Dyslexia biomarker detection with Quantitative electroencephalography (QEEG) data in children: Feasibility, Acceptability, Economic impact, Pilot Study and Survey Results

Gunet Eroglu, Mertali Köprülü[1], Berdan Karabacak[2]

1 Isik University

2 Neuro Brain Academy

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

Introduction

Dyslexia awareness is not high in societies. Although 5-17% of the population is thought to be dyslexic in every global society, only 0.05 % of the population gets a formal diagnosis in Turkey. The financial situation of families of children with dyslexia is not enough to get an early diagnosis and investigate intensive remedial opportunities. Mobile apps have been used to diagnose health conditions in recent years. Worldwide solutions for dyslexia do not exist as the solutions were language-dependent. We need an easily reachable, affordable, language-independent, and worldwide acceptable biomarker detection app for dyslexia at schools because it potentially helps to prevent severe consequences through an early diagnosis that helps provide early intervention. No other research assessed the feasibility and acceptability of using this mobile app to detect dyslexia at home or school. Here we present a dyslexia biomarker detection app based on Z-scored Quantitative Electroencephalogram (QEEG) data that can be used at home or school and has accomplished a high accuracy rate in diagnosing dyslexia. The mobile app can be used at home by parents or teachers at school.

Methods

We have collected data from 207 children (96 of them have dyslexia, 111 of them are typically developing) between 7-10 years old for 40 sessions on average. The data consists of the eyes-open resting state 2-minute QEEG data from 14-channels.

Results

Using the Artificial Neural Network (ANN) machine learning method, children with dyslexia/ typically developing children (TDC) classification has been done with a high accuracy rate (98.8%). ANN yields the highest accuracy results with standardized QEEG data in the literature. A survey is created to assess the dyslexia biomarker detection app's feasibility, acceptability, and economic impact. The results have shown that the biomarker detection app is found feasible and acceptable by families, however, it is found expensive to use at home as it includes the costly

electroencephalography (EEG) headset.

Conclusion

In order for this biomarker detection method to be used extensively at home, EEG headset prices should become more affordable or this dyslexia biomarker detection method can be used at schools.

Günet Eroğlu[a], **Mertali Köprülü**[b], **Berdan Karabacak**[c]

[a] *Computer Engineering Dept, Engineering and Nature Faculty, Bahçeşehir University, Istanbul, Turkey, gunet.eroglu@healthmobilesoftware.com (Corresponding author)*

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

[c] *Research Department, Neuro Brain Academy, Kocaeli, Turkey*

Keywords: Dyslexia detection, developmental dyslexia biomarkers, QEEG Dataset, Artificial-Neural Network.

1. Introduction

Dyslexia awareness by families and teachers at school is not high in societies (Madriaga, 2007). According to the Turkish Dyslexia Association, 41,600 people out of 85 million people in Turkey are diagnosed with dyslexia which is only 0.05% of the whole population (Tokar, 2018) although the number of people with dyslexia is thought to be at least 5-17% in every society (Ozernov-Palchik & Gaab, 2016; Wagner et al., 2021). Early diagnosis and early intensive intervention are crucial for dyslexia as brain plasticity is high during childhood (Habib & Giraud, 2013). The financial situation of families of children with dyslexia is not enough to get the proper diagnosis on time and investigate intensive remedial opportunities (Peterson & Pennington, 2015). Dyslexia (as defined in The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, DSM-V), or decoding difficulty, refers to children who have difficulty mastering the relationships between the spelling patterns of words and their pronunciations (American Psychiatric Association, 2013). These children typically read aloud inaccurately and slowly and experience additional problems with spelling (Snowling et al., 2012). Dyslexia is characterized by marked hypoactivation within the reading network, disrupted functional connectivity, and differences in structural connectivity in certain fiber tracts (Kuhl et al., 2020). It's characterized by problems with accurate and fluent word recognition and poor spelling skills (Snowling et al., 2012). Other dyslexia-related implications include difficulties with reading comprehension and a lack of reading experience, both of which can hinder the development of vocabulary and knowledge (Lyon, Shaywitz, & B.A. Shaywitz, 2003).

Reading is an acquired ability for humans (Dehaene, 2005). The left hemispheric lateralization should be completed by age 7 before the child starts school and learns how to read (Koenig et al., 1990). During this period, as the child can not properly read and write, his self-esteem would go down and his academic life starts to be affected by the late intervention. With the help of computers and machine learning, the diagnosis of dyslexia can be done earlier at school or at home, and remedial interventions may be taken much beforehand (Economou, M., 2022).

In the literature, machine learning methods have been utilized to diagnose dyslexia. These algorithms use functional magnetic resonance imaging (fMRI) scans (Janković, 2022), Electroencephalography (EEG) scans (Janković, 2022), Magnetoencephalography (MEG) scans (Iraola Goiburu, 2014), and eye tracking information (Janković, 2022). Psychometric tests take 1.5- 2 hours. The psychometric tests are the Wechsler Abbreviated Scale of Intelligence (WASI, II, III&IV) (Wechsler, 1999), Wide Range Achievement Test: Revision R,3&4 (WRAT-R,3&4) (Jastak & Wilkinson, 1984), Woodcock & Johnson III (WJ-III) (Schrank, 2011), Comprehensive Test of Phonological Processing (CTOPP) (Wagner et al.,1999), and Peabody Picture Vocabulary Test: Third Edition (PPVT-III) (Dunn & Dunn, 1965), Test of integrated language and literacy skills (TILLS test) (Nelson et al., 2016), and WISC-R tests (Kaufman, 1994). Although fMRI scans and eye tracking data yield higher accuracy results than other methods, they are expensive solutions. It is hard to use MRI scans and eye tracking methods to collect data from a 7-year-old child as they require expensive equipment. Although it is a cheaper way to collect data, the algorithms that use EEG scans do not yield a high accuracy result to go to the market in the literature.

