This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

An Ensemble-based Machine Learning Technique for Dyslexia Detection During a Visual Continuous Performance Task

Mona Zaree [a] , *Maryam Mohebbi* [a,*], *Reza Rostami* [b]

[a] Department of Biomedical Engineering, K. N. Toosi University of Technology, Tehran, Iran.
[b] Department of Psychology, University of Tehran, Tehran, Iran.

Abstract

Background: Dyslexia is a neurological disorder which affects the learning of individuals suffering from it, especially children. It causes reading and writing difficulties, leading to anger, frustration, poor self-esteem, and other negative emotions. Early diagnosis of dyslexia can be extremely useful for dyslexic children since their learning requirements can be correctly handled.

Method: This study aimed to propose an efficient method for dyslexia detection via evoked-related potentials (ERPs) during a Visual Continuous Performance Task (VCPT). To this end, after extracting the latency and amplitude of the event-related potential components and a statistical analysis between the two groups, the number of data dimensions is reduced by the KS-test and principal component analysis (PCA). This method used the ensemble learning classification technique to classify the normal and dyslexic groups, including five classifiers, namely a Support Vector Machine (SVM), Decision Tree, Linear Discriminant Analysis (LDA), K-nearest Neighbors (KNN), and Naive Bayesian. Moreover, we have used an 8-fold cross-validation method to assess how generalized the process is and, more importantly, to have a well-performed control on overfitting.

Results: According to this approach, the overall average classification accuracy is 87.5%, and the sensitivity and specificity are 81.2% and 93.7%, respectively.

Conclusion: The suggested technique can improve children's education before they start school and have any skills in writing and reading. The most discriminative electrodes were found in the left hemisphere, which can be related to the Wernicke and Broca areas that have an essential role in the brain's reading, writing, and language-related functions.

Keywords: **Visual Continuous Performance Tasks (VCPT), Ensemble Learning, Dyslexia, Electroencephalography (EEG), Event-Related Potential (ERP).**

1. Introduction

The word " dyslexia"; is derived from the Greek root and refers to a learning disability. This word consists of the terms " dys"; which means difficult, painful, and abnormal, and "lexicos"; which means word. This condition can be defined as the inability to read, write, or spell words. It is critical to note that these children are no longer differentiated from other children physically or psychologically. Additionally, the term is often used to describe youngsters who have troubles of some magnitude regarding their performance in tasks such as reading and writing despite already having average or higher intelligence. This is considered one of the most common learning disabilities affecting children. In the United States, research conducted by the National Institutes of Health (NIH)

* m.mohebbi@kntu.ac.ir

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

has demonstrated that dyslexia affects approximately 5% to 20% of the population. Some people may have less severe cases, while others may have more serious ones [1].

People with dyslexia understand words and use them in conversation, but they are unable to recognize and comprehend written words as well as the correct meaning of those words. It is common for these individuals to have difficulty spelling words. As a result, negative emotions such as frustration, anger, anxiety, depression, and low self-esteem are produced. This disorder is associated with impairments in processing speed, short-term memory, visual-auditory perception sequence, speech-language, and motor skills [2].

Therefore, early detection of dyslexia is crucial to providing them with the support they require and fostering their development. According to well-recognized and experienced psychologists, people with dyslexia are either classified as "visually impaired" or "hearing impaired". People with visual dysfunction have difficulties grasping visual concepts, resulting in confusion as one of the side effects. It is also difficult for visual dyslexics to recognize complex shapes and to comprehend and assemble a series of visual events, which makes it difficult for such patients to identify similar letters and words. People with auditory dysfunction are unable to understand the meaning of words simply by hearing them. In this case, the patient cannot distinguish between identical letters when listening. This is because their auditory nervous system is not functioning correctly.

Dyslexic children, particularly those who are visually impaired, have difficulty orienting themselves spatially, distinguishing right from left and top from bottom, and eye-hand coordination. The diagnosis of dyslexia is typically based on oral and written assessments [3]. These results are used to make clinical judgments. The main problem with this method is that it is not objective because the clinical decision depends on the expert's analysis, so it is time-consuming and error-prone [4]. In addition, most of the benchmarks are designed for children already learning to read but are failing, causing the diagnosis to be made at an earlier age. Early diagnosis is the focus of research, and personalized intervention would have theoretical and practical repercussions [5]. There is currently a challenge involved in the development of objective diagnosis methods, and with an active research area focused on identifying biomarkers, it can be enthusiastic about being discussed in detail and elaboration that may be utilized to increase diagnosis accuracy by employing neuropsychological responses related to reading or language processing processes [6]. Also, as the biological factors, parameters, and processes of dyslexia are poorly understood, these biomarkers might reveal unknown aspects of dyslexia related to its neural basis [7]. This can contribute to learn more about the differences between dyslexic people and normal subjects. This is especially useful when making individualized intervention tasks in detail. Also, it can provide diagnostic tools for those who can read and those who are unable (early diagnosis).

Among the best-performed tests for examining executive functions in dyslexics is the Continuous Performance Test (CPT), which does not require reading and writing. Visually cued Go/NoGo tasks are defined by a two-stimulus scenario, with the first stimulus representing the cue and the second the aim. Some tasks demand the subject's active participation. Assessments of dyslexia, including sustained attention, inhibitory control, and cognitive processing stability, are important and significant to take into account regarding the testing of these basic qualities. The CPTs provide a quantitative behavioral assessment of executive function performance.

Researchers have developed various techniques for detecting and classifying dyslexia and its markers by using datasets prepared from diverse sources. These sources include web-based or mobile-based games, standardized psycho-educational tests, Magnetoencephalography (MEG) scans [8], Magnetic resonance imaging (MRI) [9], Positron emission tomography (PET) scans [10], eye movement tracking [11], images and video taken during cognitive/phonological tasks, and electroencephalogram (EEG) scan. The EEG recording method is reliable for acquiring data in many fields since it poses no radiation risks, is non-invasive, has a high temporal resolution that can be easily adapted to pre-readers, and is more cost-efficient than other neuroimaging techniques. Several studies of EEG have focused on eye close and open tasks, memory processing tasks, and event-related potential (ERP) recording [12], [13], [14].

Perera et al. used an EEG headset to gather 32 English-speaking dyslexic brain activity signal patterns. They then used a support vector machine (SVM) classifier and achieve an accuracy of 78.2% [15]. Al-Barhamtoshy et al. obtained EEG data from 80 schoolchildren in Saudi Arabia. They employed Artificial Neural Networks (ANN) for classification and achieved an accuracy of 89.6% [16]. Taskov et al. investigated differences in the topological properties of EEG resting state networks [17]. Also, the functional connectivity of reading neural networks in

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224. https://doi.org/10.1016/j.bspc.2023.105224

