these evaluations (85% to 89%). It is complex to compute how these models can generalize the provided samples and lack sample replication [26] – [30]. However, these methods show some drawbacks in terms of prediction. The proposed model aids in resolution of the above said issues.

3. Methodology

A dyslexic (Kaggle dataset for dyslexia prediction) and ADHD (https://fcon_1000.projects.nitrc.org/indi/adhd200/) predictive model is analyzed in this research.The suggested system design is presented in Fig 1. Preprocessing is performed over the data to remove null values with zeros. Blanks and null values are ignored when extracting and classifying features. After processing, feature set and sub-set are derived. Using statistical methods, progressive and regressive saccades were recovered from different events, including fixation length and the raw data.

A subset of the characteristics in the dataset is chosen as principal components using the algorithm for choosing features in Principal Component Analysis (PCA). When the analysis was performed, the model considered five main factors. Data is split into training and testing samples in 80% and 20%. The PSO technique is employed to adjust the feature weights and to create an optimized kernel that combines quadratic and linear kernels. To construct a dyslexia classification model, a better kernel is used. The procedure is iterated using various subsets for training and testing. With three iterations using different random divisions of the complete dataset, the technique is designed to reduce the variance calculated by cross-validation (CV). Averaging the CV performances across three iterations yields the final evaluation. The final assessment shows the suggested model's projected accuracy.

3.1. Pre-processing $Lx, Ly\ stands\ f\ Rx, Ry, and\ t$

The following reading can be found in the eye tracker's raw data: $t, Lx, Ly, Rx, and\ Ry$. Here, the coordinates of the eyeballs on the computer screen are x and y.Here, *Lx, Ly* stands for the coordinates of the left eye. *Rx, Ry* stands for the coordinates of the right eye and for viewing time in seconds, the right eye, the left eye, and the left eye. When participants blink, the data could contain null values. When blinking or when the eye is not focused on the computer screen, the reading positions of the eye are noted as blank. For feature extraction and categorization, blinks must be eliminated or ignored. Hence the default value is used in place of these values. During this test, zero is used to replace any missing values.

3.2. Feature extraction

From the bare eye, a substantial number of objective and systematic features were extracted from the tracking data after analysis to retain the entire ocular signal. Using statistical methods, different events, including fixation length and progressive and regressive saccades, from the raw data. Fixation, saccades, transients, and distortions are the four states of identification. A fixation state is defined as when a user maintains a fixed gaze for at least 50 milliseconds. When reading, which is when they move quickly the eyes are in a saccade condition,. Saccades that progress and retreat have been noticed. When the eye's horizontal and vertical locations are below 0.5 degrees, it is in a transient state. When the subject blinks, it is said to be distorted because both the horizontal and vertical signals are discarded. Both fixation and saccades have a specific set of qualities. The following variables are measured. (1) Fixation and saccade duration. (2) The eye's specific location throughout the test. (3) The average eye position's standard deviation. (4) The separation between the two positions (5) of the eyes is determined by the average of the two eye locations. Both the horizontal and vertical measurements of the parameters above are required. Each parameter's mean and standard deviation were calculated, yielding 75 characteristics. These characteristics completely caught the eye movement signal quantitative characteristics.

To extract features, the PCA method is utilized.It examines the variance and correlationwith orthogonal linear transformations. The transformation of the valuesis typically from strongly correlated into uncorrelated values known as principal components. It aids in highlighting the data's key characteristics, making it simple to analyze and visualize. Based on the variance, the provided data is remapped to new coordinates. Using R programming, the PCA technique is implemented in this work. In the outcome, there are five dimensions or primary components. PCA helps to reduce background noise and provides a clear understanding of the ocular aspects by removing duplicated and irrelevant features.It is performed ten times to prevent biasing or too optimistic selection. It aids in identifying the most crucial factors when determining the major components since it demonstrates the structural connection between the components and the variables.

The $cos2$ value indicates the precision of the variable representation while the elevated $cos2$ indicates effective variable representation. The variable factor map projects the 75 characteristics and their correlations. The accuracy of a variable's representation depends on how close it is to the circle. The initial components provide less weight to variables near the plot's centre.The percentage of variation that the primary components, when written as Eigenvalues, preserve is the variable contribution in calculating a certain principle component. The first principle component's Eigenvalues are higher and gradually get less as we move on to the subsequent principal components. The connected line segments display each primary component's Eigenvalues. The first two primary components have been found to preserve 60% of the variations found in the data.

3.3. Classifier

Fig 1 illustrates the two successive steps of proposed diagnostic system. A linear and L_1regularized SVM is usedanL_2 regularised in the initial stage SVM is used in the second stage together with several kernels like RBF and linear. Initial model eliminates extraneous features by provisioning coefficients to zero. Several subsets of features are produced by deleting certain traits, depending on the value of its hyper-parameter, C_1. We manually adjust C_1 with different values to identify the collection of discrete values from which different feature subsets could be derived. We declare these discrete values as the hyper-parameter space for C_1upon looking through these distinct C_1 values. Then, to obtain the optimal subset of characteristics, the best value of C_1 must be found among the defined finite discrete values of C_1. The successive SVM model which serves as a prediction model, receives best traits. The successive model's kernelC_2, and gamma G hyper-parameters must also be optimized. It is crucial to talk about the two models' formulations to comprehend how L_1 and L_2regularization affects SVM performance and how