A computer model known as an artificial neural network (ANN) is made up of many processing components that accept inputs and produce results in accordance with predetermined activation functions. An artificial neural network (ANN) is a non-parametric machine learning technique that focuses on identifying the discriminant function to identify dyslexia. The discriminant function is the function that divides the two data sets into two different subsets. It does not make any assumptions about the data set. The k-means algorithm finds the k number of centroids and then assigns each data point to the closest cluster, minimizing the size of the centroids. The centroids are the center points of the clusters. Finding the centroid is what "means" in the k-means algorithm indicates: averaging the data. A supervised machine learning approach called Support Vector Machine (SVM) is used for both classification and regression. Although we also refer to regression issues, categorization is the most appropriate term. Finding a hyperplane in an N-dimensional space that clearly classifies the data points is the goal of the SVM method. On EEG scan data sets, Al-Barhamtoshy and Motaweh (2017) used k-means, ANN, and fuzzy logic classifiers that classify the data sets into a spectrum instead of binary to reach an accuracy of 89.6%, 89.6%, and 85.7 percent, respectively. The three models were shown to have an overall accuracy of 81.1%, 62% precision, 100% recall, and 76.6% F-score. Precision is the amount of information that is conveyed by a value, whereas accuracy is the measure of the correctness of the value in correlation with the information. Recall or sensitivity is the fraction of the total amount of pertinent models that were retrieved. In statistical analysis of binary classification, the F-score or F-measure is a measure of a test's accuracy. Karim et al. (2013) employed a Multilayer Perceptron to detect dyslexia signs by collecting brain waves in the resting state. A multilayer perceptron (MLP) is a fully connected class of feedforward artificial neural networks (ANN). With eyes closed, the accuracy was 85%, and with eyes open, it was 86%. Using machine learning, Frid and Breznitz (2018) investigated the variations in ERP signals between proficient readers and readers with dyslexia. They used SVM, ANN, and Principal Component Analysis (PCA) to get a 78% accuracy. Principal component analysis (PCA) is a popular technique for analyzing large datasets containing a high number of dimensions/features per observation, increasing the interpretability of data while preserving the maximum amount of information, and enabling the visualization of multidimensional data. In the literature, the higher accuracy scores reached with ANN on fMRI scans and DTI scans are 94.8% (Chimeno et al., 2014). The higher accuracy results

with eye tracker data were 95.6% with SVM (Benfatto et al., 2016). The higher accuracy result for dyslexia detection with handwriting data was 77.6% (Spoon et al.,2019). The higher accuracy results with psychometric test scores were 99% (Khan et al., 2018).

This study tests the feasibility, acceptability, and economic impact of the dyslexia biomarker detection app that we have developed with ANN with a survey of families of children with dyslexia. This is the first research that investigated the use of dyslexia biomarker detection software in families who have children with dyslexia.

Our research questions are stated as follows:

- **Research question 1:** Are Z-scored QEEG recordings from 14-channel band power data of both children with dyslexia and TDC classified with a high accuracy rate (>90%) by ANN?
  - H0: Z-scored QEEG recordings from 14-channel band power data of both children with dyslexia and TDC are not classified with a high accuracy rate (>90%) by ANN.
  - H1: Z-scored QEEG recordings from 14-channel band power data of both children with dyslexia and TDC are classified with a high accuracy rate (>90%) by ANN.
- **Research question 2:** Is the dyslexia biomarker detection algorithm feasible to use as a mobile app?
  - H0: The dyslexia biomarker detection algorithm is not feasible to use as a mobile app.
  - H1: The dyslexia biomarker detection algorithm is feasible to use as a mobile app.
- **Research question 3:** Is the dyslexia biomarker detection algorithm acceptable to families?
  - H0: The dyslexia biomarker detection algorithm is not acceptable to families.
  - H1: The dyslexia biomarker detection algorithm is acceptable to families.
- **Research question 4:** Is using the dyslexia biomarker detection algorithm at home affordable including the EEG headset?
  - H0: Using the dyslexia biomarker detection algorithm at home is not affordable including the EEG headset.
  - H1: Using the dyslexia biomarker detection algorithm at home is affordable including the headset.
- **Research question 5:** Is using the dyslexia biomarker detection algorithm at school affordable including the EEG headset?
  - H0: Using the dyslexia biomarker detection algorithm at school is not affordable including the EEG headset.
  - H1: Using the dyslexia biomarker detection at school is affordable including the EEG headset.

2. Methods and materials

2.1. Participants

96 children with pure dyslexia (Mage= 8.85, SD = 1.56; 76 males, 20 females; the ethnic group is white) who used a mobile app module at home formed our experiment group. 111 children in the control group were developing typically (Mage= 8.80, SD = 1.60; 80 males, 31 females; the ethnic group is white). The participants are recruited randomly with social media ads. The children in the experimental group were diagnosed with dyslexia by doctors. Psychiatrists examine whether the individuals met the DSM-V dyslexia criteria. The children chosen to participate in the experiment were chosen at random with social media ads if they conform to the inclusion criteria. The inclusion criteria for children with dyslexia are not being on medication, being aged between 7-10, and having no comorbid situations like ADHD, or atypical autism. The participant's primary goal in the study is to use the software at home. Individuals were instructed to perform a 2-minute resting state EEG measurement for the purpose of data collection.