individuals with dyslexia was studied by T-Ahram et al. to determine differences in the topological features of the groups before and after undergoing visual training [18]. Other researchers who have utilized EEG datasets for the detection of dyslexia by the application of machine learning techniques are Karim et al. and Fried and Breznitz [19], [20]. Fraga-González et al. also used graph theory to analyze EEG functional connectivity networks in dyslexic and normal readers while performing an auditory-visual test [21]. Dushanova et al. studied the effect of training with visual tasks on functional connectivity in dyslexia [22]. Also, EEG connectivity analysis was used by Shaywitz et al. to investigate the relationship between dysfunctional connectivity and reading dysfluency in the neural network [23]. In another research, Zainuddin et al. analyzed the writing-related EEG signals of 20 dyslexics and 10 non-dyslexics and reached 83% and 89% accuracy, respectively, using KNN and Extreme Learning *Machines* (ELM) techniques [24]. Additionally, Nicolas et al. recorded the EEG during an auditory task and achieved an accuracy of 72.9% using an SVM classifier [25]. Pavlos et al., using Random Forest (RF) classification, identified dyslexia with an accuracy of 87.6%. Three different circumstances were used as tasks, including auditory recognition, visual discrimination, and visual discrimination with background music. The content in the three circumstances consisted of letters that were comparable in appearance (κ, γ, χ) or sound (f, v, θ, ð) [26]. In another study, Hanafi et al. employed long-short-term memory neural networks (LSTM) to categorize dyslexic children's EEG data. Researchers captured EEG data while the students did a reading task. The results demonstrated that the LSTM network classified the EEG signals with an accuracy of 86.2% [27].
It can be observed that the majority of the prior EEG-based investigations employed skills like reading and writing that must be taught at school. In our study, due to the importance of precise and early diagnosis of dyslexia before school, we used the visual continuous performance task (VCPT), in which reading and writing skills are not necessary for participation. For this purpose, we investigated the potential of ERP components as a diagnostic tool for classifying normal and dyslexic individuals. After extracting the Latency and amplitude of ERP components, we performed the statistical analysis and examined the differences between the two groups. Then, the distinguishing and effective features were selected with a method consisting of the Kolmogorov-Smirnov test (KS-test) and principal component analysis (PCA). Additionally, the ensemble learning algorithm was used to improve the dyslexia diagnosis accuracy. To the best of our knowledge, this is the first research that investigates the feasibility of using the characteristics of ERP components in the VCPT to detect dyslexic children. This study differs from others in the field since it has the ability to diagnose dyslexia by explainable features before the start of school and does not necessitate any particular skills (reading or writing) or knowledge.

2. Method and analysis

This study employed a method based on machine learning and the differences in EEG signals to diagnose dyslexia. This section introduces the dataset and explains how to record the EEG signal. Also, the extracted signal properties and the details of ensemble learning will be explained.

2.1 Participants

Forty-four right-handed participants took part in the study. The subjects were 22 normal subjects, of which 12 were males and 22 dyslexics (18 males). The mean age of the normal group was 94.3 ± 3.3 months and the mean age of the dyslexic group was 95.12 ± 2.8 months. The participants attended the Atieh clinical neuroscience center with the first report of low school spelling scores and disability in learning. All subjects were native Persian speakers with normal vision and hearing and did not differ in chronological age. Since other skills have not entirely fixed dyslexia, this age range was chosen. Thus, it makes sense to look into its natural causes and effects. There was no history of neurological disease or clinical evidence of hyperactivity problems among the subjects. In the Atieh Clinical Neuroscience Center, subjects were diagnosed as dyslexic using Wechsler Intelligence Scale (version 5) [28]. The families of all participants submitted written informed consent and agreed to the treatment strategy, and

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

the study was approved by Tehran Medical University's ethics committee. The participants could take part in a way that was easy for them, and all of them told the researchers about any possible harmful effects of the treatment.

2.2 EEG recording

A non-invasive EEG signal was obtained in the Atieh-clinical neuroscience center following the diagnosis of dyslexia. A Mitsar EEG recorder was used with a 19-channel EEG cap and a sampling rate of 500 Hz to record the data, which had FDA and CE approval. The EEG signals were referenced to linked ears and obtained while performing the task. The placement of the electrodes followed the international 10-20 system. Figure 1 shows the EEG channels' spatial setup [29]. Before every recording session, the operator set each channel's impedance to ensure it was the best it could be.

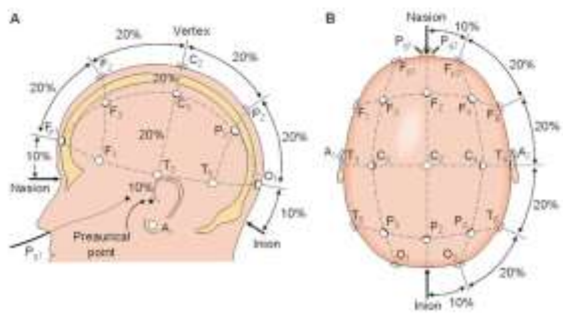


Figure 1: International 10–20 system for EEG [29].

2.1 Behavioural Task

The Visual Continuous Performance Task was performed using the standard protocol [30]. This task is an adaptation of the two-stimulus visual GO/NOGO paradigm. During the test, subjects sat in comfortable armchairs with armrests. The pictures were shown in a "pseudo-random" organization in the center of a computer display that was 1.5 m away from the eyes of the subjects. On a 17-inch monitor, the stimuli were shown utilizing the Psytask software (Mitsar Ltd.). Before every session, subjects were given a thorough explanation of the test, and 10 to 20 trials were performed. During the test, there was a 5-minute break. If needed, the subjects took a short break after every 200 trials.

20 images of animals, 20 images of plants, and 20 images of humans (presented in conjunction with an artificial "novel" sound) were chosen from three categories of visual stimuli. All visual stimuli were chosen to be the same size and brightness. Four different categories of trials were used (see Figure 2): GO trials (animal-animal), NOGO trials (animal-plant), IGNORE trials (plant-plant), and NOVEL trials (plant-human) [31]. The trials included presentations of paired stimuli with inter-stimulus intervals of 1000 msec; in addition, the duration of each trial was 3000 msec. The stimulus lasted 100 msec.

The trials were organized into four distinct blocks. Each block selected a unique combination of five animal, five plant, and five human stimuli. Every block contained 100 pairs of stimuli that were presented in a "pseudo-random" way, with the same probability for each trial group. The aim was to press a button as quickly as possible when GO trials came up (Event 1). The Animal-Plant (A-P) pairs were the NoGo conditions, which meant that the person should not answer (Event 2). Thus, following the presentation of the first Go stimulus, the participant was prepared to click the button, and the second stimulus prevented the anticipated motion. For the plant-plant (P–P) and plant-human (P–H) tests, it was supposed that the first stimulus would indicate that no action preparation was necessary and that the trial could be disregarded (Events 3 and 4).

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

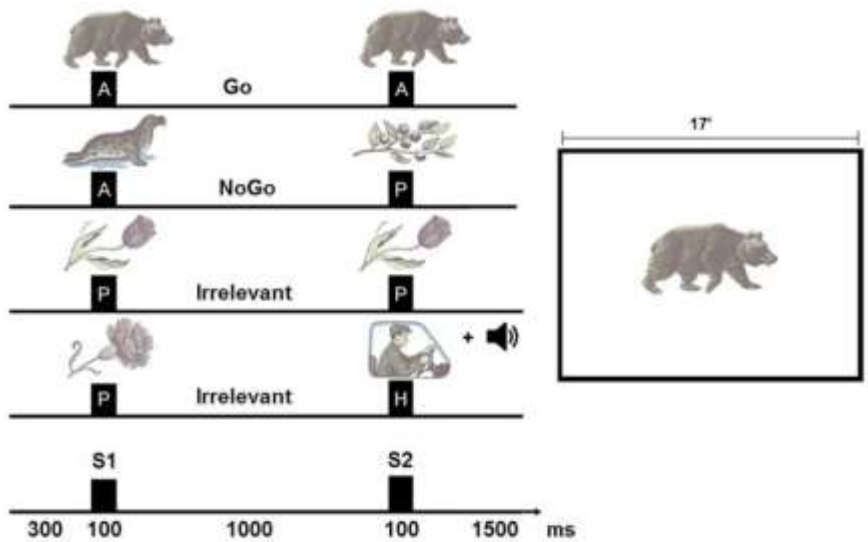


Figure 2: The cued Go/NoGo task [31]. Left: GO trials (A-A) require the participant to click the button NOGO trials need the suppression of a planned activity. IGNORE trials are stimulus pairs that begin with a plant image and require no action preparation. Novel trials are stimuli coupled with no action required, consisting of the presentation of a novel sound and a second human image stimulus. At the bottom, time intervals are shown. Each trial lasted 3000 msec in total, with the first S1 and S2 stimuli starting at 300 msec and 1400 msec into the trials, respectively. The stimulus presentation lasting was 100 msec. Right: The magnitude of the stimuli in relation to the 17-inch computer screen is shown [32].

