Open Peer Review on Qeios

Can we use a biomarker detection algorithm to measure the effectiveness of 14-channel neurofeedback in dyslexia?

Gunet Eroglu, Raja Abou[1]

1 Isik University

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

The most common neurological diversity that children experience is dyslexia and it manifests itself in reduced reading ability. There is a genetic predisposition for dyslexia, and more recent theories explain it as a disruption in left hemispheric lateralization that reduces effective reading and writing. A software for smartphones called Auto Train Brain helps children with dyslexia to improve their reading comprehension and reading speed. Measuring the efficacy of the mobile app training was done manually with psychometric tests beforehand and we use a biomarker detection software to measure the efficacy of the neurofeedback. Machine learning (ML) techniques have recently been used to classify children with dyslexia and typically developing children (TDC). The data consists of 100 sessions of 2-minute resting-state eyes-open 14-channel Quantitative Electroencephalography (QEEG) data from 100 children with dyslexia and that of 100 TDC. We used the dyslexia biomarker detection software to assess the efficacy of 14-channel neurofeedback that was applied with Auto Train Brain. The results have shown that 30% of the sessions of children with dyslexia were classified as electrophysiologically normal, and 61% of the children with dyslexia were classified as electrophysiologically normal for at least 1 session after 20th sessions of neurofeedback.

Günet Eroğlu[*a], **MHD Raja Abou HARB**[b]

[a] *Computer Engineering Dept, Engineering and Nature Faculty, Bahçeşehir University, Istanbul, Turkey*

[b] *Computer Engineering Dept, Engineering and Nature Faculty, Işık University, Istanbul, Turkey*

**** Corresponding author:***

Tel.: +905323265582.

E-mail address: gunet.eroglu@healthmobilesoftware.com

Keywords: Dyslexia detection, developmental dyslexia biomarkers, QEEG, Artificial-Neural Network, Auto Train Brain.

1. Introduction

Developmental dyslexia is a specific learning disorder that affects reading ability and is not related to intelligence (American Psychiatric Association, 2013). It has a strong genetic component, with dyslexic individuals being more likely to have dyslexic children (Francks et al., 2002; Van Bergen et al., 2012). Other potential causes of dyslexia include maternal stress, autoimmune responses and infections during pregnancy (D'Souza et al., 2016), and an overactive stress response leading to early maturation at the expense of neuroplasticity (Kershner, 2019, 2021). Developmental dyslexia may also be characterized by delays in the lateralization of the brain, specifically in the left hemisphere (Yılmaz & Akyüz, 2021). It is important for children to establish left-brain dominance before entering school in order to facilitate learning (Kershner, 2020).

There are different subtypes of dyslexia, which can be characterized by various underlying factors such as auditory processing deficits (Schulte-Körne, 2010), visual processing impairments (Lassus-Sangosse et al., 2008), attentional deficits (Bednarek et al., 2004), abnormal eye movements (Bellocchi et al., 2013), and irregularities in processing (Horwitz, 1998). Dyslexia is characterized by difficulty in accurately translating written letters (graphemes) into corresponding sounds (phonemes) (Rubinsten et al., 2006). It may also involve abnormalities in certain subcortical structures (Geschwind & Galaburda, 1985). Research suggests that poor integration of auditory and visual information may be a diagnostic feature of dyslexia, possibly resulting from a weakness in the early binding of audiovisual speech in the sensory cortices (Kershner, 2021). This model suggests that dyslexia may be caused by dysfunction in the large-scale ventral and dorsal attention networks that control such binding, and that excessive activity in these networks due to the Locus coeruleus-norepinephrine system may interfere with multisensory integration and impair the acquisition of reading by disrupting grapheme-phoneme conversion (Kershner, 2021).

A new theory in evolutionary developmental biology suggests that early exposure to stress may be a cause of dyslexia. This theory combines research in the neurobiology of stress with the Extended Evolutionary Synthesis (EES) framework, which focuses on the role of genes in maturation and the evolution of reading skills in literate societies (Kershner, 2020). According to this theory, chronic stress can overactivate the hypothalmic-pituitary-adrenal stress axis, disrupting the balance between epigenetic, stress-induced, and cognitive-growth genetic programs in the brain (Kershner, 2020). This can compromise neuroplasticity and learning potential, particularly in the prefrontal cortex and networks associated with dyslexia (Kershner, 2020). The theory also suggests that stress can interfere with the coordination of afferent signals with cognitive performance in the insular cortex, amygdala, and hippocampus, and can over-recruit the brain's Default Mode network (Kershner, 2020). Additionally, stress can amplify the release of stress hormones from the Locus coeruleus/norepinephrine system, which can impair the entrainment of oscillations in the lower phonological frequencies of speech (Kershner, 2020). These changes in the brain's cellular and regional networks may contribute to the development of dyslexia. This evidence supporting a stress-growth imbalance in dyslexia is preliminary, but holds promise for a new understanding of the neurobiology of dyslexia and potential ways to reduce its prevalence (Kershner, 2020).

There is also evidence to suggest that inflammation may play a role in dyslexia. Inflammation is a natural response of the body to injury or infection, and it involves the activation of immune cells and the release of inflammatory molecules (Yektas et al., 2022). Some research has found that individuals with dyslexia have higher levels of inflammation in the

brain, and that inflammation may contribute to the neurocognitive changes associated with dyslexia (Estes et al., 2016; Jiang et al., 2018; Kuban et al., 2015; Yektas et al., 2022; Eroglu, 2022). However, the specific mechanisms by which inflammation and stress may contribute to dyslexia are still not fully understood and require further research.

Multiple studies have demonstrated that children with dyslexia exhibit slower brain waves at FC5 and F7, and do not experience changes in beta-1 activity while reading, in areas related to Broca's area (FC5, which is involved in speech production and articulation) and the Angular gyrus (CP5, P3, which is involved in understanding semantics and mathematics) (Klimesch et al., 2001). Additionally, there is reduced activity in the left parieto-occipital area (P7, O1) (Rippon & Brunswick, 2000). Children with dyslexia may also have increased sluggish activity in the right temporal and parietal regions (P8 and T8) (Arns et al., 2007), and reduced activity in the left temporal region (Thornton & Carmody, 2005). People with both dyslexia and ADHD may also exhibit increased sluggish activity in the frontal region of the brain. The alpha and beta bands show increased coherence in the right temporal region (T3 and T4), while the delta and theta bands show increased coherence across both hemispheres (Arns et al., 2007). However, there is reduced coherence in the delta, theta, and alpha bands between P7 and O1. Difficulties in the gamma band and reduced functioning connections have also been linked to dyslexia (Fraga González, 2018; Kraus, 2012). The reduced activity in the left temporal region is associated with difficulties in auditory processing and reading (Thornton & Carmody, 2005). These abnormalities in various brain regions provide evidence of underlying neurological dysfunction in the temporo-occipital and parietal-temporal regions, particularly in ventral attentional networks, and may significantly impact dyslexia.

Children with dyslexia may require a significant amount of time and support, including proper nutrition, education, and assistance, in order to catch up to their peers in reading ability. This is because dyslexia is often accompanied by deficits in phonological processing, including the ability to consciously manipulate speech sounds (phonological awareness), temporarily store phonological information in verbal short-term memory, and quickly retrieve long-term phonological representations (McDougall et al., 1994). These difficulties may make it challenging for children with dyslexia to accurately process and comprehend written language.

We focused on the study of the previous works in the publications that specialized in classifying dyslexia among children through supervised machine learning techniques using EEG datasets.

A lot of works in the literature worked on Support Vector Machines (SVM) classifiers in their proposals. (Zainuddin et al., 2018) focused on Radial Basis Function (RBF) kernel in SVM and the reached accuracy was 0.91. EEG data were recorded from 33 participants with the age ranging from 7 to 12 years old, 8 of them were normal, 17 were poor dyslexics and 8 were capable dyslexics. This data was acquired with the assistance of the Dyslexia Association of Malaysia and Rakan Dyslexia Malaysia group. (Rezvani et al., 2019) reached 0.95 of accuracy with their proposal, and the data was collected using 44 participants, 29 of them were dyslexic and the rest were good readers. The studied age was the age of third grade. (Kheyrkhah Shali & Setarehdan, 2020) focused on identifying the brain parts that are related to reading to improve them. Data was taken from the Atiyeh clinic center. The sample group consisted of 30 primary school students equally separated between dyslexic and good reader children. They used 19 channels for EEG signals and sampled at 250 Hz.