2.2. Materials

2.2.1. Electroencephalography

We have used EMOTIV's EPOC-X headsets in the experiments. The internal sampling rate of the headset was 2048 samples/s per channel. The EEG data is down sampled to 128 samples per second for each channel. Before the training, the EMOTIV EPOC-X headset was calibrated on children with dyslexia using EMOTIV APP mobile app. For all analyses in this work, theta (4–8 Hz), alpha (8–12 Hz), beta-1 (12–16 Hz), beta-2 (16–25 Hz), and gamma (25–45 Hz) bands of the 14-channel EEG data were recorded throughout the tests. Data from the delta (0–4 Hz) band was absent from the EMOTIV headset interfaces. It has been demonstrated that EMOTIV EPOC-X provides high-quality QEEG data (Badcock et al, 2013). In the dataset, 70 features (frequency band data; theta, alpha, beta1, beta2, gamma from 14 channels) indicate whether a person has dyslexia or not. The information is gathered using electrodes on the mobile app (AF3, F3, F7, FC5, T7, P7, O1, O2, P8, T8, FC6, F8, F4, AF4). The labels for the characteristics reflect where the electrodes are placed.

2.2.2. Survey related to the usage of mobile app

The survey to assess the app’s feasibility and acceptance by the families consists of 6 questions: (1) Is the diagnosis shown on the mobile app correspond to that of the psychiatrists? (2) Is it easier to use the app on a child? (3) Do you suggest this app to any other people? (4) Would you prefer this app for dyslexia classification at home or school? (5) Is the price of acquiring the solution (EEG headset and software subscription- 1450 USD) high/ moderate/ low at home? (6) Is the price of 1-time measurement (children with dyslexia/ TDC classification- 100 USD) at school high/ moderate /low?

106 people participated in the survey. These are middle-aged parents (35- 45 years old, 60 female, 46 male). The socioeconomic status of the parents was middle class.

2.2.3. Socioeconomic status criteria for Socioeconomic situation

To evaluate the socioeconomic status of the children, a survey of their parents was conducted. The poll includes inquiries about occupation, education (primary school, high school, and university), and income (basic income 6,000 TL, mid-level income 6,000 TL to 20,000 TL, high income >20,000 TL) (staff, blue-collar workers, white-collar workers). The study's inclusion criteria were that the participants were not on medication and did not have any other comorbid disorders than dyslexia and they came from middle-SES families. They lived in many cities around Turkey.

2.3. Procedures

Each participant sat on a chair while the electrodes were inserted for the QEEG assessment. The distance between the participant and the mobile phone screen was 0.5 meters. The participants were asked to be in the resting state position while their eyes are open. The QEEG measurements were taken for 2 minutes using the EPOC-X and the app. The participant was requested to complete the QEEG measurement while relaxing and with their eyes open. This study has a limited sample size and is set up as a repeated measurement.

2-minute resting-state eyes-open theta, alpha, beta1, beta2, and gamma-band powers were measured (40 sessions per subject on average). This study uses a small sample size and multiple measures (8,301 sessions and 207 participants). The typically developing children's 2-minute resting-state eyes open QEEG data are gathered, and the data for the experimental and control groups are balanced to have an equal number of instances in each group.

2.6. Statistical Analysis

Python/Google Collab, Sci-kit Learn, and TensorFlow ML libraries were used to conduct the statistical and data analysis. K-Folding, Cross-validation, and confusion matrix generator functions derived from ML libraries. Mat plot library stands for plotting learning, validation, and ROC curves. The averaged 2-minute QEEG band power data (continuous data) are acquired from fourteen electrode channels each session. For the whole data set, the Z-scores are computed using the equation $z = (x-m)/s$ for each QEEG band power (including the experiment and the control group) for all the data points (x) gathered from children with dyslexia and typically developing children. Because EMOTIV does not offer Z-scores, m and s stand for the sample's mean and sample standard deviation, respectively. Outliers (>5 or -5) were removed from the analysis. The missing values were replaced with the mean of the featured data. The data is labeled by a computer scientist following the doctor’s diagnosis of the participant. The children with dyslexia and typically developing children’s data were balanced with pre-processing operations and adjustments. The binary classification with a supervised ML model is applied.

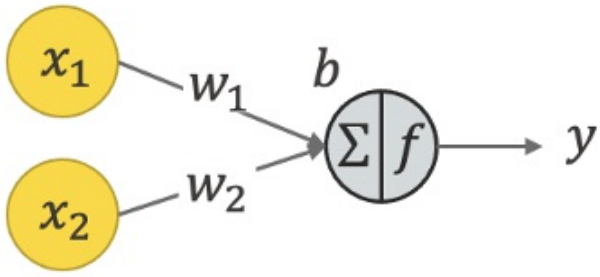

Figure 1. *Perceptron and Neural Network Illustration*

In figure 1, the illustration of a neural network was presented; where x1 and x2 represent inputs; w1 and w2 as weights; b is the bias, and y is the output of the neural network. The task is to sum the multiplication of input and weights with the addition of bias value. Right after that process, we have a function called the activation function that results in the scalar value of sigma notation through a transfer function.

Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1-f(x))$
TanH		$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2+1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1+e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$

Figure 2. *Activation Functions*

In figure 2, there are several activation functions that can be used defining our neural network. Since our data ranges between negative and positive values, rather than directly use an activation function called sigmoid neglecting negative

values and making them a 0, the data set is smoothed with activation functions like, -tanh, soft plus, arctan, logistic, or linear.