Reaction time (RT) is the average time for correct answers computed across trials for each subject. A response was considered valid if it happened concerning the appropriate second stimulus and occurred between 200 and 1000 msec following the presentation of the second stimulus. Commission and omission errors were also calculated for each subject. An omission error occurs when the button for event 1 is not pressed, while a commission error occurs when the button for events 2, 3, and 4 is pressed instead.

2.2 Pre-processing

Because of the low signal-to-noise ratio and many artifacts found in EEG recordings, preprocessing the signals is an essential step. Therefore, EEG signals were preprocessed to eliminate artifacts caused by eye blinking and variation in impedance associated with movements. This stage was done using the EEGLAB toolbox [33]. A band-pass FIR filter was utilized to remove frequencies lower than 0.5 Hz and higher than 45 Hz from the dataset to filter the low and high-frequency noises. Data were re-referenced to the average reference before analysis in the following step. Since the eye blinking affected the recorded EEG signals, these artifacts were eliminated by blind source separation using Independent Component Analysis (ICA) [34], [35].

Event-related potentials are electrophysiological responses to voluntary activity or an external stimulus. ERPs are obtained by averaging the spontaneous EEG over multiple occurrences of the same event. Assuming the event's beginning is known, the ERP is a time series encompassing the succeeding brain reaction, referred to as an epoch. ERP can be used to indicate the differences between how the subject's brain reacts to different stimuli in a particular examination context [36]. The epoch interval used depends on the purpose of the investigation. The epoch period was chosen from the beginning to the end of stimulation, i.e., 0 to 3000 msec after stimulation. Finally, 100 msec before the stimulation was chosen due to the baseline correction.

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224. https://doi.org/10.1016/j.bspc.2023.105224

2.3 ERP components Extraction

In task-based EEG studies, the recorded signals include evoked potentials (Eps) and other unrelated information. ERPs refer to averaged EEG responses that are time-locked to visual, somatosensory, or auditory stimuli. The averaging of single trials highlights the information related to the presentation of stimuli. Therefore, the averaged ERPs allow us to focus on ERP-related components while disregarding EEG sources spontaneously [32].

ERP waveforms consist of a series of peaks (P) and nulls (N) related to the set of underlying components [37]. Each of the components, P100, P200, P300, N100, N170, N400, and the last positive component (LPC) was extracted from each channel, and then the amplitudes and latencies of the components were determined separately for each condition and subject. Therefore, each ERP component provided two properties that resulted in 14 attributes for each channel. Because each person's signal had 19 channels, a feature vector with 266 elements was calculated for each subject.

2.4 Statistical analysis of extracted features

The Kolmogorov-Smirnov test was used to investigate the statistical significance of the group differences regarding behavioral performance and the amplitudes and latencies of the components [38]. The p-value for selected features was less than 0.05.

2.5 Feature selection

A project in machine learning depends heavily on feature selection; therefore, feature selection is essential for various reasons in a study. It provides for faster training of the machine learning algorithm while reducing the model's complexity and making it easier to evaluate. Selection of the proper subsets' feature decreased overfitting and improved accuracy. The number of extracted features in this study is far higher than in the supplied dataset. It lowered the dimensions to find the most crucial dyslexic features and increased the classifier's accuracy. Therefore, the feature selection method was used to make the dimension of the feature space smaller.

After selecting the subset of features using the KS-test, which makes a statistically significant difference between the two classes, we need to decrease the number of features again because these features did not produce favorable results in classification. After that, the PCA method was used to reduce the extracted feature's dimension. This method is a statistical technique that identifies the dimensions of a data covariance matrix using eigenvalue decomposition and helps reduce such datasets' dimensionality, increase interpretability, and minimize information loss. Briefly, the data is linearly transformed into a new coordinate system where many variations in the data can be visible with fewer dimensions than the initial data [39].

2.6 Ensemble learning

The predictions of multiple weak learners are combined in ensemble learners. Different strategies for predicting weak learners can be combined. In our method, we used the voting model and assessed its performance. In the voting method, different classifiers are trained with the same data, and the prediction output for each classifier is given as 1 or 0 [40]. The majority of votes received (1 or 0) is considered the final output. We utilized SVM, Decision Tree (DT), K-Nearest Neighbor (KNN), Bayesian classification (Nb), and Latent Dirichlet Allocation (LDA) as our ensemble learning models, shown in Figure 3. These categories are explained below.

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

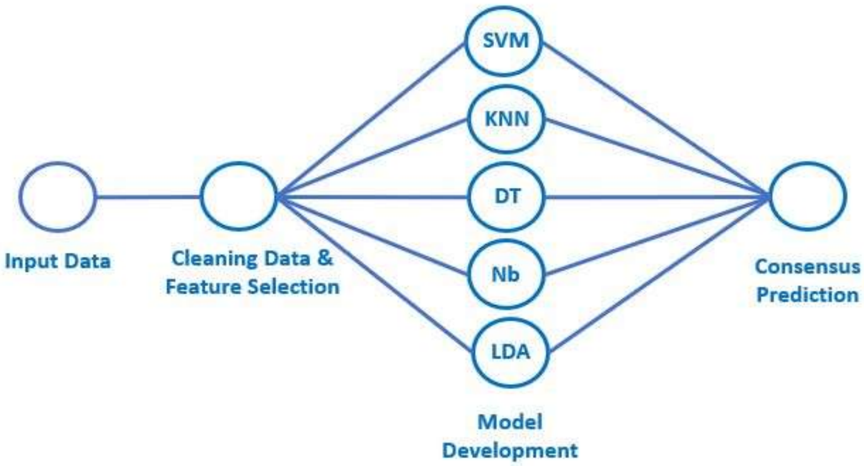


Figure 3: Overview of the modeling process ending with the voting of prediction.

SVM: Support-vector machine is a supervised learning model associated with learning algorithms utilized in machine learning to classify. SVMs can perform a linear classification and efficient non-linear classification using the kernel trick, which implicitly maps the input data into high-dimensional feature space [41].

DT: A decision tree is a supervised non-parametric learning method for regression and classification [42]. The purpose is to create a model that predicts the value or class of a target data by learning simple decision rules derived from the data features. Each tree node operates on an attribute, and the leaves that extend from this attribute are used to inspect the class label based on the attribute's value.

KNN: The k-nearest neighbor method is a classification technique that is often used. This algorithm categorizes new objects based on their properties and training data. The Euclidean equation formula is utilized to classify objects based on training data with the nearest distance to a new item [43], [44].

LDA: The linear discriminant analysis method is one of the most widely used classification algorithms. The LDA technique operates through the calculation of variance values within and between classes [45]. This algorithm must compute scatter matrices both within and between categories.

Nb: Naive Bayes is a well-known machine learning classifier that classifies a new data point using training-phase-computed probabilities [46]. The Naive Bayes theory is based on assumptions of independence. Compared to SVM models, it takes less time to train. It takes a data set’s prior probability and likelihood and figures out the posterior probability [47].

K-fold cross-validation is a resampling method for assessing machine learning models on a limited sample of data. The procedure has only one parameter called k, which refers to the number of groups a given data sample will be split into. One fold is used for testing, while the others are for training. Every sample is utilized at least once in training and testing [48].

3. Results

3.1 Statistical analysis of ERP components

As mentioned in the last section, seven components of ERP were taken out and investigated in this study based on the results of statistical tests. An extracted ERP signal associated with one of the channels in each of the four events (a to d) is shown in Figure 4. This figure shows the difference in stimulation signals between normal and

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224. https://doi.org/10.1016/j.bspc.2023.105224

dyslexic subjects. When the diagrams of dyslexics and normal people are compared visually and in terms of quality, it is clear there is a significant difference in the waveform of people with dyslexia.