(Karim et al., 2013) worked on Multi-Layer Perceptron (MLP) in solving the research problem. The Dataset was collected using Dyslexia Association of Malaysia, at Ampang Hilir and Titiwangsa branch for the dyslexic children, and the good-reader children were selected randomly from different schools. The age range was from 4 to 7 years old. the highest reach accuracy was 0.86 with eyes closed condition.

(Formoso, Ortiz, Martinez-Murcia, et al., 2021) used Naïve Bayes (NB) Classifier for classification. Data was taken from the Leeduca group at the University of Málaga. Control and experimental groups are extracted by a careful screening process from a cohort (N = 700) followed from 4 years to the second evaluation of 7 years in 20 different primary schools (Junta de Andalucía). The reached accuracy was 0.82 with the 4.8 Hz stimulus.

Two main workes focused on K-Nearest Neighbors (KNN) classifiers in solving the research problem. (Zainuddin et al., 2016) reached through this methodology 1.00 in accuracy for normal and capable dyslexic children using Euclidean with k-value at 5 for Random and Nearest rule. But no checking for overfitting was done. They collected EEG data from 21 participants between 7 and 12 years old, 7 of them were classified as capable dyslexic, while 4 of them were classified as poor dyslexic, and the rest were good readers. (Zainuddin et al., 2019) worked on KNN and extreme learning machine (ELM) classifiers to classify dyslexia. The EEG data was self-recorded. The number of participants was 20, equally separated between dyslexic and good reader children. The age range was from 7 to 12. The highest accuracy was 0.89 and it was scored by ELM.

Several of the previous works worked on testing several ML models in their design. (Alex & Larry, 2018) tested Several ML models were experimented such as Decision Trees (DT), Neural Networks (NN), and SVM, but they mentioned SVM only. The highest accuracy reached 0.78. The dataset was self-collected, and the total number of participants was 32 Hebrew native children 17 of them were dyslexic while the rest were good readers. (Ortiz et al., 2020) worked on Deep learning with SVM classifiers. The EEG data used in this work was provided by the Leeduca Study Group at the University of Malaga. The highest scored accuracy was 0.96. (Gallego-Molina et al., 2022) worked on Phase–amplitude coupling (PAC) then SVM. EEG data were obtained by the Leeduca research group at the University of Málaga. The total number of participants was 48, 32 were good readers while the rest were dyslexic. The age range was from 88 to 100 months. The highest accuracy was 0.729 for the 4.8 Hz stimulus. (Formoso, Ortiz, Martínez-Murcia, et al., 2021) used new way to combine the ML models. They used SVM to detect the most discriminant nodes and then ensemble classifier using Random Forest (RF) and Gradient Boosting classifiers. The used data was provided by the Leeduca Research Group at the University of Málaga. EEG signals were recorded using the Brainvision acticHamp Plus with 32 active electrodes (actiCAP, Brain Products GmbH, Germany) at a sampling rate of 500 Hz during 15-minute sessions while presenting an auditory stimulus to the subject. The number of participants was 48, 17 of them were good readers while the rest were dyslexic. The Area Under Curve (AUC) metric was used and the highest value reached 0.73.

From the studied previous works we can see that the highest accuracy score was 1.00 (Zainuddin et al., 2016), however the overfitting issue that ML models have was not taken into consideration. Most of the previous works focused on SVM classifiers to solve the classification problem (Formoso, Ortiz, Martinez-Murcia, et al., 2021; Gallego-Molina et al., 2022; Kheyrkhah Shali & Setarehdan, 2020; Ortiz et al., 2020; Rezvani et al., 2019; Zainuddin et al., 2018), while other ML

methods chose KNN classifier (Zainuddin et al., 2016; Zainuddin et al., 2019), NB classifier (Formoso, Ortiz, Martinez-Murcia, et al., 2021), and MLP (Karim et al., 2013). From the studied previous works, we can see that there was a limitation in choosing supervised ML models, so there was no overall comparison of the performance of other ML models. The chosen supervised ML models were not also achieved good performances in the classification problem. Most of the previous works did not take into consideration the overfitting issues that ML models may face over the given datasets. Processing time for the studied previous works was also not mentioned.

Auto Train Brain is a cutting-edge system that combines multi-sensory learning, neurofeedback from 14 channels, and special education concepts to improve reading skills in children with dyslexia (Eroğlu et al., 2021). This system includes machine learning algorithms and has been tested in clinical trials, which have shown that it is an effective method for improving reading skills in dyslexic children, as demonstrated by pre- and post-treatment differences on the Multi scale entropy and TILLS tests (Eroğlu et al., 2021; Eroğlu et al., 2022). In this study, the researchers used a machine learning-based dyslexia diagnostic system to evaluate the effectiveness of 14-channel neurofeedback using Auto Train Brain.

2. Methods and materials

2.1. Participants

Our experiment group consisted of 100 children with dyslexia who regularly utilized Auto Train Brain at home for neurofeedback (Mage= 8.85, SD = 1.56, 75 males, 25 females; ethnic group is white). There were 100 healthy children present in the experiment (Mage = 8.75, SD = 1.50; 79 males, 21 females; ethnic group white). Psychiatrists identified the children in the experiment group as having dyslexia and advised their families to use Auto Train Brain at home. Psychiatrists used the TILLS tests to determine whether the test subjects satisfied the DSM-V criteria for dyslexia. The children were picked at random to take part in the experiment by using social media advertisements. In this retrospective study, the participant's main objective is to use Auto Train Brain software as a neurofeedback device at home. Individuals were advised to conduct a 2-minute resting state QEEG assessment for the purpose of data collection prior to each neurofeedback session. The participants utilized Auto Train Brain before leaving for school in the morning. The inclusion criteria for the study was that the participants hailed from middle-SES homes, did not use any medications, and did not have any comorbid conditions aside from dyslexia and aged between 7-10. They resided in numerous cities all around Turkey. A survey of the parents of the children was done to assess their socioeconomic position. The survey asks questions about employment, education (elementary, secondary, and graduate), and income (low income 6,000 TL, middle income 6,000 TL to 20,000 TL, high income >20,000 TL) (staff, blue-collar workers, white-collar workers).

Before each neurofeedback session, 2-minute resting-state eyes-open theta, alpha, beta1, beta2, and gamma-band powers were measured (60 sessions per subject on average). This study makes use of many measurements and a limited sample size (200 participants). Using Auto Train Brain, 2-minute resting-state QEEG data from typically developing children are collected 60 times per participant, and the data for the experiment and control groups are balanced to have an equal number of occurrences in each group.

2.2. The Test of Integrated Language & Literacy Skills (TILLS)

The TILLS is a test for the assessment of oral and written language abilities in students 6–18 years of age. Published in 2016 (Nelson, Plante, Helm-Estabrooks, & Hotz, 2016), it is unique in the way that it is aimed to thoroughly assess skills such as reading fluency, reading comprehension, phonological awareness, spelling, as well as writing in school-age children. The test is originally developed in English. Turkish Dyslexia Association has translated and adapted it to Turkish. This test has been used for diagnosing learning disabilities. For 6-7 years old children, a TILLS descriptive score of less than 24 indicates learning disability with 84% sensitivity and 84% specificity. For 8-11 years old children, a TILLS descriptive score of less than 34 indicates learning disability with 88% sensitivity and 85% specificity.

The TILLS test has 2 dimensions (language and modality). For listening modality, it has (1) Vocabulary awareness, (2) Phonemic awareness, (6) Listening comprehension, (8) Following directions; for speaking modality, it has (4) Nonword repetition, (3) Story retelling, (13) Social communication; for reading modality, it has (10) Nonword reading, (11) Reading fluency, (7) Reading comprehension; for writing modality, it has (5) Nonword spelling, (12a) Written expressions- Word score, (12b) Written expression -discourse score, (12c) Written Expression – sentence combining score; for Memory, (14) Digit span forward, (15) Digit span Backward, (9) Delayed story retelling subtests. The TILLS descriptive point is the sum of all subtests' scores.