The number of true positives (TP), false positives (FP), and false negatives (FN) are used to calculate precision, recall, and F-score (FN).

- True Positive (TP): entities that are acknowledged by classification and match with ground truth
- True Negative (TN): Entities that are not acknowledged by classification and match with ground truth
- False Positive (FP): entities that are acknowledged by classification but do not match with ground truth
- False Negative (FN): entities that are not acknowledged by classification

Sensitivity (true positive rate) refers to the probability of a positive test, conditioned on truly being positive. Specificity (true negative rate) refers to the probability of a negative test, conditioned on truly being negative. F-score is the harmonic mean of precision and recall; the F1 score is common to evaluate performance across all of the classification models.

The model output is the dyslexia positivity probability score. The ML model architecture is Artificial Neural Network. The model features made in the study are 60 times repeat of gradient descent running (epoch 60), the average error is calculated and then weights are updated every 32 items (batch size 32), and loss as binary cross-entropy. Binary Cross Entropy is the negative average of the log of corrected predicted probabilities. The best model is selected among many other varying hidden layers and activation functions. The k-fold cross-validation technique has been used to evaluate the model with ten-folded cross-validation. This method is generally used to test model performance to estimate how well the model performs on unseen data. The overfitting is prevented by applying the dropout between layers. The results have been tested with an external test set which contains a diverse set of input data. The model is then converted to a TFLITE model and embedded into the mobile app.

The statistical analysis of the survey results was performed using SPSS. For the analysis of the data, primarily descriptive statistical methods are used. Mean ± standard deviation is used in the specification of numerical data, and the % ratio is used in the representation of categorical data. The "Pearson Correlation Test", which is a parametric test, is used in the analysis of the relationship between normally distributed data, and the "Spearman Correlation Test", which is a non-parametric test, is used in the analysis of the relationship between data that is not normally distributed.

3. Results

In this research, we have designed a Machine Learning algorithm to classify dyslexia and assessed the effectiveness with a survey of families with dyslexia. This study is a repeated assessment with a limited sample size (8,301 sessions) (207 participants: 96 of them have dyslexia, and 111 of them are typically developing).

The artificial neural networks achieved high accuracy and low-loss functions. The results show the performance of suggested preprocessing methods and the artificial neural network models achieved 98.8% accuracy with 0.05 loss with a 95% of confidence interval (CI, k-fold cross-validation applied), which is the state-of-the-art in literature for dyslexia

biomarker detection with EEG scans (Research Question 1, H1). In addition to that, this study concludes models which have additional preprocessing techniques like minimum-maximum scaler, standard scaler, etc. reduced the accuracy slightly from 98.8% to 98.63%. F1 score and loss become 0.986 and 0.07 respectively.

Table 1. Top 3 ANN results on the data set to detect the dyslexia biomarkers

Model Architecture	Accuracy	F1	Loss	AUC
1st tanh , 2ndsoftsign, 3rd tanh , 4thsigmoid,	**0.9880**	**0.983**	**0.05**	**0.98**
min max Scaler + 1st *tanh* , 2nd*softsign* 3rd *tanh* , 4th*sigmoid*	0.9863	0.986	0.07	0.96
1st *tanh* , 2nd*softsign* 3rd *linear* , 4th*sigmoid*	0.9818	0.982	0.09	0.95

In Table 1, the model architecture describes the activation functions at each level of ANN. Even though this study achieves high-accuracy results, it is also important to examine the cases of people who were misclassified by the model. According to the confusion matrix in Figure 3, the model claims 72 TDC sessions for children with dyslexia sessions. On the other hand, the model states that 27 children with dyslexia sessions as TDC sessions out of 8301 samples. In this study, the number of false negatives is more important than the number of false positives. Although misclassifying TDC sessions as children with dyslexia -false positives- sessions would not create problems when it comes to the training, it is essential to consider children with dyslexia's sessions classified as TDC -a false negative- session predictions.

Figure 3. *Confusion matrix result of the best model*

The receiver Operating Characteristic (ROC) Curve, which stands for sensitivity and specificity, of the ANN model indicates the capability of the model to discriminate dyslexia and the results were very promising (Figure 4).

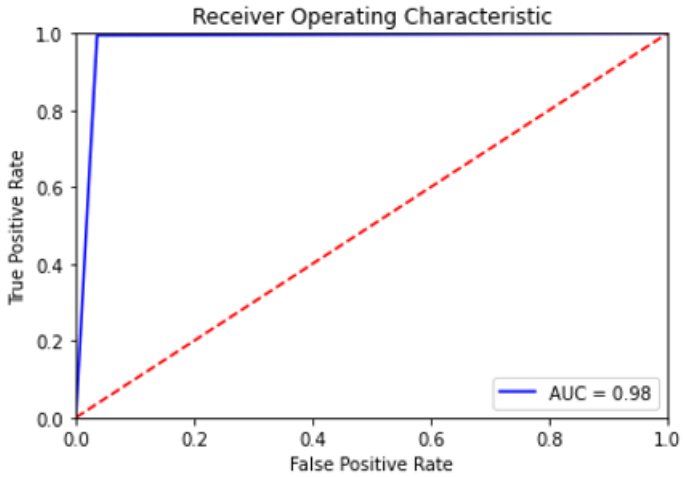

Figure 4. *Receiver Operating Characteristic Curve*

This study achieved a 0.98 area under curve (AUC) score. According to the ROC Curve definition, an AUC value between 0.8 and 0.9 is considered excellent, and more than 0.9 is outstanding. According to our ANN model, the most important features are theta power band values when it comes to the firing frequency of the perceptron.