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

they approach the issue of classifying features. The following is how L_2regularized SVM model is expressed:

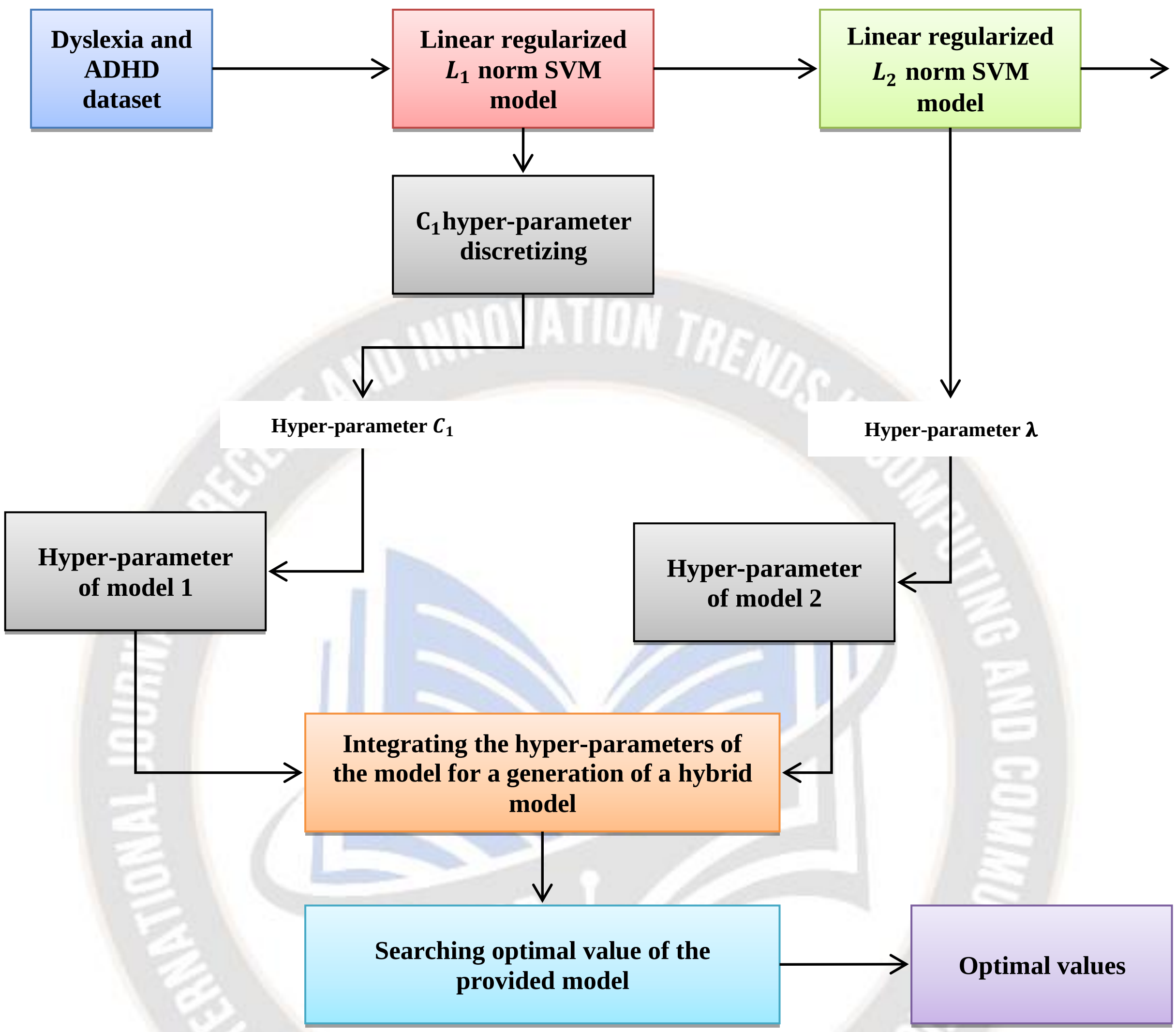

Fig 1 Block diagram of the proposed model

1) L_2SVM

SVMisadoptedin various categorization issues, including bioinformatics. To optimize, the model attempts to find an ideal hyper-plane given the distance for any class about the nearby training data points. Due to their high ability to generalize to new, unknown data items, lack of local minima, non-linear decision boundary, and reliance a limited set of hyper-parameters, SVM models are frequently utilized in classification applications. Here, $S = \{(x_i, y_i)|x_i \in R^P, y_i \in \{-1, 1\}_{i=1}^{k}$, where x_i is the i^{th} instance and P is the dimension of every instance/feature vector, which is the definition of S. $y_i$also stands for the class designation. The class label for dyslexia and ADHD disease classification problems might either be -1 or 1. The hyper-plane denoted by $f(x) = w^T * x + b$ where b is the bias and w is the SVM model learns the weight vector. The SVM model's hyper-plane maximizes the margin and reduces classification error. The distance to one of the closest positive and negative examples is added to determine the margin. The $\frac{1}{||w||_2^2}$hyper-plane thus maximizes the margin distance.The SVM model includes slack variable set,$\xi_i = 1, \ldots, k$ and penalty parameter, C to balance the minimization of $||w||_2^2$ with the misclassification error minimization. The following statement makes this clear:

$$\min_{w,b,\xi} \frac{1}{2} ||w||_2^2 + C \sum_{i=1}^{k} \xi_i \qquad (1)$$

$$\begin{cases} y_i(wx_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1,2,\dots,k \end{cases}$$

The regularizer term is L_2-norm, and the ξ slack variable gauges the level of misclassification.