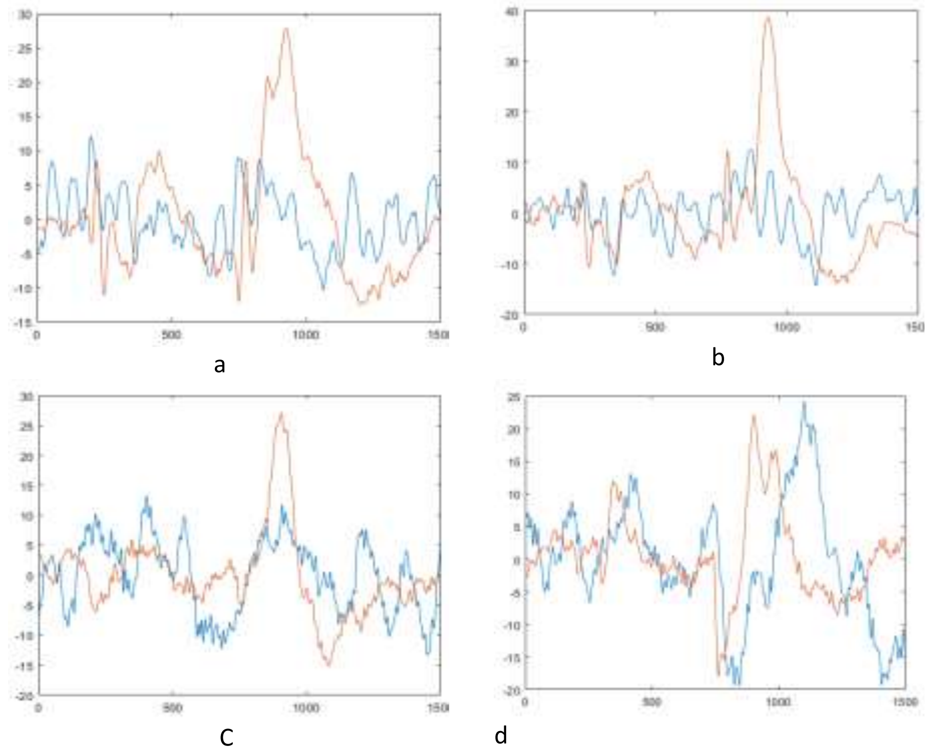


Figure 4: An extracted ERP signal associated with one of the channels in each of the four events (Event1 to Event4 [a to d]) in normal (red) and people with dyslexia (blue).

P100 Component:

The statistical test results for channels with significant differences are shown in Table 1. The P100 component differs significantly between the normal and dyslexic groups in the Cz, P4, and Pz channels, as seen in this table. This component varies in response to alternation in the stimulus. Because of this, it has to deal with new stimuli and be tested on them. Thus, in Events 1 and 2, we saw the most significant change in the amplitude of the P100. This indicates that a higher detection power was considered in these two events. Also, P100 latency varies significantly depending on the similarity or distinctiveness of the uploaded images. According to the Table, the results showed the difference in the delay time of this component between events 1 and 4, which have the most variance and similarity in the center.

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224. https://doi.org/10.1016/j.bspc.2023.105224

Table 1: **Statistical Results of the P100 component between dyslexic and normal subjects**

Events	**Amplitude Latency**	**Channel**	**Mean±SD Dyslexia**	**Mean±SD Normal**	**P-value**
Event 1	Amplitude	Cz	1.80±6.73	6.37±9.77	0.0356
	Latency	Cz	100.81±13.48	110.27±11.76	0.0356
Event 2	Amplitude	P4	4.31±4.23	8.45±6.70	0.0138
Event 4	Latency	Pz	103.18±14.02	109.63±11.37	0.0363

P200 component:

Table 2 displays the statistical test outcomes for channels with appreciable variations. This table shows that variations in this component have been seen across several channels. This component is related to aspects of perceptual processing. This is different from P300, which is more common in complex situations. P200 is more prevalent when simple stimulus features define goals. The most significant differences in this P200's amplitude and latency could be seen in Events 3 and 4. Since the first image was not an animal, it was easier for the subjects to process the stimulus than in the first and second events. The frontal lobe, the center of logic, reasoning, and memory, showed the most variation. According to the definition of dyslexia in the introduction, in this region, there was a considerable discrepancy between the groups.

Table 2: **Statistical Results of the P200 component between dyslexic and normal subjects.**

Events	**Amplitude Latency**	**channels**	**Mean±SD Dyslexia**	**Mean±SD Normal**	**P-value**
Event 1	Amplitude	Fp2	6.57±16.56	7.70±7.63	0.0356
	Latency	O2	202.63±32.43	176.36±24.90	0.0138
Event 2	Latency	F8	186.09±42.67	165.90±31.74	0.0133
		T3	200.45±47.33	170±36.01	0.0138
		T5	186.36±40.86	162.9±26.72	0.0138
		O2	207.18±32.22	181.18±33.27	0.0131
Event 3	Amplitude	Fp1	8.01±15.95	8.64±7.99	0.0366
		F7	3.22±7.23	4.74±3.31	0.0015
		F3	3.43±6.21	4.94±3.22	0.0138
		F8	3.97±6.26	5.27±2.82	0.0049
		P4	6.59±3.09	9.23±5.14	0.0049
		Cz	8.23±8.19	11.10±6.55	0.0356
		Pz	7.59±5.27	12.84±6.13	0.0035
Event 4	Latency	P3	240.18±46.15	207.72±48.42	0.0356
		C4	193.45±43.84	167.63±19.95	0.0158
		T5	204.18±52.95	174.81±40.94	0.0482
		T6	197.90±38.40	175.63±22.19	0.0356

P300 component:

Table 3 shows no significant difference between the normal and dyslexic groups in the amplitude of P300. However, there were differences in the latency of this component on several channels. Repetitive and low-repetition target stimuli generate this component. In the presence of low-repetition stimuli, participants were instructed to press a button. P300 reflects the processes involved in assessment, stimulus estimation, and classification. Thus, the P300 is defined by its sensitivity to the likelihood of occurrence. Its amplitude is higher when subjects try harder in the task and is lower when they are doubtful whether a stimulus is a target or not. According to this table, there were no significant differences in the amplitude of the component. Results indicate

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

that the two groups made approximately the same amount of effort. A delay in the decision will make the P300 latency longer; therefore, it is essential to note that the P300 appears after the stimulus that caused the decision. Various studies have demonstrated that the increase in latency relates to sensory processing and decision-making.

As shown in Table 3, the latency of P300 differed significantly between the dyslexic and normal groups because children with dyslexia had difficulty making decisions. The latency of this component will be longer if individuals have a perception impairment. Perception is caused by deficits in the frontal and parietal lobes, which are linked to the ability to infer and solve issues. According to the results, the differences were more prominent in the parietal lobe.

Table 3: Statistical Results of the P300 component between dyslexic and normal subjects.

Events	**Amplitude Latency**	**channels**	**Mean±SD Dyslexia**	**Mean±SD Normal**	**P-value**
Event 1	Latency	C3	369.72±69.78	319.45±29.62	0.0138
		Cz	331±50.71	310.18±49.6	0.0138
Event 2	Latency	C3	397.09±49.76	365.27±49.43	0.0353
		C4	405.45±54.97	356.18±56.59	0.0138
		T4	381.27±54.97	348.63±62.83	0.0356
		T5	382.9±52.78	359±68.35	0.0356
		O2	368.81±45.84	334±47.39	0.0171
		Pz	352.54±60.95	318.54±59.87	0.0138

LPC component:

Table 4 shows only one significant difference between the normal and dyslexic groups in the C3 channel. This component is found in meaningless stimuli, shapes, and sounds by definition. The LPC amplitude did not substantially differ between the two groups. As shown in Table 4, the latency of Event 4 is different in the central electrodes because the sound in this event had no special meaning to the participants.

Table4: Statistical Results of the LPC component between dyslexic and normal subjects.

Events	**Amplitude Latency**	**channels**	**Mean±SD Dyslexia**	**Mean±SD Normal**	**P-value**
Event 4	Latency	C3	545.18±46.56	521.63±37.85	0.0356

N100 component:

The findings for this component are shown in Table 5. In this study, no results were observed regarding the difference in the N100's amplitude between two groups of dyslexics and non-dyslexics. However, concerning the latency of this component in Event 4, there was a significant difference between these two groups in Cz and Fz channels related to the central parts of the frontal lobe. The observed discrepancy is because N100 occurs in both auditory and visual stimuli, which indicates the different functioning of the dyslexic brain to auditory stimuli.