2.3 Electroencephalography (EEG) recording

The EPOC-X headsets from EMOTIV were utilized throughout the tests. The headset's internal sampling rate for each channel was 2048 samples per second. Prior to down-sampling, the principal artifacts and alias frequencies were eliminated from the EEG data and each channel was sampled at 128 times per second. Two additional channels were employed as controls, and a total of 128 samples per second were applied to each of the 14 EEG channels. To ensure that each electrode could deliver high-quality EEG data, the EMOTIV EPOC-X headset was calibrated on the participants' scalps using the EMOTIV APP mobile applications prior to the training. The 14-channel EEG data were acquired in the theta (4–8 Hz), alpha (8–12 Hz), beta-1 (12–16 Hz), beta-2 (16–25 Hz), and gamma (25–45 Hz) bands for all analyses in this work. The EMOTIV headset interfaces lacked data from the delta (0–4 Hz) band. It has been shown that EMOTIV EPOC-X delivers high-quality QEEG data (Badcock et al, 2013). There are 70 features in the dataset. On the Auto Brain Train device, electrodes are used to collect the data (AF3, F3, F7, FC5, T7, P7, O1, O2, P8, T8, FC6, F8, F4, AF4).

2.4. Neurofeedback treatment protocol and multi-sensory learning method

Neurofeedback and multi-sensory learning techniques are employed in the Auto Train Brain smartphone app. The EMOTIV EPOC-X headsets are compatible with it. It is a non-invasive technique that helps people of all ages gradually improve their brain function. The only side effect that was occasionally seen was a headache. Traditional neurofeedback training may have surprising results when applied to the wrong brain regions or when the wrong neurofeedback protocols are used for the subject's condition. Before each neurofeedback session, Auto Train Brain's AI-assisted algorithms analyze the 2-minute resting state eyes-open QEEG data, and based on the results, the optimum neurofeedback protocol

is offered for the subsequent session. 1700 healthy people aged 4 to 80 gave norm data to Auto Train Brain.

This norm data is used to compare the subject's QEEG readings to those of TDC throughout the neurofeedback session. Neurofeedback training is carried out on the brain regions that are the most dysfunctional. In prior studies and clinical trials for kids, Auto Train Brain was found to be useful with the intention of minimizing adverse effects (Eroğlu et al., 2021). It provides online, real-time visual and aural neurofeedback while reading and interpreting QEEG signals from 14 channels. Within this software, a system and approach for increasing learning capacity are provided. We believed the following strategy might be helpful because dyslexia is described in the literature as a disconnection disorder in the left temporal area. On dyslexic children, this neurofeedback technique has already been tried in a clinical setting.

- Reduce absolute theta waves at FC5 if they are higher than the age-matched norm theta; and/or,
- Reduce absolute theta waves at T7- P7- O1 if they are higher than the age-matched norm theta; and/or, • Find the channels with the highest absolute power of theta waves in the left hemisphere and reduce absolute theta; and/or,
- Find the channels with the highest absolute power of theta waves in the right hemisphere and reduce absolute theta.

A green arrow on the screen denotes positive reinforcement, whereas a red arrow and a "beep" sound denote unfavorable comments. When a favorable reward is received, the score that is visible on the screen increases. If the participant's absolute theta levels are over the norm, a red arrow is shown on the screen, and they are encouraged to try to change it to a green one. Neurofeedback sessions typically last for 30 minutes. After the neurofeedback session, a phoneme-grapheme matching alphabet teaching approach is provided. The fact that the Auto Train Brain incorporates neurofeedback with multi-sensory learning ideas sets it apart from previous neurofeedback systems. A distinctive technique is the 14-channel neurofeedback system. According to Eroğlu et al. (2021), utilizing Auto Train Brain to lower theta reduces the dyslexia disconnection syndrome and improves the left lateralization of the brain. The results of the TILLS test show that reducing QEEG theta band power practically improves coherence across multiple brain regions.

2.5. Study Design

An app for mobile devices that may be used at home is the Auto Train Brain solution. The QEEG assessment electrodes were implanted while each participant sat in a chair. The participant was placed 0.5 meters away from the phone's display. Before each neurofeedback session, the resting-state QEEG readings were obtained for two minutes using the EMOTIV EPOC-X and the Auto Train Brain app. The individual was asked to relax and keep their eyes open while doing the QEEG assessment. This study is designed as a repeated measurement.

2.6. Statistical Analysis

The QEEG band power data was collected from 14 electrode channels during each session and averaged over 2 minutes. Z-scores were calculated for each QEEG band power using the equation $z = (x-m)/s$, where m and s represent the mean and standard deviation of the sample, respectively. Outliers (values greater than 5 or less than -5) were removed and missing values were replaced with the mean of the featured data. The data was labeled by a computer scientist based on

a psychiatrist's diagnosis of the participant, which did not change throughout the neurofeedback sessions. The data was balanced and pre-processed, and a binary classification using a supervised machine learning model was applied. The model, which was an artificial neural network, output a probability score for dyslexia positivity. The model was chosen based on the best performance among various hidden layers and activation functions, and was evaluated using the k-fold cross-validation technique with ten folds. To prevent overfitting, dropout was applied between layers. The results were also tested with an external set of diverse input data. Python, Google Collab, Sci-kit Learn, and TensorFlow machine learning libraries were used for statistical and data analysis, including k-folding, cross-validation, and the creation of confusion matrices. The Mat plot library was used to plot learning, validation, and ROC curves.

2.7. Experimental Methodology

Figure 1 summarizes the methodology that we used to reach our results. After collecting the dataset and validating it, the QEEG records for the first 20 sessions are used in training the ML models after splitting it into training and testing according to the ratio 80:20. After the training process, testing the trained ML models was conducted. In order to check the effectiveness of the neurofeedback treatment protocol, a validation process was conducted using the QEEG records of children with dyslexia in the sessions from 21 to 100. Lastly, the results were analyzed.

To have better knowledge of the effectiveness of using specific parts of the brain in the classification process, the experiment is repeated for the following brain sections:

- Left Hemisphere (LH) represented by the electrodes: AF3, F3, FC5, T7, P7 and O1,
- Left Frontal (LF) represented by the electrodes: AF3, and F3,
- Left Temporal (LT) represented by the electrodes: T7, and FC5,
- Left Occipital (LO) represented by the electrodes: O1, and P7,
- Right Hemisphere (RH) represented by the electrodes: AF4, F4, FC6, T8, P8, and O2,
- Right Frontal (RF) represented by the electrodes: AF4, and F4,
- Right Temporal (RT) represented by the electrodes: T8, and FC6,
- Right Occipital (RO) represented by the electrodes: O2, and P8.

The used Supervised ML models in the conducted training are: Logistic Regression (LR), Random Forests (RFs), Support Vector Machines (SVMs), Gradient Boosting (GB), Light Gradient Boosting Machine (LGBM), and lastly hard and soft voting ensemble classifiers used a combination of the separately mentioned classifiers except for RF classifier since it had a negative effect on the learning process. Please refer to Appendix I for more details about the used parameters.