2) L_1SVM

The complexity of the model is reduced by substituting the L_1norm for the L_2norm as the regularizer or penalty function. Because the L_1-norm SVM can automatically suppress irrelevant or noisy data, it can be utilized for feature selection. The vector $w's$ elements that match the traits that will be removed are shrunk as in Eq. (2):

$$\min_{w,b,\xi} \frac{1}{2}||w||_1 + C\sum_{i=1}^{k}\xi_i \quad (2)$$

$$\begin{cases} y_i(wx_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1,2,\dots,k \end{cases}$$

Some of the fitted coefficients or w in 2 components will be sparse solutions or exactly zero for sufficiently lesser C. The L_1 −regularized linear SVM model can choose features thanks to this characteristic. Additionally, different fitted coefficients will become zero if the value of the C hyper-parameter is changed. As a result, several feature subsets will be obtained. Therefore, we must look for hyper-parameter $C's$ ideal value, which will produce the ideal subset of features.

It is clear from the discussion above that we are working with two models stacked on top of one another. Both of the models have their hyper-parameters, as already mentioned. In this study, The hyper-parameter of the first model and the second model, i.e. L_1regularizedand L_2regularized SVM serves as a prediction modeland represented by C_2. In contrast, the linear SVM, which serves as a model for feature selection, is marked by C_1.Another hyper-parameter known as the kernel is also present in the second model. If a linear kernel is used, the sole hyper-parameter for the second model will be C_2. On the other hand, if an RBF kernel is used, a second SVM model will contain an additional hyper-parameter named G. In any situation, both models' hyper-parameters need to be tuned. As a result, we have two optimization problems to solve: one involves optimizing the hyper-parameters of the first model, and the other involves optimizing the second model's hyper-parameters. While the second model's optimization will produce an optimized predictive model, $C_1's$ optimization will provide an optimal subset of characteristics.

In this work, we create a hybrid grid by combining the hyper-parameters from the two models. In other words, the initial hyper-parameter for the first model will be each point's initial coordinate on the hybrid grid.C1.whereas the hyper-parameters of the second and third models, $C2\ and\ G$, will be the second and third coordinates, respectively. As a result, the hybrid grid's points can all be represented as $(C1, C2, G)$. The optimized iterations of the two models will be produced at the ideal point of the hybrid grid.In other words, the hybrid grid's ideal location corresponds to the ideal feature set and the ideal predictive model, which will perform well when applied to the ideal feature set.

4. Numerical results

Here, numerous experiments are doneto assess the model efficiency. The regularized linear SVM model L_1 and L_2 are stacked in the first experiment. AnL_1regularized linear SVM model is stacked on top of an RBF kernel-equipped L_2regularized SVM model in second experiment. The third experiment is doneto compute how the proposed model performs when compared to other machine learning models. The CPU is an Intel (R) Core, and Windows7 64-bit operating system (TM) i3-2330M is clocked at 2.20GHz.Here, the experiments are simulated using the Python programming language and toolkit.

The suggested linear SVM model L_1 and the regularized linear SVM model L_2were used to train the classification model. Traditional SVM typically achieves 90% accuracy, but the suggested model achieves 95% accuracy. The findings demonstrate that the model developed in this work outperformed the conventional Linear SVM in terms of F1, recall, sensitivity, accuracy, precision, and other metrics. The suggested model's accuracy findings are comparable. Table 1 and Table 2compare the results of the two classification models. In this experimental setup, the performance indicators are calculated as follows:

$$Accuracy = \frac{\sum Correctly\ classified\ instances}{\sum all\ instances} \quad (3)$$

$$Specificity = \frac{TN}{FP + TN} \quad (4)$$

$$Sensitivity\ (recall) = \frac{TP}{TP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$False\ positive\ rate = \frac{FP}{FP + TN} \tag{7}$$

$$FNR = 1 - sensitivity \tag{8}$$

$$F1 = 2 * \frac{precision * recall}{precision\ + recall} \tag{9}$$

Fig 2 to Fig 4 shows a visualization of the testing data's expected and actual class values. When employing Linear SVM, the observed class is represented in $the\ x$ −axis, and the y-axis represents the expected class. People with dyslexia fall into Class 1, and non-dyslexia into Class 0, respectively. Fig 5 to Fig 7 shows a visualization of the testing data's expected and actual class values. When utilizing Hybrid Kernel SVM-PSO, the projected class is represented by the y −axis, while the X-axis represents the observed class. Dyslexiafalls into Class 1 and non-dyslexia into Class 0, respectively.