Table 2. The performance results of the ANN network

Metric	Measured Value
Accuracy	98.80%
F1 Score	98.33%
Sensitivity	99.19%
Specificity	97.39%

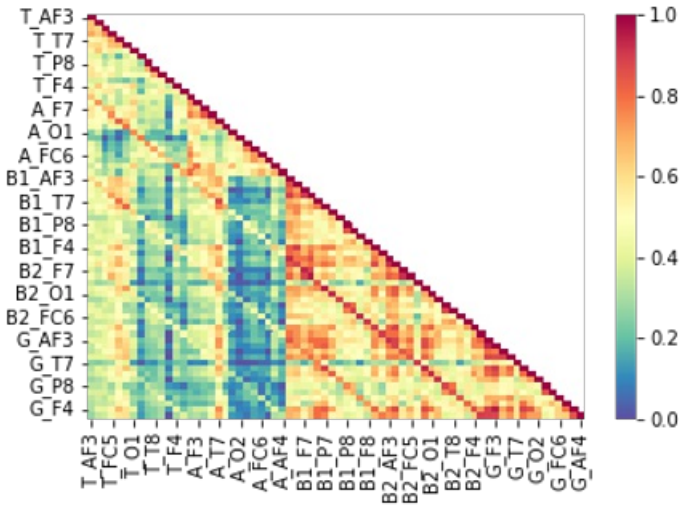

Figure 5. *Children with dyslexia, data correlation matrix of 70 features for x and y axis, Electrode places: AF3, F3, F7, FC5, T7, P7, O1, O2, P8, T8, FC6, F8, F4, AF4 for theta, alpha, beta1, beta2, and gamma bands*

Children with dyslexia data have a low correlation between alpha band power values at O2, P8, T8, FC6, F8, F4, and AF4 electrode locations and beta-1, beta-2, and gamma band values at most of the electrode locations ($p<0.001$) (Figure 5).

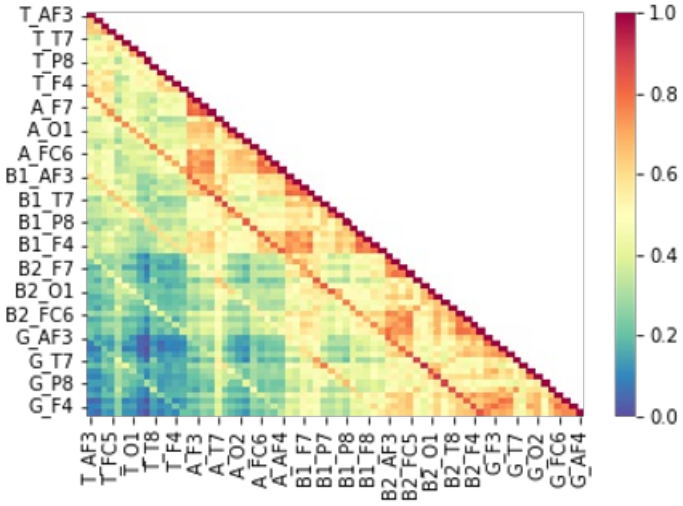

Figure 6. *Typically developing children (TDC) data correlation matrix of 70 features for x and y axis, Electrode places: AF3, F3, F7, FC5, T7, P7, O1, O2, P8, T8, FC6, F8, F4, AF4 for theta, alpha, beta1, beta2, and gamma bands*

According to figure 6, TDC data has a low correlation between theta band power values at all electrode locations and beta-2 and gamma band power values at most of the electrode locations ($p<.001$).

The survey results (N= 106) indicate that 76% of the respondents found the results of the diagnosis correspond to that of the psychiatrists (Research Question 2, H1); 82% of the respondents find the app easily usable at home (Research Question 3, H1); 80% of the respondents suggest the other families use the app. 45% of the respondents prefer to use the app for dyslexia/TDC classification at home (Research Question 4, H0). 80% of the respondents find the total cost of ownership (the price of the EEG headset and the software subscription) high for usage at home (Research Question 4, H0). 15% of the respondents think the price of the one-time measurement (dyslexia/TDC classification) at school is expensive (Table 3, figure 7) (Research Question 5, H1).

Table 3. Survey results on the usage of dyslexia diagnosis mobile app

Survey Questions	Results
Feasibility	76%
Acceptability	82%
Referral	80%
Prefer dyslexia detection usage at home	45%
Price high for usage at home	80%
Price high for one-time measurement at school	15%