Table 5: Statistical Results of the N100 component between dyslexic and normal subjects.

Events	**Amplitude Latency**	**channels**	**Mean±SD Dyslexia**	**Mean±SD Normal**	**P-value**
Event 4	Latency	Cz	126.45±29.21	130.45±38	0.0356
		Fz	118.18±31.44	112.81±27.31	0.0363

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

N170 component:

Table 6 shows several significant differences between the normal and dyslexic groups in the four channels (P4, T6, Fp1, and C3). In Event 4, this component had a different latency in the left hemisphere's brain between dyslexics and non-dyslexics. N170 is frequently sensitive to face processing. This explains the difference in N170 components' latency between the two groups in Event 4. By definition, it is a measure of the similarity of stimuli. In Event 1, the two pictures were more similar to each other, so it is clear that we should have observed a variation in amplitude between these groups. The results of the statistical test in Table 6 confirm this analysis. This difference is found in the temporal and parietal lobes of the brain.

Table 6: Statistical Results of the N170 component between dyslexic and normal subjects.

Events	**Amplitude Latency**	**channels**	**Mean±SD Dyslexia**	**Mean±SD Normal**	**P-value**
Event 1	Amplitude	P4	-2.14±4.14	-4.46±4.43	0.0356
		T6	-2.41±3.48	-4.82±3.92	0.0356
Event 4	Latency	Fp1	186.9±28.74	171.81±31.51	0.0356
		C3	194.45±32.43	183.36±42.69	0.0138

N400 component:

The results for the N400 component are shown in Table 7. Since N400 is the brain's natural reaction to meaningful stimuli such as visual and auditory, dyslexics constitutionally had differences in the amplitude and latency of this component compared to normal subjects. According to the results of this investigation, statistical analyses showed that Events 1, 2, and 4 had different amplitudes, as shown in Table 7. In addition, similar stimuli might trigger a response in the N400 component. Therefore, the difference in amplitude grows linearly with the similarity of the stimuli, which is why Event 1 had the most considerable amplitude diversity. These variations in amplitude and latency were more prominent in the brain's center and right hemispheres. The right side controls memory, attention, reasoning, and solving problems and is responsible for image processing and movement on the left side of the body. As well as a person's ability to detect items and comprehend their relationship to space is helped by the center of their brain.

Table 7: Statistical Results of the N400 component between dyslexic and normal subjects.

Events	**Amplitude Latency**	**channels**	**Mean±SD Dyslexia**	**Mean±SD Normal**	**P-value**
Event 1	Amplitude	P4	3.07±5.31	7.03±6.44	0.0356
		Cz	5.20±5.40	10.89±8.12	0.0085
	Latency	C3	380.18±98.45	449.54±77.95	0.0138
		O2	399.36±98.04	457.54±62.02	0.0354
Event 2	Amplitude	C3	4.02±5.02	2.74±4.37	0.0356
Event 3	Latency	C4	341.45±99.60	314.36±90.54	0.0354
Event 4	Amplitude	F4	10.32±11.67	11.37±4.72	0.0370
		T3	7.08±10.48	9.45±5.01	0.0356
	Latency	Fp1	303±78.87	325.63±58.1	0.0353
		Fp2	303.54±83.4	327.72±56.67	0.0049

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

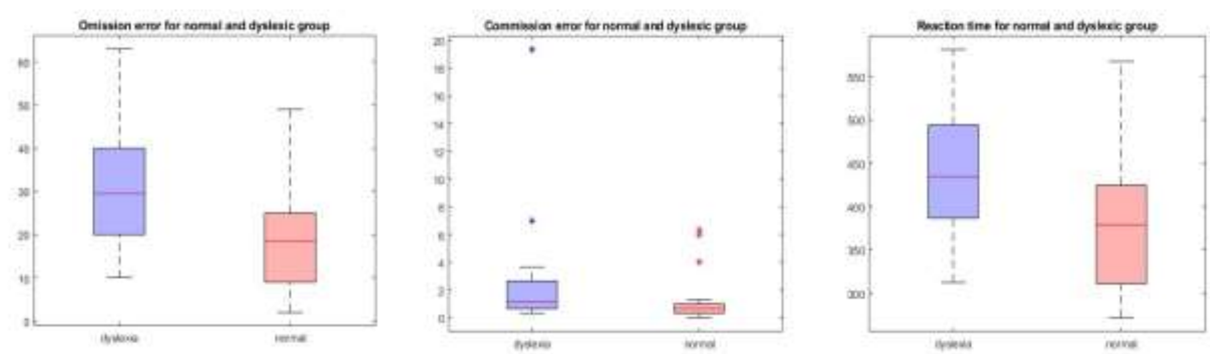
3.2 Behavioral analysis

We acquired two types of mistakes to study the participants' behavior in reaction to the stimulus: omission error and commission error. A box plot in Figure 5-a demonstrates the omission errors in both the dyslexic and normal groups. This error had a lower mean in the normal group than in the dyslexic group, as seen in the boxplot below. According to the description of dyslexia, people with dyslexia have more mistakes and reactions to stimuli than normal people, shown as a box representation of commission error in Figure 5-b. The number of normal people’s errors was lower than dyslexics, indicating that the subjects diagnosed with dyslexia pressed the mouse more than usual. These two errors show that diagnosing stimulus is much harder for dyslexic people than for those who do not have the disorder.

Participants' average reaction time to stimulation is one of the essential criteria in the behavioral analysis for those who have taken part in the study. Figure 5-c depicts a diagram illustrating this criterion in further detail. People with dyslexia had an average reaction time longer than normal ones. People without dyslexia responded on average in 386.68 msec, while people with dyslexia took 437.5 msec. Table 8 shows both commission and omission errors as well as reaction time.

Table 8: Behavioral results on VCPT task for dyslexic and normal subjects.

	Dyslexia	**Normal**
	Mean±std	**Mean±std**
Omission error	31.14+14.08	18.54±11.8
Commission error	3.59±3.44	2.09±4.08
Reaction time(msec)	437.5±71.55	386.68±87.86



a **b** **c**

Figure 5: **Omission error (a), commission error (b), and reaction time (c) for both dyslexic and normal subjects.**

3.3 Classification results

We implemented our proposed approach using MATLAB. Seven ERP components were extracted from each channel in every subject after preprocessing, and their amplitudes and latencies were estimated as signal characteristics, which resulted in 266 features from 19 recording channels. After applying the statistical test (KS-test) and finding the statistically significant features, we used the PCA technique to effectively reduce the obtained

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

feature vector. PCA is a technique for extracting strong patterns from large and complex datasets. It aids in identifying the most important features within a dataset. This approach led to the selection of five principal components, which were fed to the classifier to distinguish dyslexic and normal subjects. Figure 6 shows the average weight of each feature in 8-fold cross-validation for Event 2. In fact, this figure can be used to understand and extrapolate the characteristics that were useful in creating each PCA component. For example, in PC4, the 7th feature, which is associated with P300 latency in C3, has the highest average weight among other characteristics, indicating that it played the most important role in the creation of this component.

