Date available online: 30-06-2026
Vol. 4, Issue 1 (January - June 2026)

DOI: https://doi.org/10.24312/ucp-jeit.04.01.872

A Systematic Evaluation of Machine Learning Techniques for Dyslexia Detection Using EEG Signals

Iqra Muneer[2,*], Muhammad Kashif[1,3]

[1,*] Department of Electrical Engineering, The Islamia University of Bahawalpur, The Bahawalpur, 63100, Punjab, Pakistan.
[2] Department of Computer Science & Engineering, University of Engineering & Technology Lahore, Narowal Campus, Lahore, 54000, Punjab, Pakistan.
[3] Department of Computer Science, Comsats University Islamabad, Sahiwal Campus, Punjab, Pakistan.

*Corresponding author(s). E-mail(s): iqramuneer@uet.edu.pk;

ABSTRACT

Dyslexia is a learning disability that develops in childhood, leading to difficulty in reading, writing, and spelling, while having average or above average intelligence. The range of impact can be significant if this condition is not diagnosed early. Various methods have been tested to diagnose this condition, including psychometric assessments, facial and eye movement tracking, magnetic resonance imaging, and electroencephalograms. Electroencephalography (EEG) measures the electrical activity of the brain and contributes to understanding cognitive processes. It is particularly useful in recognizing patterns of neural activity that correlate with dyslexia, providing a cost effective and objective means of diagnosis. This paper attempts to review available diagnostic methods and highlight their shortcomings and gaps in research. It also evaluates machine learning and deep learning models for the classification of dyslexia using EEG data collected from a video-based educational paradigm. The models include Support Vector Machine, Random Forest, Logistic Regression, KNN, Decision Trees, AdaBoost, Naïve Bayes, CNN, LSTM, BiLSTM, and GRU. The best models in capturing EEG data patterns were CNN and LSTM, while the best classical baselines were SVM and AdaBoost. The proposed method showed a weighted F1 score of 0.99, proving reliability.

INDEX TERMS: Deep Learning, Machine Learning, Dyslexia, EEG based classification.

I. INTRODUCTION

Dyslexia is characterized as a variety of reading difficulties with a reading disability, typified by slow and inaccurate reading of words in a text and often accompanied by poor spelling ability. Similarly, phonological processing in the brain occurs, which involves associating phonemes with letters, and these phonemes are further organized into words, sentences, and paragraphs. Though more evidence emerged in the mid-2000s, dyslexics were said to have phonological processing problems that impeded reading and comprehension [1]–[3]. This difficulty in associating a letter with its corresponding sound results in decreased reading speed. While it takes a longer time for someone with dyslexia to read, they are often great at creating and reasoning through their thoughts. These reading deficits and the associated difficulties with word decoding usually remain into adult years, impacting academic activities in multiple areas of life.

Between the early and mid-2010s, it was clear from scientific studies conducted that dyslexia was a form of neurological disorder. In 2012, neuroimaging studies were able to ascertain that, among people who have dyslexia, there is a defect in the left anterior part of the brain that maps sound elements and processes them, thus resulting in auditory processing defects [4]. Integration of modern imaging technologies, including Positron Emission Tomography (PET), functional magnetic resonance imaging (fMRI), Magnetic Resonance Imaging (MRI), Electroencephalography (EEG), and Magnetoencephalography (MEG), have revealed smaller structural volumes in the regions of the left hemisphere of the dyslexic's brain, which is necessarily involved in the reading disorders [5].

As stated by the World Health Organization's World Report on Disability, there are about two billion men and women – around 37.5% of the total population – who experience disability in one form or another, even when the term is interpreted 'broadly' [6]. Of the most common and most important learning disabilities, the representatives of the National Center for Learning Disabilities (NCLD) make particular mention of dyslexia, dysgraphia, and dyscalculia [7].

Dyslexia is said to affect cognitive processes involved in the rational interpretation of visual symbols and various phonetic components of words. It also hinders the individual's ability to learn to read or spell, or even say words. Crucially, dyslexia should not be construed as an intellectual disability since it is a neurological condition. It is not rare, as one in every six persons is said to have it according to the estimates – in fact, it can be as low as 5% of the entire population or 17% [8]. Alongside the behavioral manifestations of dyslexia, there is also evidence to suggest that its genetic makeup can cause one to have it.

Dyslexia is more or less diagnosed by oral and written tests [7]. The patients suffering from this disorder usually have a normal IQ and don't suffer from visual disturbance as such. Most of the children suffering from dyslexia can progress academically with proper tutorship and, most importantly, emotional assistance. This disorder can also be prevented by early identification, counseling, and supportive care. Dyslexia is different for every person; in some cases, it is a very mild form and patients can breathe normally. There is no known drug that can effectively cure dyslexia, but through assessment and intervention at the early stages of the condition, the results are optimal.