Table 1 Evaluation metrics with regularized L_1 −norm for dyslexia

$k - fold$	**Accuracy**	**TPR**	**TNR**	**FPR**	**FOR**	**F1**
1	95	90	100	0	0.1	95
2	95	100	90	0.1	0	95
3	83	100	70	0.3	0	83
4	89	100	81	0.18	0	87
5	89	72	100	0	0.30	88
6	95	84	100	0.1	0.15	94
7	100	100	100	0	0	100
8	89	100	80	0.2	0	88
9	95	100	87	0.15	0	95
10	83	90	72	0.30	0.1	82

Table 2 Evaluation metrics with regularized L_2 −norm for dyslexia

$k - fold$	**Accuracy**	**TPR**	**TNR**	**FPR**	**FOR**	**F1**
1	95	100	89	0.1	0	95
2	100	100	100	0	0	100
3	95	89	100	0	0.125	95
4	93	100	91	0.09	0	94
5	89	100	72	0.28	0	88
6	89	86	90	0.1	0.14	88
7	100	100	100	0	0	100
8	95	100	88	0.125	0	95
9	100	100	100	0	0	100
10	95	100	89	0.1	0	95

Table 3 Comparative analysis based on L_1 −norm and L_2 −norm for dyslexia

Metrics	L_1 **−norm**	L_2 **−norm**
Accuracy	91	96
Sensitivity	95	100
Specificity	71	90
PPV	77	90
NPV	93	100

Table 4 Evaluation metrics with regularized L_1 −norm for ADHD

$k - fold$	**Accuracy**	**TPR**	**TNR**	**FPR**	**FOR**	**F1**
1	95	90	100	0	0.1	95
2	95	100	90	0.1	0	95
3	83	100	70	0.3	0	83
4	89	100	81	0.18	0	87
5	89	72	100	0	0.30	88
6	95	84	100	0.1	0.15	94

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

7	100	100	100	0	0	100
8	89	100	80	0.2	0	88
9	95	100	87	0.15	0	95
10	83	90	72	0.30	0.1	82

Table 5 Evaluation metrics with regularized L_2 −norm for ADHD

$k-fold$	Accuracy	TPR	TNR	FPR	FOR	F1
1	96	100	92	0.2	0	97
2	100	100	100	0	0	100
3	96	91	100	0	0.150	97
4	95	100	93	0.10	0	97
5	91	100	75	0.30	0	90
6	91	91	93	0.2	0.15	90
7	100	100	100	0	0	100
8	96	100	90	0.150	0	97
9	100	100	100	0	0	100
10	96	100	92	0.2	0	97

Table 6 Comparative analysis based on L_1 −norm and L_2 −norm for ADHD

Metrics	L_1 −norm	L_2 −norm
Accuracy	92	97
Sensitivity	96	100
Specificity	73	93
PPV	79	94
NPV	95	100

We used the technique described in 10-fold cross-validation for 100 cycles using several classifiers and feature sets. Table 3 to Table 5 provides a summary of the top outcomes. Learning approach is employed to create the column. Class weights for the Scikit-learn library SVM were modified in an inverse connection to the class frequencies, as indicated by the "Bal" tag. The feature sets whose names are listed in the "Feat" column. The typical proportion of accurate predictions for each 100 epochs that the algorithm generated is shown in the "Accuracy" column. The accuracy's standard deviation ratings are indicated as the error. Similarly, the "Memory" column includes average and standard deviation recall scores. The $RFFn$ feature sets produced the best SVM outcomes. Table 6 presents these outcomes. The two last columns include a list of the ideal SVM hyper-parameter settings. Class weight balancing significantly increased the recall score for the $RFF35$ model with only a minor accuracy loss. The SVM results utilizing the remaining feature sets are shown. TR feature set had the finest accuracy and recall ratings. The top results generated by the classifiers areshown. The hyper-parameters utilized by each model and optimized with grid search are listed in the two last columns. The outcomes fall short of those obtained using SVM, as shown. However, we used SVM to discover the best predictions by employing RF as a feature selection technique.

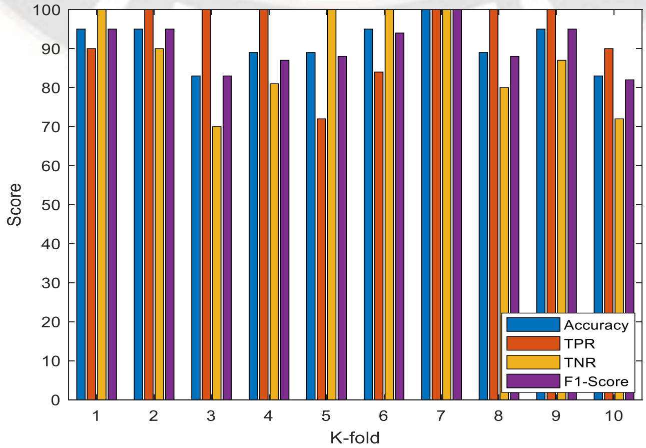

Fig 2 Evaluation metrics with regularized L_1 −norm for dyslexia

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 10 September 2023 **Revised:** 20 October 2023 **Accepted:** 30 October 2023*