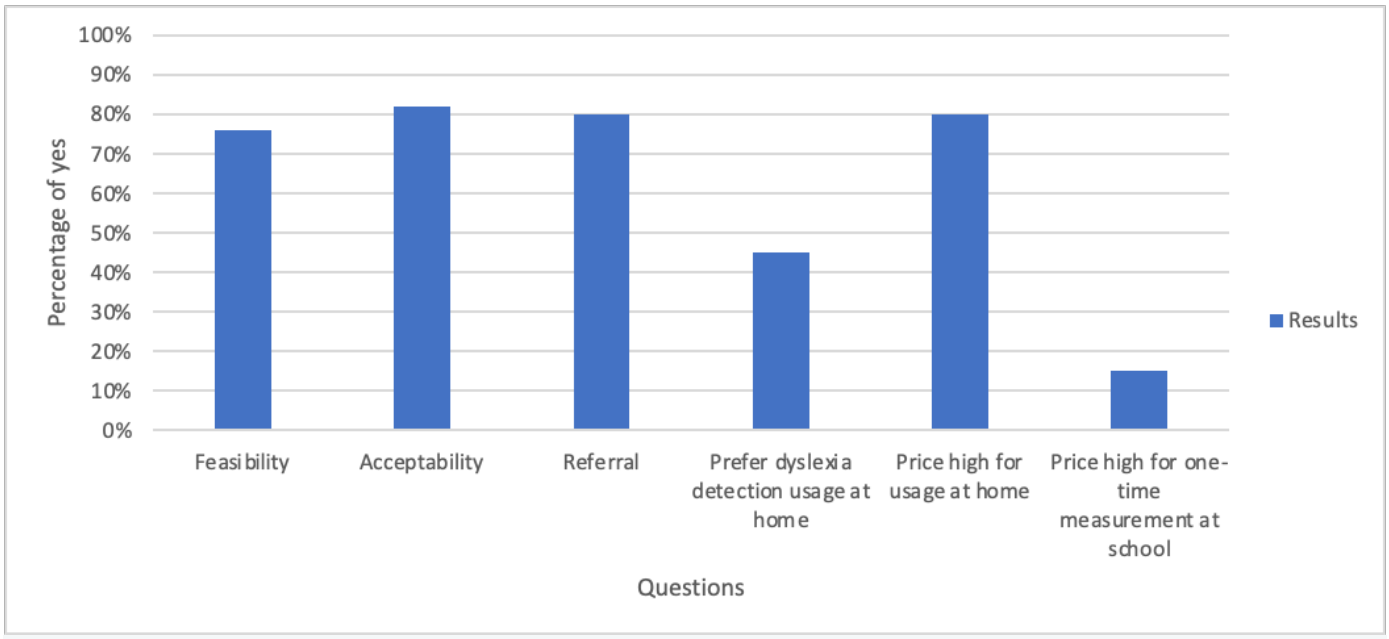

Figure 7. *Survey results as a bar chart*

4. Discussion

The novelty of this research is that Z-score normalization of 14- channel QEEG data let to the high-accuracy results and the ML model is embedded into a mobile app for everyday usage. This is the first research in the literature where ANN yields high accuracy results (98.8%) with z-scored QEEG data for dyslexia biomarker detection (Research Question 1, H1). Moreover, the feasibility, acceptability, and economic impact of the mobile app on families who have children with dyslexia are assessed with a survey in the real life afterward.

The QEEG power bands' correlation calculations between different electrode pairs for both children with dyslexia and TDC clearly show that children with dyslexia data have a higher correlation between theta, beta, and gamma signals at the left and the right hemisphere than the TDC data. According to figure 5, children with dyslexia data have lower correlations between the alpha band power at the right hemisphere with beta-1, beta-2, and gamma band power at the left and right hemispheres. These results indicate that the left lateralization of the cortex is delayed for children with dyslexia and they use the right hemisphere. Figure 6 indicates that theta band power at the left and the right hemisphere have lower correlations with the beta-2 and gamma values at the left and the right hemispheres for TDC. These data indicate that the left lateralization of the cortex is established for TDC and they use the left hemisphere.

As we collect QEEG data from 14 channels, it is expected that data may be correlated, and the covariance matrix is not diagonal. ANN still performs well under these conditions with QEEG data. Eliminating the noise and outliers in the data helped to achieve such high accuracy. On EEG scan data sets, Al-Barhamtoshy and Motaweh (2017) used ANN (89.6% percent accuracy). Karim et al. (2013) employed a Multilayer Perceptron to detect dyslexia signs by collecting brain waves in the resting state (85% accuracy). Using machine learning, Frid and Breznitz (2018) investigated the

variations in ERP signals between children with dyslexia and TDC (with ANN, 78% accuracy). Usman et al. (2020) achieved state-of-the-art CNN architecture with MRI scans at 84.6%. Our research achieved high accuracy results by the state-of-the-art ANN architecture with QEEG data at 98.8%.

The machine learning model created with ANN is converted to a TFLITE model and embedded into the Android Mobile Application on Android and iOS (Figure 8). In this way, the diagnosis of dyslexia with a high accuracy rate becomes possible with 2-minute resting-state QEEG data collected from a mobile app module (Figure 8, figure 9).

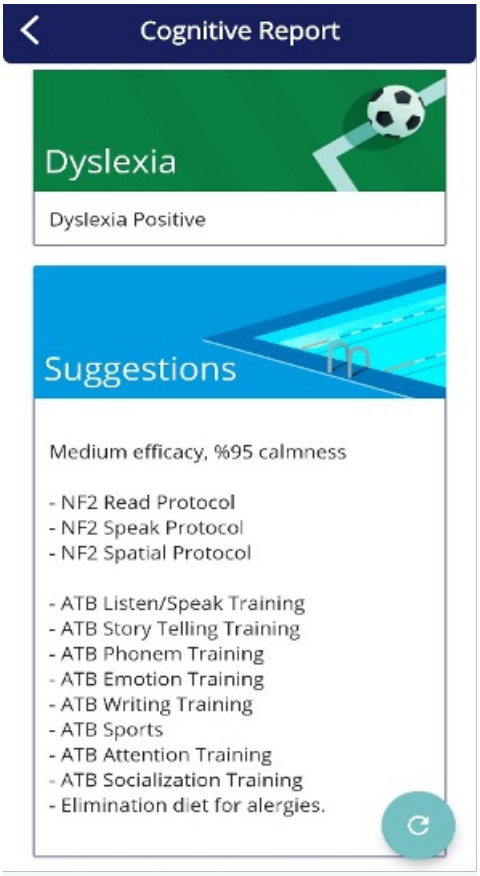

Figure 8. *The cognitive reports automatically generated after 2-minute QEEG Measurement*