An economically viable and convenient strategy for the evaluation of neurological disorders is the use of EEG (Electroencephalogram) signals as a screening procedure. Many traditional methods of imaging the brain are very costly and involve exposure to radiation, which makes them impractical for repeated use in the diagnosis of learning disabilities [9]. The EEG monitoring process consists of stimulation and recording of electrical signals emanating from the brain using electrodes placed on the scalp. The methodology of EEG-based dyslexia diagnosis includes the setting of the working parameters, time-series analysis of the signal components, selection and extraction of features, their building, and applying machine learning algorithms.

In order to carry out a proper study and analysis, the input parameters such as age, sex, and EEG channels relevant to dyslexics should be advanced. It is also important to be aware of inclusion and exclusion criteria and the layout of the experiment. Noise and artifacts should also be avoided and eliminated as much as possible in the signal's preprocessing phase after the targets are recorded, as such could lead to compromising the results. In light of the fact that EEG produces large volumes of data, the volume of this data needs to be minimized with the aim of retaining essential information. Feature extraction methods like the discrete wavelet transform (DWT) [10] and independent component analysis (ICA) are employed for such purposes. DWT separates features and classifies EEG signals into five frequency bands, which include the θ (theta band - 4–8 Hz), δ (delta band - 0.5–4 Hz), α (alpha band - 8–13 Hz), β (beta band - 13–30 Hz), and γ (gamma band - >30 Hz) [11].

In the analysis of the obtained data, it was seen that neural activity is connected with the change of the EEG frequency spectrum in the form of spatial coherence, alteration of amplitude/power, envelope changes, or the combination of these three features in a specific sequence. Research suggests that while low EEG signal frequencies coincide with arousal and movements, higher signal frequencies correlate with watching something, remembering something, expressing something, and emotional response [12]. It has been shown that every specific frequency band of EEG signal correlates with specific functions such as deep sleep (delta band), sleepiness (theta band), calmness (alpha band), thought processes (beta band), and stress or high behavioral activity (gamma band) [13].

This particularity makes the EEG specifically sensitive and selective for the following states of stress, alertness, resting, hypnosis, and sleep. There is profound EEG data where each subject's unique brainwave pattern allows for the depiction of spherical waves a subject exhibits in response to the positioning of electrodes over certain regions of the brain. This application of EEG makes it very useful for assessing specific functional deficits of the brain such as learning difficulties [14].

Importantly, EEG has some limitations as compared to other brain imaging techniques; however, it also has many advantages, and one of them is the time factor, since it takes just milliseconds for a complex brain structure to be stimulated and recorded. EEG is less expensive than most of its competitors, although it has low spatial resolution. On the whole, these advantages further justifiably extend the areas of use for EEG, allowing it to detect and diagnose children's learning disorders such as dyslexia [15].

The task has been investigated in several ways, including eye movement [16]–[19] and learning/speaking aloud using EEG, game activity-based learning without EEG, and brain imagery scans. On the other hand, speaking aloud with EEG learning, game activity-based learning [20], and video-based tasks with EEG have mostly been limited to one or two machine learning algorithms. Moreover, both tasks have not been investigated with deep learning models.

In this regard, the research took a quantitative approach where different techniques of machine learning and deep learning were deployed to examine the features extracted from EEG signals related to a learning activity.[1]

Within this investigation, we intend to employ several machine learning and deep learning approaches. For classification tasks, we depict several machine learning techniques such as Support Vector Machines (SVM), Random Forest (RF), Logistic Regression (LR), K-nearest neighbors (K-NN), Decision Trees, AdaBoost, Naive Bayes, Multilayer Perceptron, and Gradient Boost Classifier, along with deep learning models, which were designed, developed, evaluated, and compared.

Last but not least, we have developed and applied a variety of deep learning methods including LSTM and Bi-LSTM, CNN, GRU, and Bi-GRU in the task of dyslexia detection. To our best knowledge, an exhaustive study has never been done and reported before.

A. Description and Justification of the Pie Chart

In Figure 1, the concept of the pie chart is exploited, as the chart showcases the indications surpassing different dyslexia prediction techniques briefly referred to in the literature, as well as the outcomes of the recent study conducted. Each section of the pie chart depicts the percentage of accuracy of the different methodologies used in detecting dyslexia, highlighting the importance of the progress this study managed to achieve.

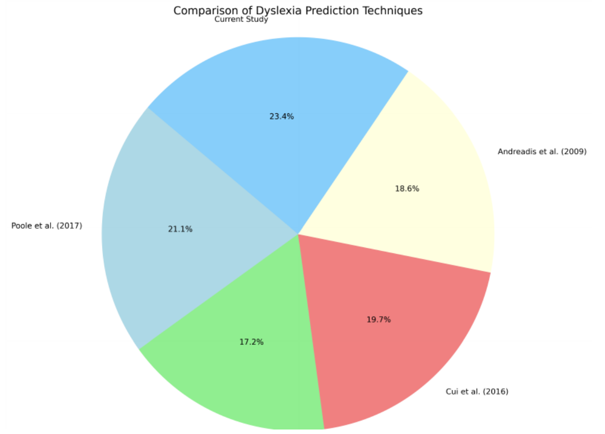

Figure 1: Performance wise Comparison of dyslexia Prediction Technique.