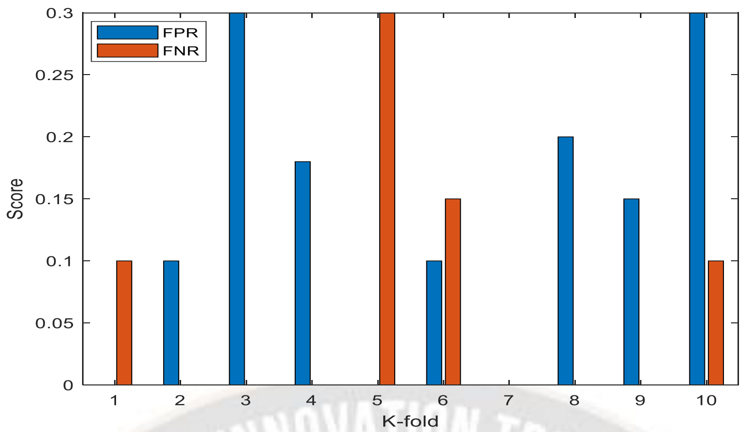

Fig 3 FPR and FNR evaluation with regularized L_1 −norm for dyslexia

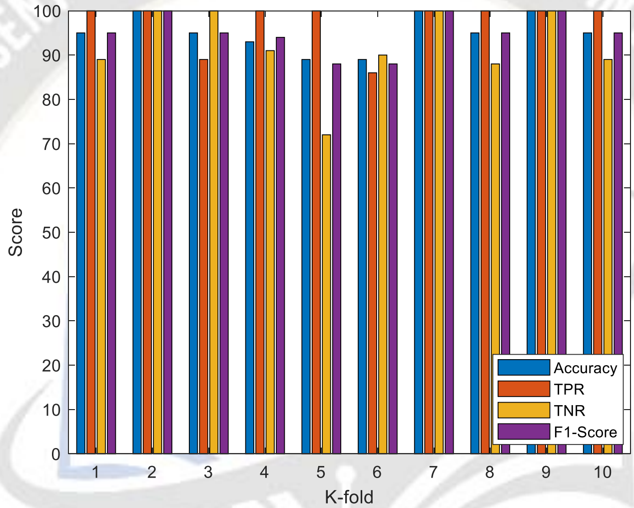

Fig 4 Evaluation metrics with regularized L_2 −norm for dyslexia

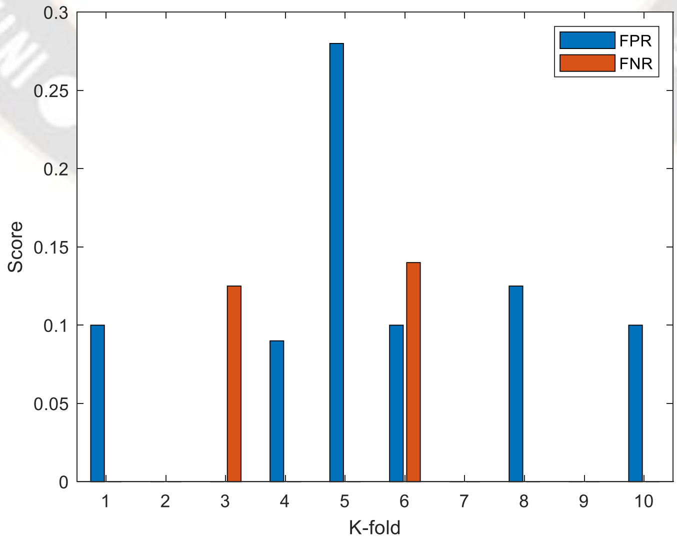

Fig 4 FPR and FNR evaluation with regularized L_2 −norm for dyslexia

Article Received: *10 September 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *30 October 2023*