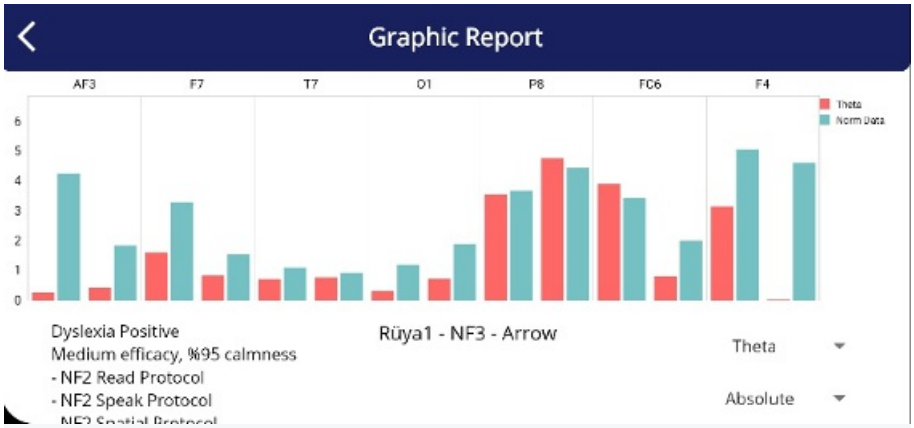

Figure 9. *The example cognitive reports that have been generated after a 2-minute QEEG Measurement*

With the advancement of technologies, dyslexia detection starts at the age of 7 with 2-minute resting-state QEEG data with a mobile app. This measurement is standardized and an objective assessment. The child sits quietly for 2 minutes in a chair, the EEG headset is placed in 2 minutes, and the resting state EEG measurement was accomplished in 2 minutes, the results (whether he has dyslexia or not) are immediately ready and can be seen on the screen. If there would be a possibility of dyslexia, remedial actions can be taken as soon as possible. The biomarker detection app catches some non-diagnosed dyslexic cases; we find it very important to receive an early remedial training. The families found the mobile app a feasible and acceptable solution for dyslexia detection (Research Questions 2 &3, H1). If it is a first-time measurement, they can get an assessment with this app at school which will cost less than buying the EEG headset. In this way, the first-time assessment at school becomes very economic for the families too.

Although technology is ready to diagnose dyslexia with a high accuracy rate and improve the symptoms of dyslexia with mobile health solutions, families still think the economic impact of the whole solution is high to use at home (Research Question 4, H0). Learning disabilities in generations impact the SES of families. The families who participated in our survey have middle SES, but they still find the solution expensive. Either EEG headset prices should come down to make it possible for all families with dyslexia to use them at home, or the solution should be available at school (Research Question 5, H1). As the dyslexia market is a niche (only 10% of society has dyslexia), EEG headset prices may not possibly be down in the short run. The awareness of dyslexia and the possible training methods are not high in society. Schools in Turkey stay neutral in dyslexia biomarker detection at school; the burden of getting an early diagnosis and intensive remedial work is on the shoulders of families and the state. The families need assistance by trusted parties like school management and the state to help them to decide on the right intervention methods for their children.

Dyslexia biomarker detection is not enough for families, when they see new technological advancements, they require a training roadmap for their child to overcome this condition as fast as possible. Using ML methods to predict the next possible training or the level/scale of dyslexia will be our next challenge to add to the mobile app.

4.6. Limitations of the Study

The study's first drawback is that it only includes 207. The trial should have included more participants; that would have been optimal. The likelihood of placebo effects is the study's second restriction. Children who get one-on-one interactions and specialized therapies may enhance their functioning primarily due to the social and environmental influence of those interventions, according to Gaab et al. (2019). Because the control group was not given an alternative intervention, placebo effects might be a substantial source of improvement. In the near future, we will repeat the experiment with more participants.

The study's third drawback is the maturation effects. Throughout their growth, all children's brains undergo major changes. As a result, QEEG modifications are anticipated to be influenced by maturation during the next six months.

5. Conclusions

Dyslexia biomarker detection apps could provide new methods for dyslexia training. In this work, we created a classification scheme for dyslexia diagnosis based on generated Z-scores from QEEG data and ANN. Families will gain quicker, objective evaluation of their children with dyslexia by using the mobile app at home or at school as finding an expert in dyslexia may not be easy and reachable for families.

Ethics approval and consent to participate

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

Availability of data and material

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability

None.

Acknowledgments

We want to express our sincere appreciation to the families that took part in the study; without their commitment and help, we might not have been able to finish it.

Declarations

Competing interests

None

Funding

None

Authors' contributions

G.E. wrote the main manuscript text including the title, abstract, introduction, materials and methods, results, and discussion. G.E prepared tables and figures.

M.K. wrote the results and enhanced the materials and methods with the ANN method and algorithm. M.K. prepared tables and figures.

B.K. collected the data from normal children.

All authors reviewed the manuscript.

References

- Al-Barhamtoshy, H. M., & Motaweh, D. M. (2017, February). Diagnosis of Dyslexia using computation analysis. In*2017 International Conference on Informatics, Health & Technology (ICIHT)* (pp. 1-7). IEEE.
- American Psychiatric Association. (2013). Anxiety disorders. In Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780896425596
- 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.
- 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.
- Bosch-Bayard, J., Peluso, V., Galan, L., Valdes Sosa, P., & Chiarenza, G. A. (2018). Clinical and electrophysiological differences between subjects with dysphonetic dyslexia and non-specific reading delay. *Brain sciences*, *8*(9), 172.
- Chen, H., Uddin, L. Q., Guo, X., Wang, J., Wang, R., Wang, X., ... & Chen, H. (2019). Parsing brain structural heterogeneity in males with autism spectrum disorder reveals distinct clinical subtypes. *Human brain mapping*, *40*(2), 628-637.