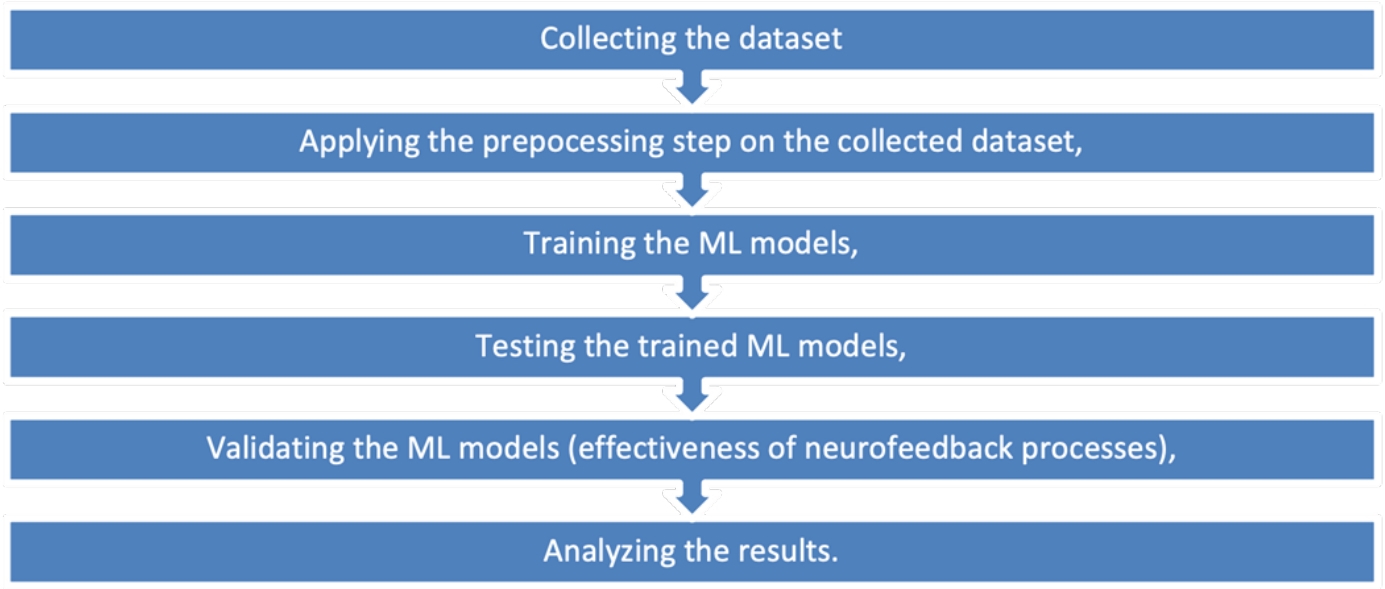


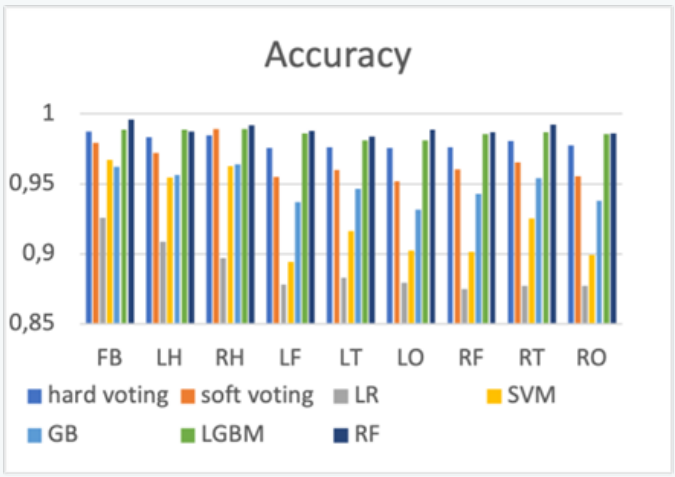
Figure 1. described below shows the methodology that we use

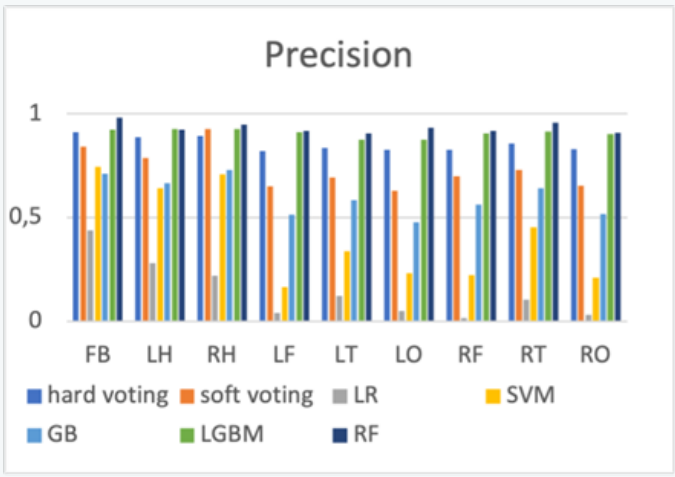
3. Results

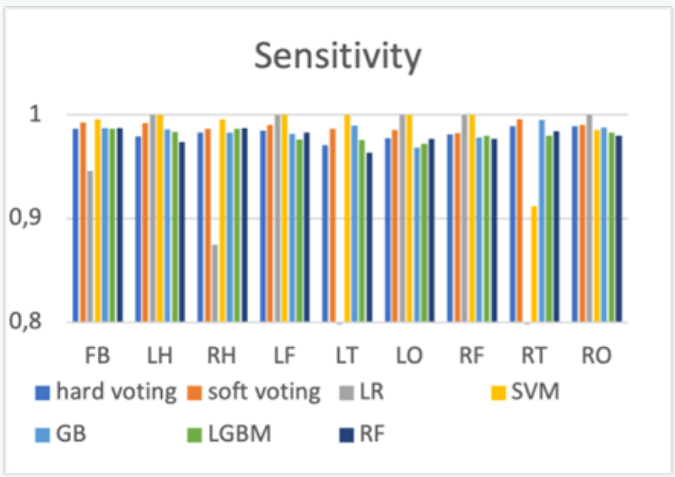
According to the results, 30% of the sessions for children with dyslexia were classified as electrophysiologically normal using the dyslexia biomarker detection algorithm. Additionally, 61% of the children with dyslexia were classified as electrophysiologically normal in at least one session after the 20th session of neurofeedback.

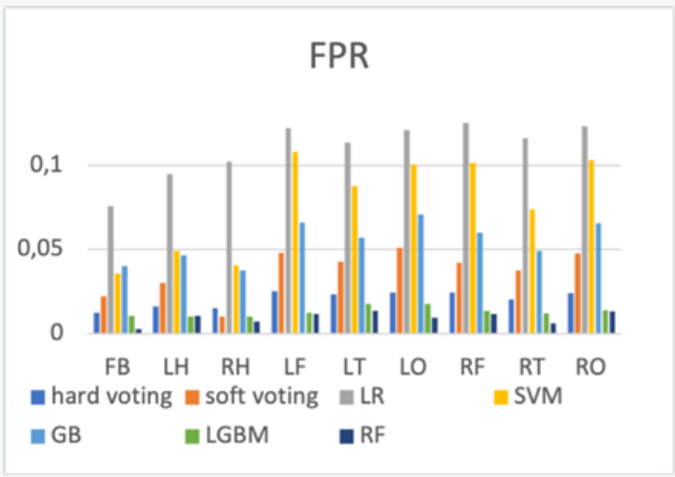
In this research we have created supervised ML algorithms for dyslexia biomarker detection. All the results in this section are rounded to 4 digits after the floating point. Table 1 shows the results of the classifying process. The mentioned results in Table 1 are based on the ML evaluation measurements: accuracy, sensitivity, specificity, precision, False Positive Rate (FPR) and Mean Squared Error (MSE). Table 2 shows the implementation time details in seconds.

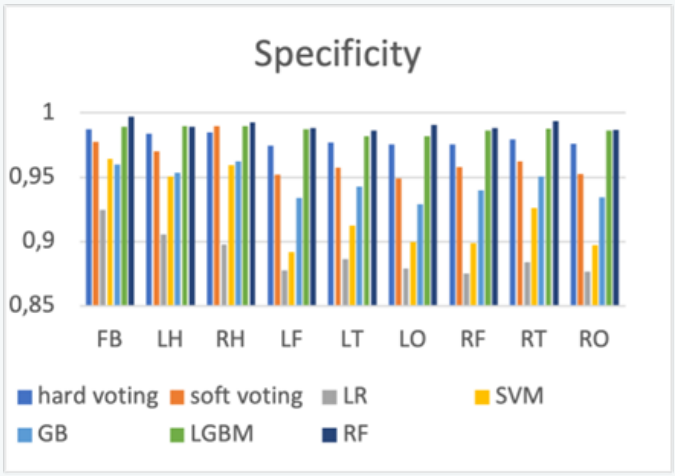
Table 1. Classification Results.