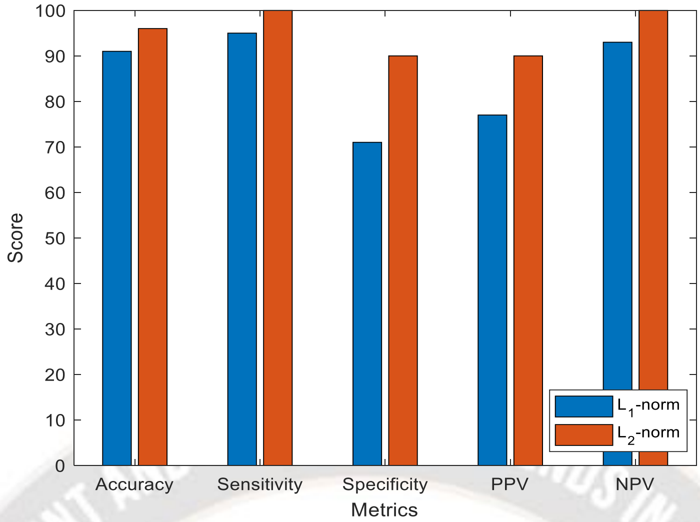

Fig 5Comparative analysis based on L_1 −norm and L_2 −norm for dyslexia

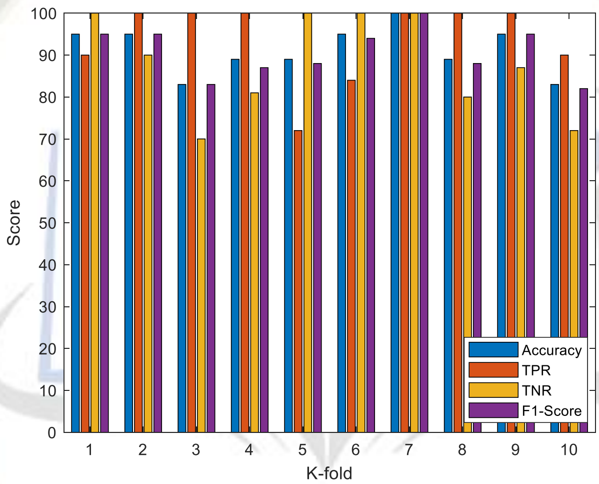

Fig 6 Evaluation metrics with regularized L_1 −norm for ADHD

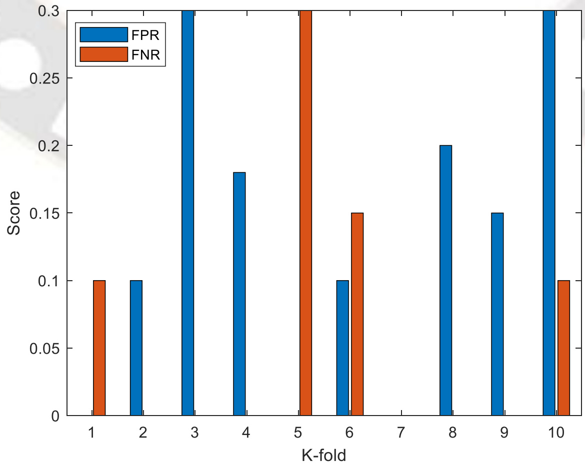

Fig 7 FPR and FNR evaluation with regularized L_1 −norm for ADHD

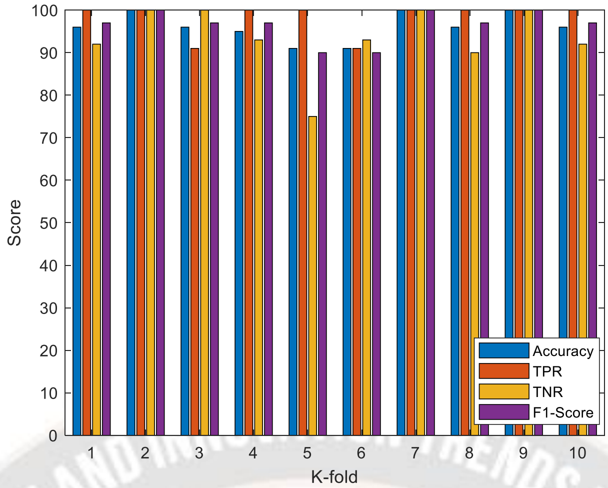

Fig 8 Evaluation metrics with regularized L_2 −norm for ADHD

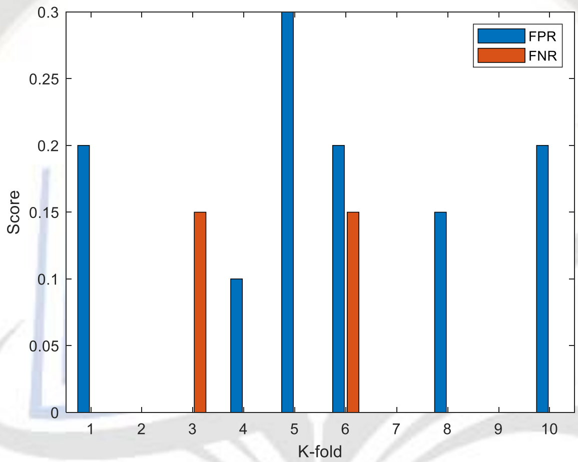

Fig 9 FPR and FNR evaluation with regularized L_2 −norm for ADHD

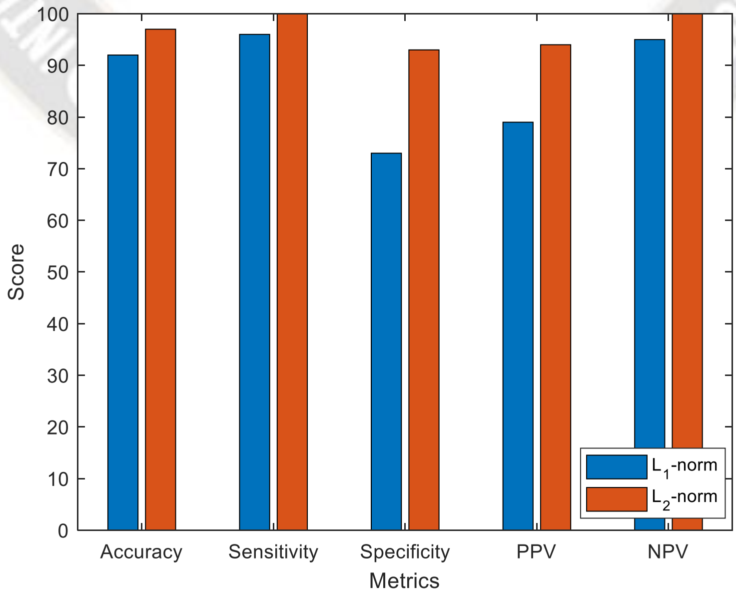

Fig 10 Comparative analysis based on L_1 −norm and L_2 −norm for ADHD

4.1. Discussion

We can conclude by examining the $RFFn$ sets' most significant aspects which may be useful when creating new dyslexia and ADHD screening tests. The number of features picked for top 10 features in every fold associated with each sentence is shown in Fig 8 to Fig 10. The total number of characteristics selected is $10 \times folds \times cycles$ which equals $10\times 5 \times 100 = 5000$. 2900 (or 58%) had something to do with the initial phrase on the task-related query page. When the context has not yet been provided, readers must largely rely on their syntactic language abilities, word identification, and decoding abilities in the first sentence. It is well-recognized that dyslexia severely impairs these skills. To put it another way, the discovery suggests that dyslexic readers use context as a compensatory reading method substantially more frequently.