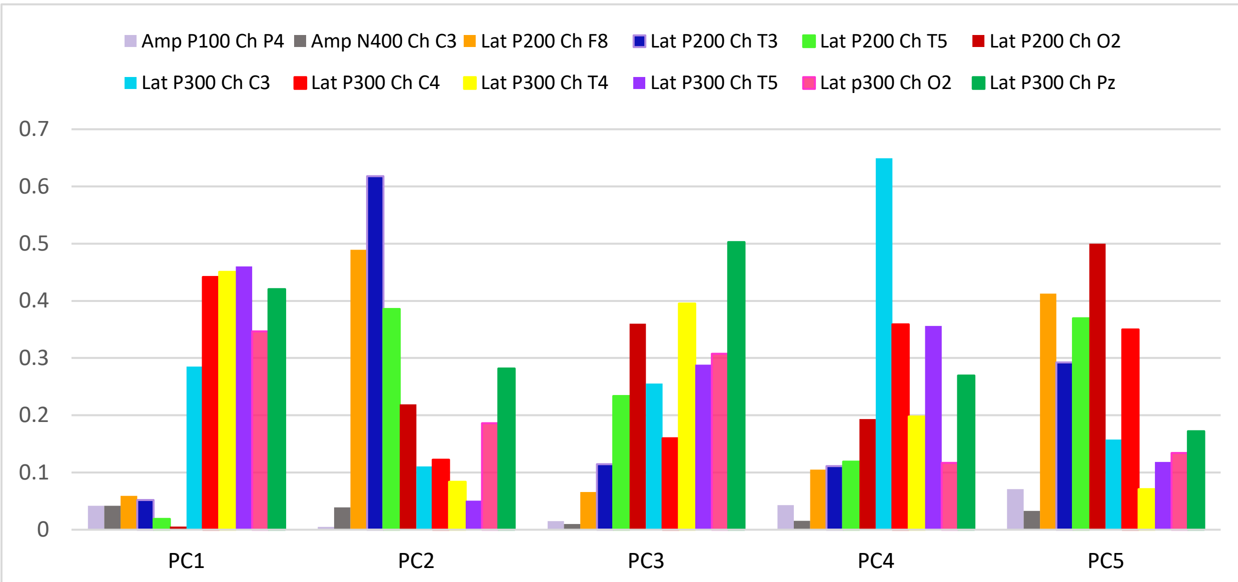


Figure 6: Mean weight of extracted features in PCA components. The diagram displays the mean of the extracted features from 8-fold cross-validation that were represented in distinct PCA components.

The results obtained from the experiment are presented in Tables 9 and 10. We used common operating metrics, such as accuracy, sensitivity, specificity, and F1 score, to evaluate the performance of various models. Table 9 demonstrates the performance of single classifiers such as KNN, SVM, nb, LDA, and DT, as well as Table 10, which shows the performance of the voting method of ensemble classifiers on the dataset by using 8-fold cross-validation. Event 2, which deals with the person's suppression of the initial decision, had an accuracy of 87.5%, a sensitivity of 81.2%, a specificity of 93.7%, and an F1-score of 86.92%, which were excellent findings in the diagnosis of this disorder.

In addition, leave-one-out cross-validation (LOOCV) was applied. LOOCV is K-fold cross-validation, with K equal to N, the number of data points in the set. That means that the classifier is trained N times on all the data except for one point, and a prediction is made for that one point. The results are presented in Table 11, and the F1-score for event 2 is equivalent to 80%. Results from LOOCV are acceptable and verify that the suggested algorithm is an appropriate method for detecting dyslexia.

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

Table 9: Classification results of each classifier using 8-fold cross-validation.

	KNN		
	Accuracy% mean±SD	**Sensitivity%** mean±SD	**Specificity%** mean±SD
Event 1	65.6±10.2	62.5±20.4	68.7±22.3
Event 2	78.1±12.1	68.7±26.7	87.7±12.2
Event 3	56.2±9.4	68.7±18.8	43.7±27.3
Event 4	59.3±12.9	50±27.3	68.7±22.3
	Tree		
	Accuracy% mean±SD	**Sensitivity%** mean±SD	**Specificity%** mean±SD
Event 1	75±12.9	62.5±18.8	87.5±11.6
Event 2	81.2±12.9	68.7±20.4	93.7±11.1
Event 3	62.5±10.2	56.2±22.3	68.7±20.4
Event 4	68.7±13.6	56.2±20.4	81.2±11.9
	LDA		
	Accuracy% mean±SD	**Sensitivity%** mean±SD	**Specificity%** mean±SD
Event 1	81.2±12.9	87.5±11.5	75±20.4
Event 2	84.3±13.3	75±25.8	93.7±9.6
Event 3	56.6±10.2	62.5±18.8	68.7±23.1
Event 4	71.8±12.9	68.7±25.8	75±26.7
	SVM		
	Accuracy% mean±SD	**Sensitivity%** mean±SD	**Specificity%** mean±SD
Event 1	75±13.6	87.5±10.9	62.5±22.3
Event 2	80.5±12.5	77.7±14.9	83.3±11.2
Event 3	62.5±11.1	68.7±20.4	62.5±22.3
Event 4	65.6±13.6	62.5±22.7	68.7±25.8
	Nb		
	Accuracy% mean±SD	**Sensitivity%** mean±SD	**Specificity%** mean±SD
Event 1	75±13.6	75±22.3	75±37.7
Event 2	75±12.1	81.2±10.6	68.7±22.3
Event 3	65.6±12.9	62.5±18.8	68.7±20.4
Event 4	75±13.6	75±20.4	75±22.3

Table 10: Classification results based on ensemble learning using 8-fold cross-validation.

	Accuracy% mean±SD	**Sensitivity%** mean±SD	**Specificity%** mean±SD
Event 1	75±13.6	81.2±12.6	68.75±18.8
Event 2	**87.5±8.7**	**81.2±11.8**	**93.7±6.9**
Event 3	59.37±12.9	62.5±13.6	56.25±20.6
Event 4	71.8±12.9	68.75±16.9	75±10.3

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

Table 11: Classification results based on ensemble learning using LOOCV.

	Accuracy%	Sensitivity%	Specificity%
Event 1	71.8	81.2	62.5
Event 2	**81.25**	**75**	**87.5**
Event 3	65.6	62.5	68.7
Event 4	68.7	62.5	75

4. Discussion

Dyslexia is a neurodevelopmental condition affecting individuals of any race or ethnicity in any part of the world. The disorder is often referred to as a learning impairment, and it is differentiated by core cognitive abnormalities in the linguistic system, namely in phonological processing, as well as functions related to the magnocellular visual pathways ([49], [50], [51], [52]). As access to objective diagnostic information is limited, in clinical practice, diagnoses are frequently based on the patient's subjective observations and perceptions and the clinician's diagnostic criteria. This study aims to contribute to the enhancement of objectivity in dyslexia diagnosis. We investigated if features of various ERP components could be used to discriminate between children with dyslexia and normal subjects accurately. To achieve this, we used the visual two-stimulation GO/NOGO task (VCPT) to obtain ERP responses under four different task situations. Then, the ERPs were computed by applying the method of averaging, and after being preprocessed, the ERP signal was decomposed into a set of seven components. And then we used statistical and behavioral analysis and machine learning techniques. Finally, we could diagnose this disorder with 87.5% accuracy and 81.2% sensitivity in Event 2.

The five best components were selected to investigate the best features of the PCA technique. The latency characteristics of P300 and P200 waves in T3, T5, Pz, T4, C4, and O2 channels had a key role in developing these five best components. In cognitive studies, the P200 wave in brain-evoked potentials is significant because it reflects high-level perceptual processing and attention (such as visual attention, memory, and language). The P300 wave also reflects the processes involved in evaluating or classifying the stimuli. The more substantial contribution of these two waves in generating the best five components for classification indicates that people with dyslexia perform poorly in perceptual processing, recognition, and classifying stimuli.

According to the results of the Ks-test in this study, the most discriminative electrodes with the most repetitions, such as C3, Fp1, and T5 were found in the left hemisphere, which can be related to the Wernicke and Broca areas. Previous research has indicated that these areas have an essential role in the brain's reading, writing, and language-related functions, as well as dyslexia anomalies [53]. Also, this result has been confirmed in research [20] and [54] that most of the differences between dyslexic and normal children are located in the left hemisphere.