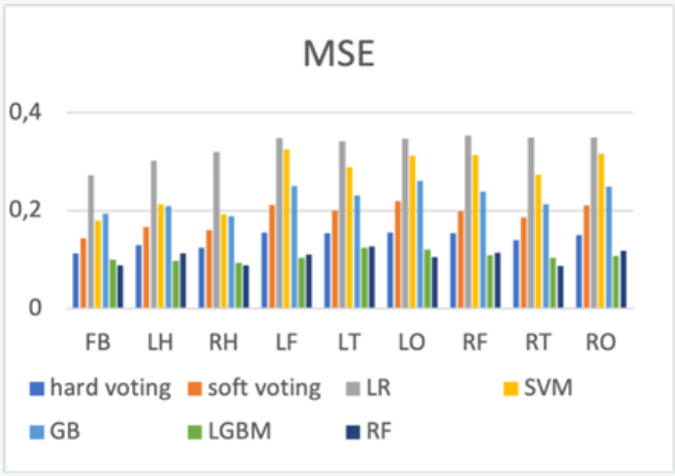
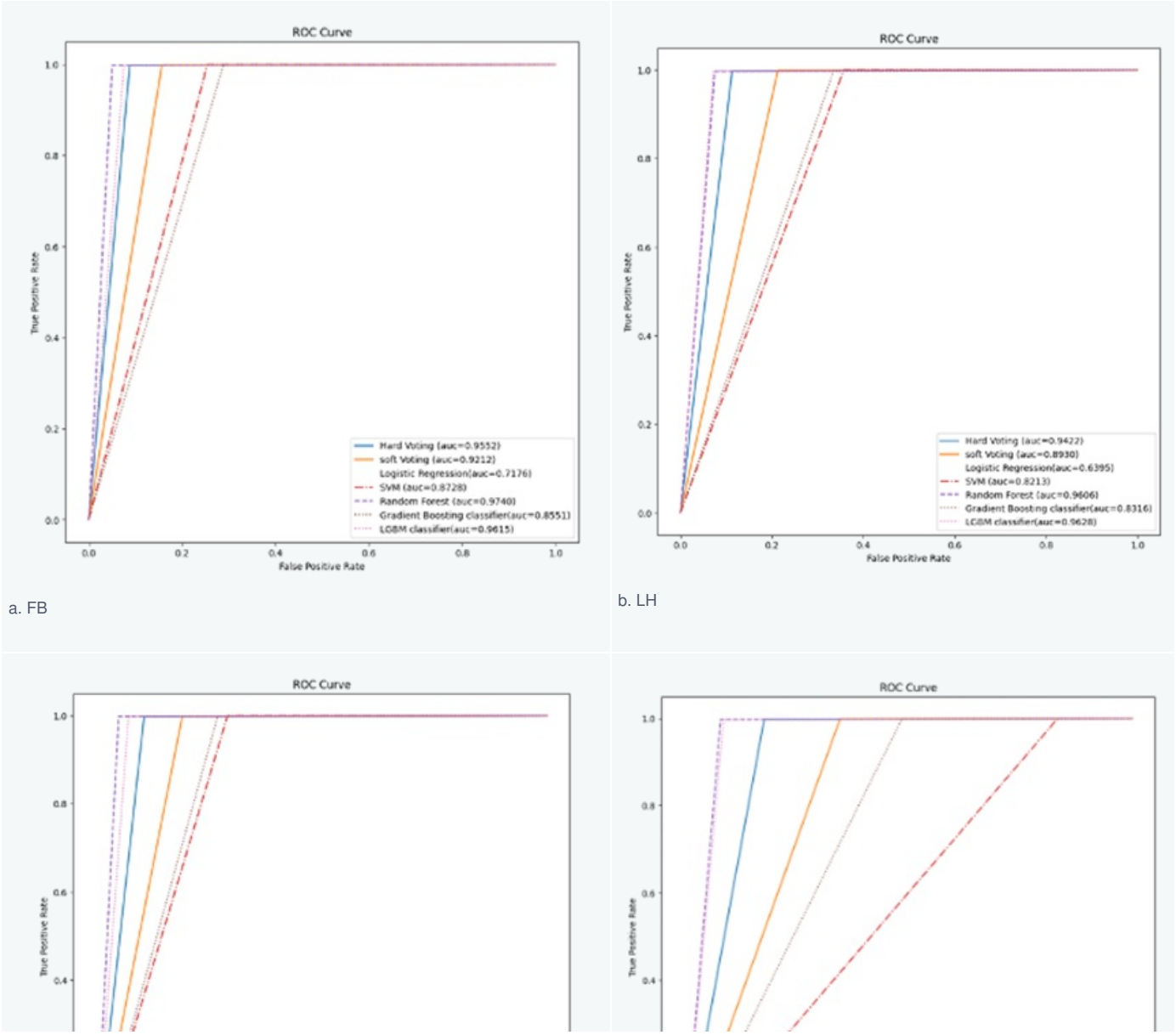


Table 2 Implementation time.

	FB	LH	RH	LF	LT	LO	RF	RT	RO
hard voting	5.5871	2.767	2.6854	1.4475	1.6323	1.492	1.4643	1.4356	2.3205
soft voting	5.6782	2.5382	2.5947	1.3643	1.4542	1.3681	1.3608	1.4852	2.9106
LR	0.4875	0.4243	0.3703	0.1563	0.1528	0.1872	0.1556	0.192	0.1374
SVM	2.2846	2.2941	2.1608	2.473	2.1793	2.6697	2.4271	1.9177	3.3719
GB	0.362	0.2812	0.279	0.275	0.3259	0.269	0.2738	0.2597	0.4214
LGBM	2.5294	1.0677	1.1097	0.4787	0.4889	0.4574	0.4854	0.4888	1.4723
RF	0.6788	0.538	0.5108	0.4314	0.4063	0.4831	0.4028	0.3844	0.45

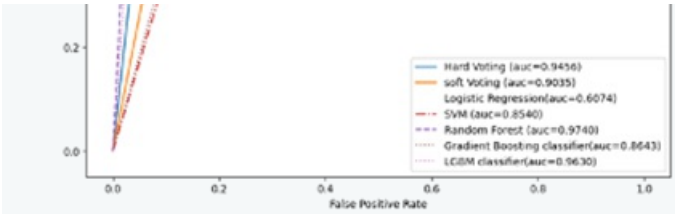
Table 3 contains the ROC curves and AUC values for the used ML models.

Table3. ROC Curves and AUC for the Used Supervised ML.