Another key in feature value appeared to be the several trials $T2 - T11$. The number of trial-specific features selected displays the top 10 spots in each fold rotation. The first real trial occurs most frequently (32%) among the $T2$ data components demonstrating a high relevance in correctly identifying the two classes. The feature count fluctuates significantly within the same range for the remaining trials, with low spots in $T6\ and\ T9$. The lack of context and a cognitive framework for the information-seeking task among the participants is thought to be the cause of $T2's$ high importance. Once more, fluent readers are less affected by dyslexics' lack of text context than they are.It may help to describe the significance of $T2$ trial features during two class's separation because participants needed to see enough trials to understand the format/context of the content. As the experiment progresses, the background is formed, making it more difficult to discriminate between difficult and average readers.

The frequency with which the most crucial traits were chosen is shown in Fig 10. We can observe that the first sentence's properties (marked by "F") and those most frequent appearances are from $T2$, proving their relevance as previously established. In addition, we can see that examining the first bin is common when examining the feature names of saccadic features that correspond to the histogram bin numbers. The last bin is the most significant in terms of features retrieved from fixation data. These results imply that the longest fixations and the shortest saccades are most helpful for classification.We can also observe that characteristics derived from saccadic data have a higher value than those derived from fixation data. Only one of the 35 most crucial traits comes from a conventional transition matrix. Hence the usage of transition matrices did not significantly aid in categorization. The remaining three features in this list are represented by the number of fixations generated within the highlighted text ($T2F - F, T10T - T, and\ T3D - D$).

In this experiment, the most pertinent features are chosen in the first and second stage using L_1 regularized linear SVM and L_2 regularized SVM with RBF kernel. Predictive modelling is performed with RBF kernel. For $K = 8$, the greatest accuracy of 92% is achieved using only 8 features. The optimum subset of attributes is $F2, F3, F7, F8, F9, F11, F12,$ and $F13$.The ideal feature subset not only increases the predictive model's potential and decreases its temporal complexity, shortening the model's training period. Table 6 lists the findings for several feature subsets for various hyper-parameters. The last row of the Table depicts a situation when only the L_2 regularized SVM model, the second SVM model is utilized. This example specifies the standard SVM model. Therefore, it is evident from the experimental findings that the proposed strategy enhances a traditional SVM model's performance by 3.3%.

5. Conclusion

We created a classifier in this study to recognize eye movement analysis of dyslexic readers and brain disorders known as ADHD. Importantly, we used an arbitrary reading fluency score threshold to designate dyslexia in this article, which makes the categorization process intrinsically challenging. Instead of employing the simple averages of the measures of fixation and saccade, our feature extraction uses gaze patterns to supplement the transition matrices that are often utilized. Aregularized linear SVM model L_1 and the regularized linear SVM model L_2 classifier incorporating the most important aspects of eye movement, when chosen via SVM achieved an accuracy score of 97% and an accuracy of 95% in both models. The outcome is encouraging, and a more in-depth examination of the feature's importance offers knowledge that may be applied to direct future research toward efficient and accurate dyslexia screening systems.

REFERENCES

1. C. Christo, J. M. Davis, and S. E. Brock, Identifying, Assessing, and Treating Dyslexia at School. Springer Science & Business Media, 2009.
2. M. Iwabuchi, R. Hirabayashi, K. Nakamura & N. K. Dim, "Machine learning based evaluation of reading and writing difficulties," Stud. Health Technol. Inform., Vol. 242, pp. 1001–4, 2017.
3. C. Galdi, M. Nappi, D. Riccio & H. Wechsler, "Eye movement analysis for human authentication: a critical survey," Pattern Recognit. Lett., Vol. 84, pp. 272–83, 2016. doi:10.1016/j.patrec.2016.11.002
4. L. Rello and M. Ballesteros, "Detecting readers with dyslexia using machine learning with eye tracking measures" in Proceedings of the 12th Web for All Conference, ACM, pp. 16, May 2015.
5. C. H. H Yang, J. H. Huang, F. Liu, F. Y. Chiu, M. Gao, W. Lyu & J. Tegner, "A Novel Hybrid Machine learning model for Auto-classification of Retinal diseases" arXiv Preprint ArXiv:1806.06423, 2018
6. P. Charoen, W. Kongprawechnon & K. Tungpimolrut, "A hybrid particle swarm optimization-SVM classification for automatic cardiac auscultation,"