- 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.
- Dehaene, S. (2005). Evolution of human cortical circuits for reading and arithmetic: The “neuronal recycling” hypothesis. *From monkey brain to human brain*, 133-157.
- Dunn, L. M., & Dunn, L. M. (1965). Peabody picture vocabulary test--.
- Economou, M. (2022). Neuroanatomical outcomes of early literacy intervention in kindergarteners at risk for dyslexia.
- Frid, A., & Manevitz, L. M. (2018). Features and machine learning for correlating and classifying between brain areas and dyslexia. *arXiv preprint arXiv:1812.10622*.
- Gaab, J., Kossowsky, J., Ehlert, U., &Locher, C." Effects and components of placebos with a psychological treatment rationale–three randomized-controlled studies". Scientific reports, 9(1) ((2019): 1-8.
- Habib, M., & Giraud, K. (2013). Dyslexia. Handbook of clinical neurology, 111, 229-235.
- Iraola Goiburu, I. (2014). MEG-Based Functional Connectivity Biomarkers of Dyslexia (Doctoral dissertation).
- Janković, M. M. (2022, May). Biomarker-based approaches for dyslexia screening: A review. In 2022 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC) (pp. 28-33). IEEE.
- Jastak, S., & Wilkinson, G. S. (1984). *The wide range achievement test-revised*. Jastak Associates.
- Khan, R. U., Cheng, J. L. A., & Bee, O. Y. (2018). Machine learning and Dyslexia: Diagnostic and classification system (DCS) for kids with learning disabilities. *International Journal of Engineering & Technology*, *7*(3.18), 97-100.
- 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.
- King, W. M., Giess, S. A., & Lombardino, L. J. (2007). Subtyping of children with developmental dyslexia via bootstrap aggregated clustering and the gap statistic: comparison with the double-deficit hypothesis. *International journal of language & communication disorders*, *42*(1), 77-95.
- Kuhl, U., Neef, N. E., Kraft, I., Schaadt, G., Dörr, L., Brauer, J., ... & Skeide, M. A. (2020). The emergence of dyslexia in the developing brain. *Neuroimage*, *211*, 116633.
- Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron*, *77*(6), 1002-1016.
- Lyon, G. R., Shaywitz, S. E., &Shaywitz, B. A. (2003). A definition of dyslexia.*Annals of dyslexia*, *53*(1), 1-14.
- Madriaga, M. (2007). Enduring disablism: Students with dyslexia and their pathways into UK higher education and beyond. *Disability & Society*, *22*(4), 399-412.
- Nelson, N. W., Plante, E., Helm-Estabrooks, N., & Hotz, G. (2016). Test of integrated language and literacy skills (TILLS). *Baltimore, MD: Brookes*.
- Ozernov-Palchik, O., & Gaab, N. (2016). Tackling the ‘dyslexia paradox’: reading brain and behavior for early markers of developmental dyslexia. Wiley Interdisciplinary Reviews: Cognitive Science, 7(2), 156-176.
- Peterson, R. L., & Pennington, B. F. (2015). Developmental dyslexia. *Annual review of clinical psychology*, *11*(1), 283-307.

- Schrank, F. A. (2011). Woodcock-Johnson III Tests of Cognitive Abilities.
- Sharma, Sagar. “Activation Functions in Neural Networks.” Medium, Towards Data Science, 4 July 2021.
- Snowling, M. J., & Hulme, C. (2012). Annual Research Review: The nature and classification of reading disorders–a commentary on proposals for DSM-5. *Journal of child psychology and psychiatry*, *53*(5), 593-607.
- Spoon, K., Crandall, D., & Siek, K. (2019, June). Towards Detecting Dyslexia in children’s handwriting using neural networks. In *Proceedings of the International Conference on Machine Learning AI for Social Good Workshop, Long Beach, CA, USA* (pp. 1-5).
- Tokar A. https://www.ntv.com.tr/turkiye/disleksili-cocuk-sayisi-120-binin-uzerinde,FOF3AiVEc02EhGwNDcw0Sw, 2018.
- Usman, O. L., & Muniyandi, R. C. (2020, May 20). CryptoDL: Predicting dyslexia biomarkers from encrypted neuroimaging dataset using energy-efficient residue number system and deep convolutional neural network. MDPI. Retrieved July 28, 2022.
- Wang, J., Barstein, J., Ethridge, L. E., Mosconi, M. W., Takarae, Y., & Sweeney, J. A. (2013). Resting state EEG abnormalities in autism spectrum disorders. *Journal of neurodevelopmental disorders*, *5*(1), 1-14.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (1999).*Comprehensive test of phonological processing: CTOPP*. Austin, TX: Pro-ed.
- Wagner, R. K., Beal, B., Zirps, F. A., & Spencer, M. (2021). A model-based meta-analytic examination of specific reading comprehension deficit: how prevalent is it and does the simple view of reading account for it?. Annals of Dyslexia, 71(2), 260-281.
- Wang, J., Xing, G. G., Li, X., & Wan, Y. (2016). Enhanced gamma oscillatory activity in rats with chronic inflammatory pain. *Frontiers in neuroscience*, *10*, 489.
- Wechsler, D. (1999). Wechsler abbreviated scale of intelligence--.