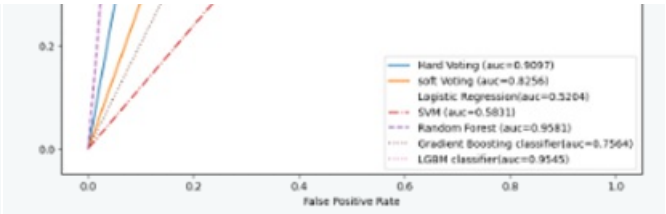


a. FB

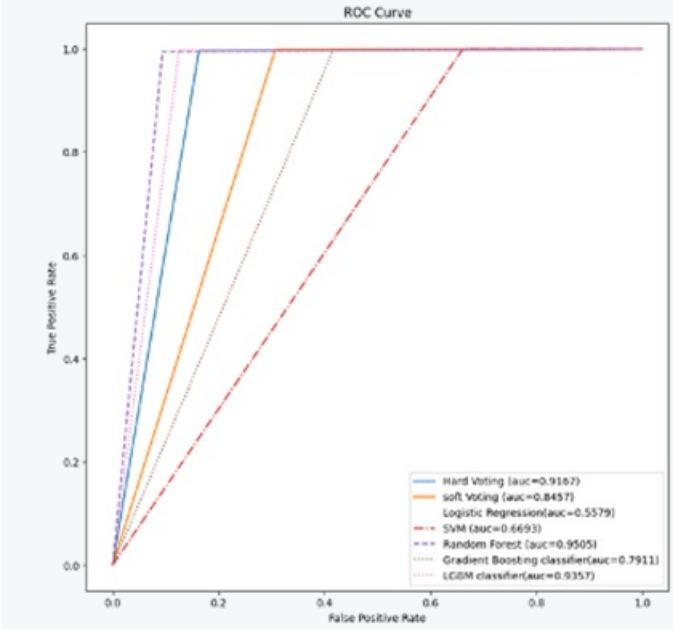
b. LH



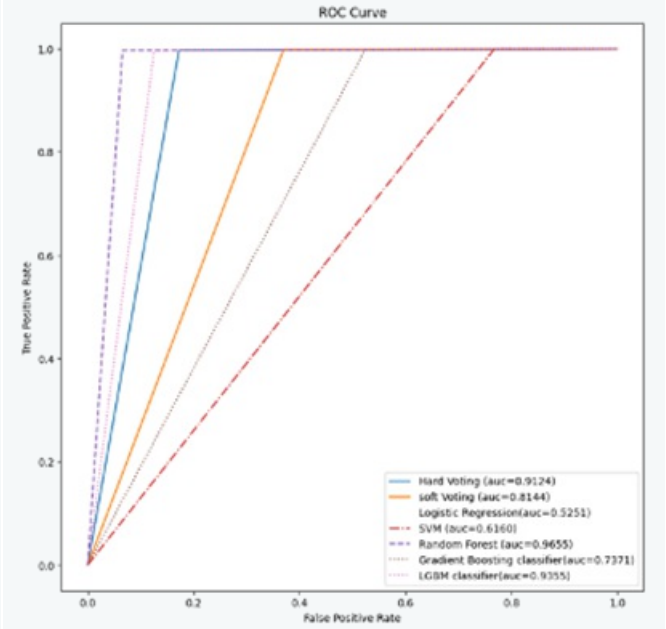
c. RH



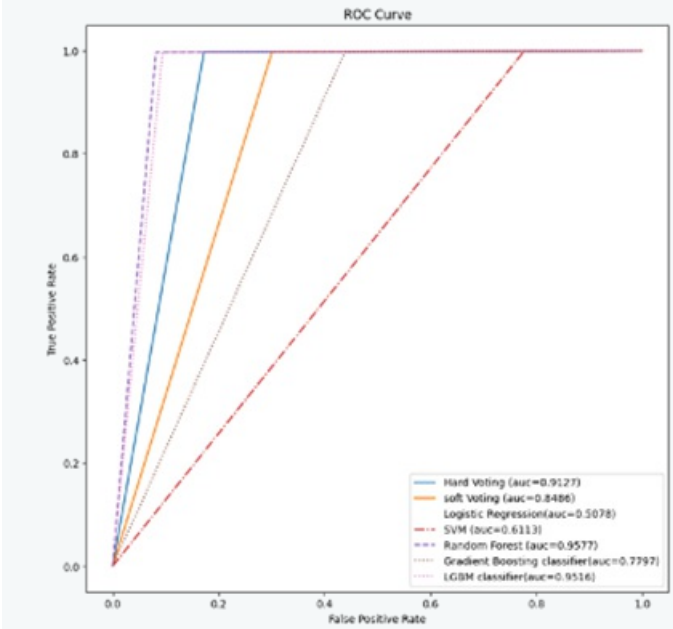
d. LF



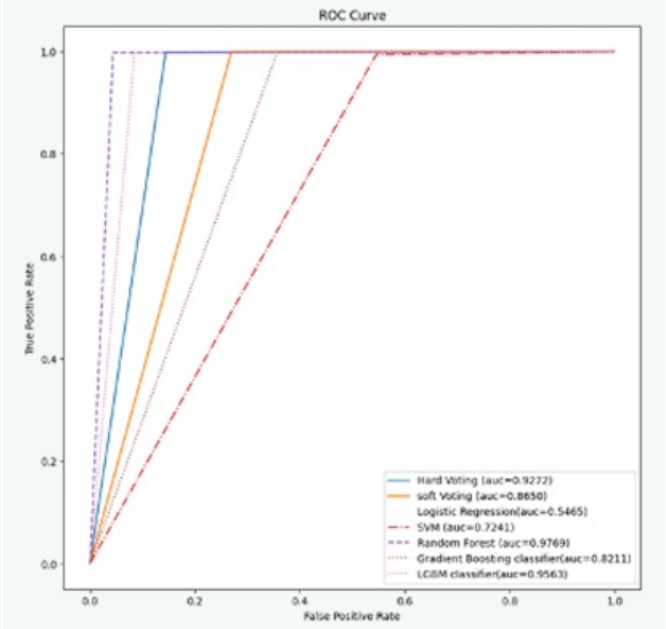
e. LT



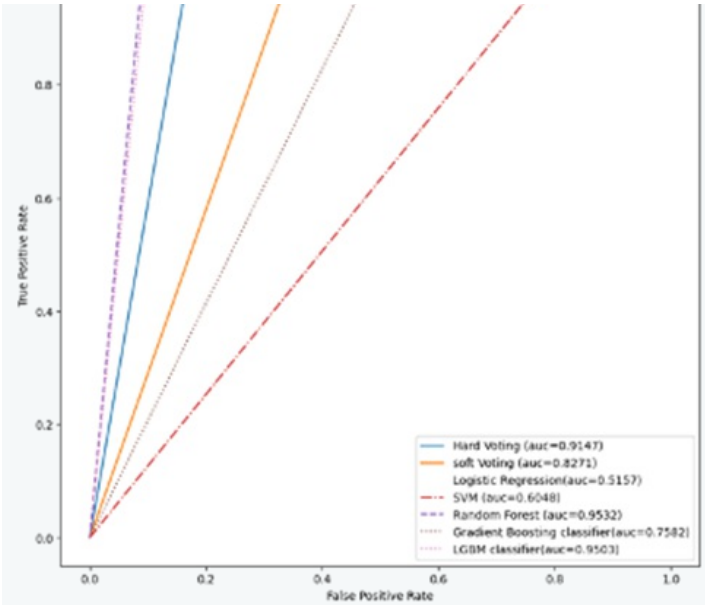
f. LO



g. RF



h. RT



i. RO

For the validation results, the bar chart in Figure 2 shows the percentage of the dyslexic cases that were classified as TDC after the neurofeedback procedure from the sessions of 21 to 100.

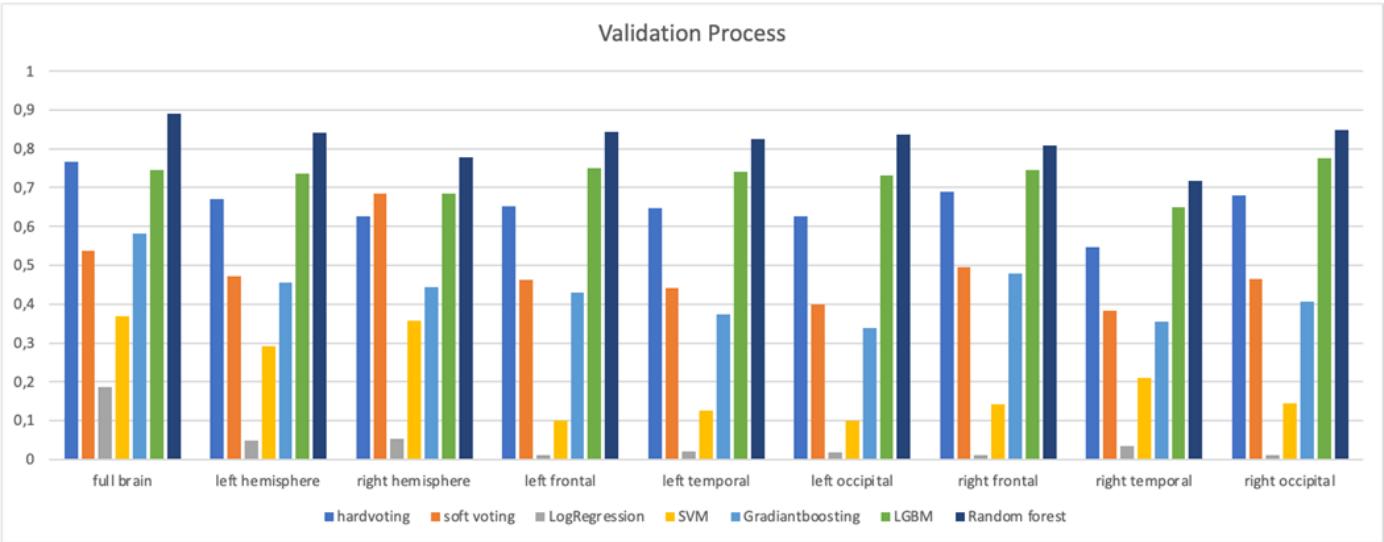


Figure 2. The Validation Process of the Supervised ML Models.

4. Discussion

The novelty of this research is that the dyslexia biomarker detection software was used to assess the efficacy of 14-

channel neurofeedback with Auto Train Brain.

Prior research has demonstrated the efficacy of 14-channel neurofeedback with Auto Train Brain in improving reading comprehension in individuals with dyslexia. In a clinical trial, the experimental group showed a significant increase in reading speed from 38 to 65 and an increase in TILLS descriptive point from 20.25 to 24.12 after 60 sessions of neurofeedback (Eroglu et al., 2020). Statistical analysis revealed that the simple effect of time on reading comprehension subtests was significant for the experimental group [F (1, 14) = 4.98, p=.042], and post-hoc tests showed that neurofeedback training was statistically more effective at improving reading comprehension than special education. The experimental group experienced a 70% improvement in reading comprehension, going from an average score of 3.06 (SD = 4.22) to 5.20 (SD = 4.41), while the control group regressed by 10%, going from an average score of 7.12 (SD = 3.18) to 6.36 (SD = 4.22) (Eroglu *et al*., 2020).