Songklanakarin Journal of Science & Technology, Vol. 39, 2017.
7. J. Bingel, M. Barrett & S. Klerk, "Predicting misreadings from gaze in children with reading difficulties" in Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications, pp. 24–34, 2018.
8. S. Kim & R. Wiseheart, "Exploring text and Icon graph interpretation in students with dyslexia: An Eye-tracking study," Dyslexia, Vol. 23, pp. 24–41, 2017. doi:10.1002/dys. 1551
9. H. Jarodzka and S. Brand-Gruwel, "Tracking the reading eye: towards a model of real-world reading," J. Comput. Assist. Learn., Vol. 33, pp. 193–201, 2017. doi:10.1111/jcal.12189
10. J. Hautala, C. Kiili,, Y. Kammerer, O. Loberg, S. Hokkanen & P.H Leppanen, "Sixth graders' evaluation strategies when reading internet search results: an eye-tracking study," Behav. Inf. Technol., Vol. 37, no. 8, pp. 761–73, 2018. doi:10.1080/0144929x.2018.1477992
11. M. N. Benfatto, G. O. Seimyr, J. Ygge, T. Parnell, A. Rydberg & C. Jacobson, "Screening for dyslexia using eye tracking during reading,". PLoS one, Vol. 11, 2016.
12. Jolliffe, T. Ian and J. Cadima, "Principal component analysis: a review and recent developments," Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, Vol. 374, pp. 2065, 2016
13. O. W. Samuel, G. M. Asogbon, A. K. Sangaiah, P. Fang, and G. Li, ''An integrated decision support system based on ANN and fuzzy_AHP for heart failure risk prediction,'' Expert Syst. Appl., vol. 68, pp. 163–172, Feb. 2017.
14. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, ''Computer aided decision making for heart disease detection using the hybrid neural network-genetic algorithm," Comput. Methods Programs Biomed., vol. 141, pp. 19–26, Apr. 2017.
15. Manogaran, R. Varatharajan, and M. K. Priyan, ''Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system,'' Multimedia Tools Appl., vol. 77, no. 4, pp. 4379–4399, 2018.
16. Chen H, Chen W, Song Y, Sun L, Li X (2019a) EEG characteristics of children with attention-deficit/hyperactivity disorder. Neuroscience 406:444–456
17. De Dea F, Ajcˇevic´ M, Stecca M, Zanus C, Carrozzi M, Cuzzocrea A et al. (2019) A Big-data-analytics framework for supporting the classification of ADHD and healthy children via principal component analysis of EEG sleep spindle power spectra. Procedia Comput Sci 159:1584–1590
18. Kaur S, Singh S, Arun P, Kaur D, Bajaj M (2020) Phase space reconstruction of EEG signals for classifying ADHD and control adults. Clin EEG Neurosci 51(2):102–113
19. Khoshnoud S, Nazari MA, Shamsi M (2018) Functional brain dynamic analysis of ADHD and control children using non-linear dynamical features of EEG signals. J Integr Neurosci 17(1):17–30
20. Jahanshahloo HR, Shamsi M, Ghasemi E, Kouhi A (2017) Automated and ERP-based diagnosis of attention-deficit hyperactivity disorder in children. J Med Signals Sens 7(1):26
21. Li W, Zhou T, Zou L, Lu J, Liu H, Wang S (2018b) Identification of attention-deficit/hyperactivity disorder in children using multiple ERP features. Curr Bioinform 13(5):501–507
22. Mohammadi MR, Khaleghi A, Nasrabadi AM, Rafieivand S, Bel M, Zarafshan H (2016) EEG classification of ADHD and normal children using non-linear features and neural network. Biomed Eng Lett 6(2):66–73
23. Mu¨ller A, Vetsch S, Pershin I, Candrian G, Baschera G-M, Kropotov J et al. (2019) EEG/ERP-based biomarker/neuroalgorithms in adults with ADHD: development, reliability, and application in clinical practice. World J Biol Psychiatry (just-accepted):1–29
24. O¨ ztoprak H, Toycan M, Alp YK, Arıkan O, Dogˇutepe E, Karakas¸ S (2017) Machine-based classification of ADHD and non-ADHD participants using time/frequency features of event-related neuroelectric activity. Clin Neurophysiol 128(12):2400–2410
25. Riaz A, Asad M, Alonso E, Slabaugh G (2018) Fusion of fMRI and non-imaging data for ADHD classification. Comput Med Imaging Graph 65:115–128
26. Sherubha, "Graph-Based Event Measurement for Analyzing Distributed Anomalies in Sensor Networks", Sådhanå(Springer), 45:212, https://doi.org/10.1007/s12046-020-01451-w
27. Sherubha, "An Efficient Network Threat Detection and Classification Method using ANP-MVPS Algorithm in Wireless Sensor Networks", International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8 Issue-11, September 2019.
28. Sherubha, "An Efficient Intrusion Detection and Authentication Mechanism for Detecting Clone Attack in Wireless Sensor Networks", Journal of Advanced Research in Dynamical and Control Systems (JARDCS), Volume 11, issue 5, Pg No. 55-68
29. Balasubramanian, S., Devarajan, H. R., Raparthi, M., Dodda, S. B., Maruthi, S., & Adnyana, I. M. D. M. (2023). Ethical Considerations in AI-assisted Decision Making for End-of-Life Care in Healthcare. PowerTech Journal, 47(4), 168. https://doi.org/10.52783/pst.168
30. Vahid A, Blaschke A, Roessner V, Stober S, Beste C (2019) Deep learning based on event-related EEG differentiates children with ADHD from healthy controls. J Clin Med 8(7):1055
31. [30] Saad JF, Kohn MR, Clarke S, Lagopoulos J, Hermens DF (2018) Is the theta/beta EEG marker for ADHD inherently flawed? J Atten Disord 22(9):815–826