There are classification algorithms for dyslexia based on classic machine learning (ML) models ([55-59], [19], [15]) and deep learning (DL) techniques ([9], [60]). DL algorithms have become more common in recent years, but ML algorithms are still very popular among experts due to the fact that they may be better suited for simpler issues that don't require a high degree of complexity. Moreover, ML models work better with smaller datasets and have more interpretability. In this study, the machine learning approach and the ensemble learning algorithm were applied.

According to the classification results reported in the previous section, the proposed algorithm for Event 2 performs better than the other events. As mentioned in this event, the subject is prepared to press the mouse in addition to processing the first stimulus. However, because the subjects did not see the picture of the animal, they must suppress their desire to respond. According to the findings of the classification performance results, people with dyslexia cannot frequently control their first decision and cannot analyze stimuli correctly and on time. Also, after Event 2, the classification in Event 1 resulted in more excellent performance than in other scenarios, which means that people with dyslexia in this instance also had less ability to identify and react to what was happening than normal people. These aspects are related to significant disparities in the abilities of people with dyslexia to concentrate, develop cognition, and analyze information compared to normal people. Because children do not know how to read and write before they start school, the parts of the ERP signal from the VCPT task can help make a more accurate diagnosis of dyslexia and the long-term effects of dyslexia in adulthood. This can help prevent serious problems in the education of children with dyslexia.

The several approaches to diagnose dyslexia that use EEG signals are outlined in Table 12. Previous research has frequently focused on reading and writing skills, as described in the introduction; hence, these methods and results are helpful for individuals who possess both of these talents. Because early detection of this disorder before school is critical, the task of VCPT, a type of visual stimulation that does not require reading or writing skills, was used in this study.

Table 12: Comparison of different EEG-based dyslexia detection techniques.

Authors	No. of subjects	Language	Stimulus	Classification Method	Performance (%)
Frid and Brenzite(2012) [56]	50	Hebrew	Auditory	SVM	Acc = 85 Sens = 74 Spec = 82
Karim et al. (2013) [19]	6	Malay	Resting-state	MLP	Acc (Eye-close) = 84.9 Acc (Eye-open) = 86
Al-Bahramotoshy et al. (2017) [16]	80	Arab	Gibson-test	K-means algorithm ANN Fuzzy logic classifier	Acc = 81.1 Precision = 100 F-score = 76.6
Perera et al. (2018) [15]	32	English	Typing and writing	SVM	Acc = 78.1 Sens = 88.2 Spec = 66.6
Ortiz et al. (2019) [57]	48	Spanish	Auditory	SVM	Acc = 78 Sens = 66 Spec = 81
Martinez-murica (2019) [58]	48	Spanish	Auditory	SVM	Acc = 74 Sens = 59 Spec = 79
Martinez-murica (2020) [59]	48	Spanish	Auditory	Autoencoder	Acc = 72.8 Sens = 66.7 Spec = 78.9
Shahkar et al. (2021) [60]	391	Germany France Hungary Finland	cross-linguistic passive oddball	Non-linear SVM	Acc = 62.4 Sens = 52.8 Spec = 91.4
Nicolas et al. (2022) [25]	48	Spanish	Auditory	SVM	Acc = 72.9 Sens = 72.3 Spec = 74.7 AUC = 73.3

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

Pavlos et al. (2022) [26]	26	Greek	Interactive Linguistic Software	RF	Acc = 87.04 Sens = 90.9 Spec = 80.9
Present study (2023)	**44**	**Persian**	**VCPT**	**Ensemble learning**	**Acc = 87.5** **Sens = 81.2** **Spec = 93.7**

5. Conclusion

Dyslexia, a complex brain developmental disorder, has recently attracted the attention of researchers in modern neuroscience and machine learning. We developed a method for detecting dyslexia in Persian speakers. This study's feature selection method highlights the differences in EEG features between dyslexics and normals. Also, the ensemble learning technique helped to increase the accuracy of dyslexia diagnosis in this method. The method was evaluated with 44 participants and attained 87.5% accuracy and 81.2% sensitivity, specificity of 93.7%, and F1-score of 86.92% on its prediction. The technique suggested in this study can improve children's education before they start school and have any skills in writing and reading.

Refrences:

1. Wagner, R.K., et al., *The prevalence of dyslexia: A new approach to its estimation.* Journal of learning disabilities, 2020. 53(5): p. 354-365.
2. Prabha, A.J. and R. Bhargavi, *Predictive model for dyslexia from fixations and saccadic eye movement events.* Computer Methods and Programs in Biomedicine, 2020. 195: p. 105538.
3. Iwabuchi, M., et al., *Machine learning based evaluation of reading and writing difficulties*, in *Harnessing the Power of Technology to Improve Lives*. IOS Press, 2017. p. 1001-1004.
4. Billard, C. and F. Delteil-Pinton, *Dyslexia: clinical characteristics.* Archives de pediatrie: organe officiel de la Societe francaise de pediatrie, 2010. 17(12): p. 1734-1743.
5. Thompson, P.A., et al., *Developmental dyslexia: predicting individual risk.* Journal of Child Psychology and Psychiatry, 2015. 56(9): p. 976-987.
6. Caravolas, M., *The nature and causes of dyslexia in different languages.* The science of reading: A handbook, 2005. 18: p. 336-355.
7. Lieder, I., et al., *Perceptual bias reveals slow-updating in autism and fast-forgetting in dyslexia.* Nature neuroscience, 2019. 22(2): p. 256-264.
8. Lizarazu, M., et al., *Neural entrainment to speech and nonspeech in dyslexia: conceptual replication and extension of previous investigations.* Cortex, 2021. 137: p. 160-178.
9. Zahia, S., et al., *Dyslexia detection using 3D convolutional neural networks and functional magnetic resonance imaging.* Computer methods and programs in biomedicine, 2020. 197: p. 105726.
10. Conant, L.L., et al., *Differential activation of the visual word form area during auditory phoneme perception in youth with dyslexia.* Neuropsychologia, 2020. 146: p. 107543.
11. Jothi Prabha, A. and R. Bhargavi, *Prediction of dyslexia from eye movements using machine learning.* IETE Journal of Research, 2022. 68(2): p. 814-823.
12. Coch, D., *Uncoupled brain and behavior changes in lexical, phonological, and memory processing in struggling readers.* Developmental Neuropsychology, 2021. 46(1): p. 33-53.