Neurofeedback has been shown to be effective in reducing the symptoms of dyslexia through the use of EEG data displayed in real-time to the patient, allowing them to gain better mental control through operant conditioning (Ninaus et al., 2015). Research has also suggested that neurofeedback can enhance brain structure and improve cognitive function (Wing, 2001; Canolty et al., 2006; Lubar et al., 1976; Mayoral Rodrigez et al., 2022; Terrasa et al., 2020). Neurofeedback has been recognized as a "possibly efficacious" treatment by the American Psychological Association (Melnikov, 2021). While clinical trials have often been used to demonstrate the effectiveness of neurofeedback, it can be challenging to demonstrate its usefulness. Using fMRI, it is possible to show the changes in brain activity that occur after neurofeedback, although it is more difficult to do so using QEEG. In one study, an hour of neurofeedback training was associated with enhanced fractional anisotropy (FA) in the corpus callosum and improved functional connectivity in the sensorimotor resting state network, as well as increased functioning in the default network configuration (Marins et al., 2018).

In this study, we developed a dyslexia biomarker detection software with high accuracy for classifying children with dyslexia and typical developmental control (TDC) using QEEG data collected during the resting state in the latter neurofeedback sessions. The results showed that 30% of the sessions of children with dyslexia were classified as electrophysiologically normal or TDC, and 61% of the children with dyslexia were classified as electrophysiologically normal or TDC for at least one session after the 20th session of neurofeedback. These findings, obtained after 100 sessions of neurofeedback, are promising and align with the results of the clinical trial. In the future, we plan to further develop machine learning models for more fine-grained classification of dyslexia severity, rather than the current binary classification of dyslexia and TDC. This will allow us to assess the extent to which neurofeedback can improve dyslexia symptoms in individuals who are still classified as having dyslexia after training.

4.1. Limitations of the Study

The first limitation of the study, we used repeated measures from the same subjects to collect QEEG data from each session. The second limitation of the study is the possibility of placebo effects. According to Gaab et al. (2019), children who receive individualized attention and specialized therapies may improve their functioning primarily as a result of the social and environmental effects of those interventions. The maturation effects are the third limitation of the study. All

children's brains change significantly as they grow. Therefore, it is anticipated that maturation will have an impact on QEEG changes over the next six months.

5. Conclusions

In conclusion, the use of machine learning techniques in the diagnosis and training of dyslexia has shown promise in this study. The development of a classification system using QEEG data and ANN was successful in accurately diagnosing dyslexia, and the application of neurofeedback as a training showed potential for improving symptoms in individuals with dyslexia. These findings suggest that machine learning may be a valuable tool in the dyslexia training process, offering a new and innovative approach to addressing this learning disorder. However, further research is needed to fully understand the potential of these techniques and to determine their long-term effectiveness in improving outcomes for individuals with dyslexia. Despite this, the use of machine learning in the field of dyslexia research and treatment is an exciting development that has the potential to make a significant impact on the lives of those affected by dyslexia.

Declarations

Ethics approval and consent to participate

All of the participants provided their informed consent after the research ethics committee explained the experimental procedure to them, the Yeditepe University ethics committee approved the study protocol, and the clinical trial was registered with the Turkey Pharmaceuticals and Medical Devices Agency (Nbr: 71146310-511.06,2.11.2018).

Availability of data and material

The corresponding author will provide the datasets produced and/or analyzed during the current work upon reasonable request.

Code availability

None

Competing interests

None

Funding

None

Authors' contributions

The title, abstract, introduction, materials and techniques, findings, and discussion of the main manuscript text were all written by G.E. who also created the charts and tables.
M.R.A.H. enhanced the introduction in relation to the ML techniques, tools, and approaches and written results section.
M.R.A.H. created the tables and graphics.

Acknowledgments

Without their dedication and assistance, we might not have been able to complete the study, thus we would like to thank the families who participated in the experiments.

References

- Alex, F., & Larry, M. M. (2018). Features and machine learning for correlating and classifying between brain areas and dyslexia. arXiv preprint arXiv:1812.10622.
- American Psychiatric Association. (2013). Anxiety disorders. In Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780896425596.dsm05
- Arns, M., Peters, S., Breteler, R., & Verhoeven, L. (2007). Different brain activation patterns in dyslexic children: Evidence from EEG power and coherence patterns for the double-deficit theory of dyslexia. Journal of Integrative Neuroscience, 6(1), 175-190.
- Badcock, N. A., Mousikou, P., Mahajan, Y., De Lissa, P., Thie, J., & McArthur, G. (2013). Validation of the Emotiv EPOC® EEG gaming system for measuring research quality auditory ERPs. PeerJ, 1, e38.
- Bednarek, D. B., Saldaña, D., Quintero-Gallego, E., García, I., Grabowska, A., & Gómez, C. M. (2004). Attentional deficit in dyslexia: A general or specific impairment? Neuroreport, 15(11), 1787-1790.
- Bellocchi, S., Muneaux, M., Bastien-Toniazzo, M., & Ducrot, S. (2013). I can read it in your eyes: What eye movements tell us about visuo-attentional processes in developmental dyslexia. Research in Developmental Disabilities, 34(1), 452-460.
- Benfatto, M., Öqvist Seimyr, G., Ygge, J., Pansell, T., Rydberg, A., & Jacobson, C. (2016). Screening for dyslexia using eye tracking during reading. PloS one, 11(12), e0165508.
- Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Berger, M. S., Barbaro, N. M., & Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. Science, 313(5793), 1626-1628. doi:10.1126/science.1128115
- Dunn, L. M., & Dunn, L. M. (1965). Peabody picture vocabulary test.

- D'Souza, S., Backhouse-Smith, A., Thompson, J. M., Slykerman, R., Marlow, G., Wall, C., ... & Waldie, K. E. (2016). Associations between the KIAA0319 dyslexia susceptibility gene variants, antenatal maternal stress, and reading ability in a longitudinal birth cohort. Dyslexia, 22(4), 379-393.
- Eden, G. F., VanMeter, J. W., Rumsey, J. M., Maisog, J. M., Woods, R. P., & Zeffiro, T. A. (1996). Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. Nature, 382(6586), 66
- Eroğlu, G., Gürkan, M., Teber, S., Ertürk, K., Kırmızı, M., Ekici, B., ... & Çetin, M. (2020). Changes in EEG complexity with neurofeedback and multi-sensory learning in children with dyslexia: A multiscale entropy analysis. Applied Neuropsychology: Child, 1-12.
- Eroğlu, G., Teber, S., Ertürk, K., Kırmızı, M., Ekici, B., Arman, F., ... & Çetin, M. (2021). A mobile app that uses neurofeedback and multi-sensory learning methods improves reading abilities in dyslexia: A pilot study. Applied Neuropsychology: Child, 1-11.
- Eroğlu, G. (2022). Auto Train Brain increases the variance of the gamma band sample entropy in the left hemisphere in dyslexia: A pilot study. In WORLD S4 2022 (Tam Metin Bildiri/Sözlü Sunum) (Yayın No: 7773551).
- Eroğlu, G., & Arman, F. (2022). k-Means clustering by using the calculated Z-scores from QEEG data of children with dyslexia. Applied Neuropsychology: Child, 1-7.
- Estes ML, McAllister AK (2016). Maternal immune activation: Implications for neuropsychiatric disorders. Science. 353:772–777.
- Francks, C., MacPhie, I. L., & Monaco, A. P. (2002). The genetic basis of dyslexia. The Lancet Neurology, 1(8), 483-490.
- Fraga González, G., Smit, D. J., Van der Molen, M. J., Tijms, J., Stam, C. J., De Geus, E. J., & Van der Molen, M. W. (2018). EEG resting state functional connectivity in adult dyslexics using phase lag index and graph analysis. Frontiers in Human Neuroscience, 12, 341.
- Formoso, M. A., Ortiz, A., Martinez-Murcia, F. J., Gallego, N., & Luque, J. L. (2021). Detecting phase-synchrony connectivity anomalies in EEG signals. Application to dyslexia diagnosis. Sensors, 21(21), 7061. https://doi.org/10.3390/s21217061
- Formoso, M. A., Ortiz, A., Martínez-Murcia, F. J., Gallego-Molina, N., & Luque, J. L. (2021). Modelling brain connectivity networks by graph embedding for dyslexia diagnosis.
- Gaab, J., Kossowsky, J., Ehlert, U., & Locher, C. (2019). Effects and components of placebos with a psychological treatment rationale-three randomized-controlled studies. Scientific Reports, 9(1), 1-8.
- Gallego-Molina, N. J., Ortiz, A., Martínez-Murcia, F. J., Formoso, M. A., & Giménez, A. (2022). Complex network modeling of EEG band coupling in dyslexia: An exploratory analysis of auditory processing and diagnosis. Knowledge-Based Systems, 240, 108098. https://doi.org/10.1016/j.knosys.2021.108098
- García Chimeno, Y., Garcia Zapirain, B., Saralegui Prieto, I., & Fernandez-Ruanova, B. (2014). Automatic classification of dyslexic children by applying machine learning to fMRI images. Bio-medical Materials and Engineering, 24(6), 2995-3002.
- Geschwind, N., & Galaburda, A. M. (1985). Cerebral lateralization: Biological mechanisms, associations, and pathology: III. A hypothesis and a program for research. Archives of Neurology, 42(7), 634-654.

- Horwitz, B., Rumsey, J. M., & Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. Proceedings of the National Academy of Sciences, 95(15), 8939-8944.
- Jiang NM, Cowan M, Moonah SN, Petri WA (2018). The impact of systemic inflammation on neurodevelopment. Trends Mol Med. 24:794–804.
- Karim, I., Abdul, W., & Kamaruddin, N. (2013, March). 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) (pp. 1-5). IEEE.
- Kaufman, A. S. (1994). Intelligent testing with the WISC-III. John Wiley & Sons.
- Kershner, J. R. (2019). Neurobiological systems in dyslexia. Trends in Neuroscience and Education, 14, 11-24.
- Kershner, J. R. (2020). Neuroscience and education: cerebral lateralization of networks and oscillations in dyslexia. Laterality, 25(1), 109-125.
- Kershner, J. R. (2020). Dyslexia as an adaptation to cortico-limbic stress system reactivity. Neurobiology of Stress, 12, 100223.
- Kershner, J. R. (2021). Multisensory deficits in dyslexia may result from a locus coeruleus attentional network dysfunction. Neuropsychologia, 161, 108023.
- Kheyrkhah Shali, R., & Setarehdan, S. K. (2020). The effective brain areas in recognition of dyslexia. International Clinical Neuroscience Journal, 7(3), 147-152. https://doi.org/10.34172/icnj.2020.16
- Klimesch, W., Doppelmayr, M., Wimmer, H., Gruber, W., Röhm, D., Schwaiger, J., & Hutzler, F. (2001). Alpha and beta band power changes in normal and dyslexic children. Clinical Neurophysiology, 112(7), 1186-1195.
- Klimesch, W., Doppelmayr, M., Yonelinas, A., Kroll, N. E., Lazzara, M., Röhm, D., & Gruber, W. (2001). Theta synchronization during episodic retrieval: neural correlates of conscious awareness. Cognitive Brain Research, 12(1), 33-38.
- Kraus, N. (2012). Atypical brain oscillations: a biological basis for dyslexia. Trends in Cognitive Sciences, 16(1), 12-13.
- Kuban KC, O'Shea TM, Allred EN, Fichorova RN, Heeren T, Paneth N, Hirtz D, Dammann O, Leviton A; ELGAN Study Investigators (2015). The breadth and type of systemic inflammation and the risk of adverse neurological outcomes in extremely low gestation newborns. Pediatr Neurol. 52(1):42-48.
- Lassus-Sangosse, D., N'guyen-Morel, M. A., & Valdois, S. (2008). Sequential or simultaneous visual processing deficit in developmental dyslexia? Vision Research, 48(8), 979-988.
- Lubar, J. F., & Shouse, M. N. (1976). EEG and behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR). Biofeedback and Self-Regulation, 1, 293-306. https://doi.org/10.1007/BF01001170
- Marins, T., Rodrigues, E. C., Bortolini, T., Melo, B., Moll, J., & Tovar-Moll, F. (2019). Structural and functional connectivity changes in response to short-term neurofeedback training with motor.
- McDougall, S., Hulme, C., Ellis, A., & Monk, A. (1994). Learning to read: The role of short-term memory and phonological skills. Journal of Experimental Child Psychology, 58(1), 112-133.
- Melnikov, M. Y. (2021). The current evidence levels for biofeedback and neurofeedback interventions in treating depression: A narrative review. Neural Plasticity.
- Nelson, N. W., Plante, E., Helm-Estabrooks, N., & Hotz, G. (2016). Test of integrated language and literacy skills

(TILLS). Baltimore, MD: Brookes.

- Ninaus, M., Witte, M., Kober, S. E., Friedrich, E. V., Kurzmann, J., Hartsuiker, E., Neuper, C., & Wood, G. (2015). Neurofeedback and serious games. In M. Management Association (Ed.), Gamification: Concepts, methodologies, tools, and applications (pp. 83-112). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-8200-9.ch005
- Niv, S. (2013). Clinical efficacy and potential mechanisms of neurofeedback. Personality and Individual Differences, 54(6), 676-686. doi:10.1016/j.paid.2012.11.037
- Ortiz, A., Martinez-Murcia, F. J., Luque, J. L., Giménez, A., Morales-Ortega, R., & Ortega, J. (2020). Dyslexia diagnosis by EEG temporal and spectral descriptors: An anomaly detection approach. International Journal of Neural Systems, 30(07), 2050029. https://doi.org/10.1142/S012906572050029X
- Rezvani, Z., Zare, M., Žarić, G., Bonte, M., Tijms, J., van der Molen, M. W., & Raga González, G. (2019). Machine learning classification of dyslexic children based on EEG local network features. BioRxiv, 569996.
- Richardson, A. J. (2004). Long-chain polyunsaturated fatty acids in childhood developmental and psychiatric disorders. Lipids, 39(12), 1215-1222.
- Rippon, G., & Brunswick, N. (2000). Trait and state EEG indices of information processing in developmental dyslexia. International Journal of Psychophysiology, 36(3), 251-265.
- Rubinsten, O., & Henik, A. (2006). Double dissociation of functions in developmental dyslexia and dyscalculia. Journal of Educational Psychology, 98(4), 854.
- Schulte-Körne, G., & Bruder, J. (2010). Clinical neurophysiology of visual and auditory processing in dyslexia: a review. Clinical neurophysiology, 121(11), 1794-1809.
- Siegel, L. S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. Canadian Journal of Psychology/Revue canadienne de psychologie, 42(2), 201.
- Terrasa, J. L., Barros-Loscertales, A., Montoya, P., & Muñoz, M. A. (2020). Self-regulation of SMR power led to an enhancement of functional connectivity of somatomotor cortices in fibromyalgia patients. Frontiers in Neuroscience, 14, 236. https://doi.org/10.3389/fnins.2020.00236
- Thornton, K. E., & Carmody, D. P. (2005). Electroencephalogram biofeedback for reading disability and traumatic brain injury. Child and Adolescent Psychiatric Clinics, 14(1), 137-162.
- Van Bergen, E., De Jong, P. F., Plakas, A., Maassen, B., & van der Leij, A. (2012). Child and parental literacy levels within families with a history of dyslexia. Journal of Child Psychology and Psychiatry, 53(1), 28-36.
- Wechsler, D. (1999). Wechsler abbreviated scale of intelligence.
- Yektas, C., Tufan, A. E., & Kilicaslan, O. (2022). Yazıcı M, Karakaya Kaplan SE, Sarıgedik E. Elevated Monocyte Levels Maybe a Common Peripheral Inflammatory Marker in Specific Learning Disorders and Attention Deficit/Hyperactivity Disorder. Psychiatry and Behavioral Sciences, 12(3), 125-133.
- Yılmaz, S., & Akyüz, F. (2021). The relationship between speech difficulties and brain laterality in Attention Deficit Hyperactivity Disorder and Specific Learning Disorder. Acta Medica Alanya, 5(3), 250-256.
- Zainuddin, A., & Y., Khuan, L., Mansor, W., & Mahmoodin, Z. (2016). Optimized KNN classify rule for EEG based differentiation between capable dyslexic and normal children. 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), 685–688.

- Zainuddin, A., Mansor, W., Y. Lee, K., & Mahmoodin, Z. (2018). Performance of Support Vector Machine in classifying EEG signal of dyslexic children using RBF kernel. Indonesian Journal of Electrical Engineering and Computer Science, 9(2), 403. https://doi.org/10.11591/ijeecs.v9.i2.pp403-409
- Zainuddin, A. Z. A., Mansor, W., Lee, K. Y., & Mahmoodin, Z. (2019). Comparison of Extreme Learning Machine and K-Nearest Neighbour performance in classifying EEG signal of normal, poor, and capable dyslexic children. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 4513–4516. https://doi.org/10.1109/EMBC.2019.8857569