13. Bosch-Bayard, J., et al., *Resting EEG effective connectivity at the sources in developmental dysphonetic dyslexia. Differences with non-specific reading delay.* International Journal of Psychophysiology, 2020. 153: p. 135-147.
14. Xue, H., et al., *Resting-state EEG reveals global network deficiency in dyslexic children.* Neuropsychologia, 2020. 138: p. 107343.
15. Perera, P., et al., *EEG signal analysis of writing and typing between adults with dyslexia and normal controls.* Int J Interact Multi Artif Intell, 2018. 5(1): p. 62.
16. Al-Barhamtoshy, H.M. and D.E.M. Motaweh, *Diagnosis of Dyslexia using Computing Analysis.* Journal of Engineering Technology, 2017. 6(2): p. 462-482.
17. Taskov, T. and J. Dushanova. *Reading multiple EEG frequency-band networks in developmental dyslexia*. in *Intelligent Human Systems Integration 2021:* Advances in Intelligent Systems and Computing, Springer International Publishing, *2021. 1322: p. 171-180*.
18. Ahram, T., et al., *Intelligent Human Systems Integration (IHSI 2022): Integrating People and Intelligent Systems*. AHFE International, 2022. Vol. 22.
19. Karim, I., W. Abdul, and N. Kamaruddin. *Classification of dyslexic and normal children during resting condition using KDE and MLP*. in *2013 5th international conference on information and communication Technology for the Muslim World (ICT4M)*. IEEE, 2013. p. 1-5.
20. Frid, A. and L.M. Manevitz, *Features and machine learning for correlating and classifying between brain areas and dyslexia.* arXiv preprint arXiv:1812.10622, 2018.
21. Fraga-González, G., et al., *Graph analysis of EEG functional connectivity networks during a letter-speech sound binding task in adult dyslexics.* Frontiers in Psychology, 2021. 12: p. 5344.
22. Dushanova, J.A. and S.A. Tsokov, *Altered electroencephalographic networks in developmental dyslexia after remedial training: A prospective case-control study.* Neural Regeneration Research, 2021. 16(4): p. 734.
23. Shaywitz, S.E., et al., *Neural systems for compensation and persistence: young adult outcome of childhood reading disability.* Biological psychiatry, 2003. 54(1): p. 25-33.
24. Zainuddin, A.Z.A., et al. *Comparison of extreme learning machine and K-nearest neighbour performance in classifying EEG signal of Normal, poor and capable dyslexic children*. in *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2019. p. 4513-4516.
25. Gallego-Molina, N.J., et al., *Complex network modeling of EEG band coupling in dyslexia: An exploratory analysis of auditory processing and diagnosis.* Knowledge-Based Systems, 2022. 240: p. 108098.
26. Christodoulides, P., et al., *Classification of EEG signals from young adults with dyslexia combining a Brain Computer Interface device and an Interactive Linguistic Software Tool.* Biomedical Signal Processing and Control, 2022. 76: p. 103646.
27. Hanafi, M.M., W. Mansor, and A.A. Zainuddin. *Recognition of EEG signals of dyslexic children using long short-term memory*. in *AIP Conference Proceedings*. AIP Publishing LLC, 2023. 2562(1): p. 060002.
28. Wechsler, D. and H. Kodama, *Wechsler intelligence scale for children*. Psychological corporation New York, 1949. Vol. 1.
29. Meissner, N.A., *A Single-Subject Evaluation of Facilitated Communication in the Completion of School-Assigned Homework.* 2018.
30. Brunner, J.F., et al., *Long-term test-retest reliability of the P3 NoGo wave and two independent components decomposed from the P3 NoGo wave in a visual Go/NoGo task.* International journal of psychophysiology, 2013. 89(1): p. 106-114.
31. Grane, V.A., et al., *Correction: ERP correlates of proactive and reactive cognitive control in treatment-naive adult ADHD.* Plos one, 2016. 11(9): p. e0163404.
32. Mueller, A., et al. *Classification of ADHD patients on the basis of independent ERP components using a machine learning system*. in *Nonlinear biomedical physics*. Springer, 2010. 4: p. 1-12.
33. Delorme, A. and S. Makeig, *EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis.* Journal of neuroscience methods, 2004. 134(1): p. 9-21.
34. Comon, P., *Independent component analysis, a new concept?* Signal processing, 1994. 36(3): p. 287-314.

35. Li, R. and J.C. Principe. *Blinking artifact removal in cognitive EEG data using ICA*. in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2006. 4: p. 1-12.
36. Kuncheva, L.I. and J.J. Rodríguez, *Interval feature extraction for classification of event-related potentials (ERP) in EEG data analysis.* Progress in Artificial Intelligence, 2013. 2: p. 65-72.
37. Luck, S.J. and E.S. Kappenman, *The Oxford handbook of event-related potential components*. Oxford university press, 2011.
38. Massey Jr, F.J., *The Kolmogorov-Smirnov test for goodness of fit.* Journal of the American statistical Association, 1951. 46(253): p. 68-78.
39. Wold, S., K. Esbensen, and P. Geladi, *Principal component analysis.* Chemometrics and intelligent laboratory systems, 1987. 2(1-3): p. 37-52.
40. Akbulut, A., et al., *Identification of phantom movements with an ensemble learning approach.* Computers in Biology and Medicine, 2022. 150: p. 106132.
41. Cortes, C. and V. Vapnik, *Support-vector networks.* Machine learning, 1995. 20: p. 273-297.
42. Utgoff, P.E., *Incremental induction of decision trees.* Machine learning, 1989. 4(2): p. 161-186.
43. Gazalba, I. and N.G.I. Reza. *Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification*. in *2017 2nd international conferences on information technology, information systems and electrical engineering (ICITISEE)*. IEEE, 2017. p. 294-298.
44. Parvin, H., *A modification on k-nearest neighbor classifier.* Global Journal of Computer Science and Technology, 2010. 10(14): p. 37-41.
45. Altay, O. and M. Ulas. *Prediction of the autism spectrum disorder diagnosis with linear discriminant analysis classifier and K-nearest neighbor in children*. in *2018 6th International symposium on digital forensic and security (ISDFS)*. IEEE, 2018. p. 1-4.
46. Wood, A., et al., *Private naive bayes classification of personal biomedical data: Application in cancer data analysis.* Computers in biology and medicine, 2019. 105: p. 144-150.
47. John, G.H. and P. Langley, *Estimating continuous distributions in Bayesian classifiers.* arXiv preprint arXiv:1302.4964, 2013.
48. Stone, M., *Cross-validatory choice and assessment of statistical predictions.* Journal of the royal statistical society: Series B (Methodological), 1974. 36(2): p. 111-133.
49. Peters, L., H.O. de Beeck, and B. De Smedt, *Cognitive correlates of dyslexia, dyscalculia and comorbid dyslexia/dyscalculia: Effects of numerical magnitude processing and phonological processing.* Research in Developmental Disabilities, 2020. 107: p. 103806.
50. Vilhena, D.d.A., et al., *Magnocellular visual function in developmental dyslexia: deficit in frequency-doubling perimetry and ocular motor skills.* Arquivos Brasileiros de Oftalmologia, 2021. 84: p. 442-448.
51. Stein, J., *The current status of the magnocellular theory of developmental dyslexia.* Neuropsychologia, 2019. 130: p. 66-77.
52. Mascheretti, S., et al., *Selecting the most relevant brain regions to classify children with developmental dyslexia and typical readers by using complex magnocellular stimuli and multiple kernel learning.* Brain Sciences, 2021. 11(6): p. 722.
53. Paulesu, E., et al., *Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning.* Brain, 1996. 119(1): p. 143-157.
54. Oliaee, A., et al., *Extraction of discriminative features from EEG signals of dyslexic children; before and after the treatment.* Cognitive Neurodynamics, 2022. 16(6): p. 1249-1259.
55. Khan, R.U., J.L.A. Cheng, and O.Y. Bee, *Machine learning and Dyslexia: Diagnostic and classification system (DCS) for kids with learning disabilities.* International Journal of Engineering & Technology, 2018. 7(3.18): p. 97-100.
56. Frid, A. and Z. Breznitz. *An SVM based algorithm for analysis and discrimination of dyslexic readers from regular readers using ERPs*. in *2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel*. IEEE, 2012. p. 1-4.
57. Ortiz, A., et al. *An anomaly detection approach for dyslexia diagnosis using EEG signals*. in *Understanding the Brain Function and Emotions: 8th International Work-Conference on the Interplay Between Natural*

This is the accepted manuscript (postprint) of the following article:
Zaree, M., Mohebbi, M., & Rostami, R. (2023). An ensemble-based machine learning technique for dyslexia detection during a visual continuous performance task. *Biomedical Signal Processing and Control*, *86*, 105224.
https://doi.org/10.1016/j.bspc.2023.105224

and Artificial Computation, IWINAC 2019, Almería, Spain, June 3–7, 2019, Proceedings, Part I 8. Springer, 2019. p. 369-678.

58. Martínez-Murcia, F.J., et al. *Periodogram connectivity of EEG signals for the detection of dyslexia*. in *Understanding the Brain Function and Emotions: 8th International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2019, Almería, Spain, June 3–7, 2019, Proceedings, Part I 8*. Springer, 2019. p. 350-359.
59. Martinez-Murcia, F.J., et al., *EEG connectivity analysis using denoising autoencoders for the detection of dyslexia.* International Journal of Neural Systems, 2020. 30(07): p. 2050037.
60. Parmar, S.K., O.A. Ramwala, and C.N. Paunwala. *Performance Evaluation of SVM with Non-Linear Kernels for EEG-based Dyslexia Detection*. in *2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE, 2021. p. 1-6.