1) Contextualizing Existing Studies:

The graphic representation focuses on the achieved level of accuracy for completed studies by Poole et al. [21], Zahia et al. [22], Cui et al. [23], and Andreadis et al. [24]. This is a wide array of methods, ranging from simple phonological processing to more detailed neuroimaging. As a result of these studies, some methods have been developed that are more information-rich with regard to dyslexic behavior, even though their accuracy, ranging from about 72.73% to 89.2%, can be considered lacking in robustness, which calls for better approaches in the detection of dyslexia.

2) Highlighting Novelty:

This study presents a novel application that focuses on EEG-based machine learning and a deep learning (DL) model. The enhancement in accuracy, achieved through the use of a gamified learning environment and considering a large dataset of 12,811 instances, resulted in a weighted F1 score of 0.99. An even more considerable improvement over existing methodologies highlights the efficiency of coupling pragmatically interesting learning activities with powerful computing.

Unlike earlier dyslexia detection studies that mainly relied on phonological processing, eye-tracking, MRI/fMRI, or a limited number of EEG-based classifiers, this study presents a systematic evaluation of both classical machine learning and deep learning models for EEG-based dyslexia detection. The novelty of this work lies in the comparative analysis of a wide range of ML and DL techniques under a unified experimental framework using EEG signals collected from a video-based educational paradigm. This enables a clearer assessment of which model families are more suitable for capturing dyslexia-related neural patterns.

3) Who is likely to benefit from this work or these debatable findings?

These results not only show that it is possible to use the EEG signals of patients with dyslexia as a marker for prediction, but also that user engagement should be an integral part of the detection process. For the first time in this research, Support Vector Machines (SVM), Random Forest (RF) models, and Long Short-Term Memory (LSTM) networks are employed. This use represents a step toward more efficient and scalable methods for dyslexia detection.

4) How might this work or these findings impact more broad areas?

The results of this study help to lay a strong foundation for pursuing future research directions with a focus on gamification in education and the development of more advanced EEG analysis methodologies. This work clears the way for potential disorder diagnosis methods and interventions addressing children's needs at the very early stages, where prevention would be possible.

This study brings valuable contributions in both theoretical and practical aspects, including the use of machine learning in predicting dyslexia, which offers great opportunities for early screening, effective treatment, and large-scale prevention programs. This approach provides the opportunity to improve outcomes for those who are likely to develop dyslexia in the future, while promoting inclusion and accessibility in education and healthcare systems. However, it is important to carefully confront moral challenges, including data protection, fairness of algorithms, and the principles of informed consent, in order to avoid the irresponsible and unjust application of predictive modeling methods in the field of dyslexia research and clinical applications.

The remaining parts of the manuscript are organized in the following way. A discussion of the available current corpora and techniques used for dyslexia classification is presented in Section II. In Section III, further details concerning the methods proposed are provided. In Section IV, the authors disclose information on the preset experiment, while in Section V, the results and their interpretation are introduced. Finally, Section VI contains suggestions for further research and concludes the paper.

II. LITERATURE REVIEW

In this section, we outline the general tactics for early identification of learners at risk of dyslexia, integrating traditional methods and recent techniques, such as machine learning and EEG approaches, into one coherent picture.

A. Phonological Tests, and Reading Based Tests

Several studies [1], [25] demonstrate that the diagnosis of dyslexia has historically been carried out using phonological and reading-related tests that pinpoint difficulties at the phonological processing level, specifically awareness as well as manipulation of sounds, letters, and reading scripts. These tests, although helpful, do not assist end-users with accuracy and can be quite labor-intensive. Additionally, structural differences in the brains of dyslexic individuals in regions related to language processing have been observed through neuroimaging modalities like MRI and fMRI [4]. These procedures, however, are important when making a diagnosis, but are expensive and require research equipment, thus limiting their application.

B. Gamified Learning Techniques for Assisting in Dyslexia Screening

Multiple applications and games have been created for the diagnosis, support, and treatment of dyslexia [26]. Gamification techniques have been utilized to develop numerous use scenarios, applications, and frameworks [27]–[31]. These supplemental tools measure the performance of readers with the help of diverse linguistic content or measure the performance of pre-reading children by means of interactive games such as the Diesel-X game designed for preschoolers [32], [33]. Rello and colleagues, on the other hand, tackled the development of gamified tasks with the intrinsic objective of decreasing user stress during dyslexia assessment [20]. However, it is noted that though they attended to and boosted their engagement, there is limited integration of tools such as EEG, which are advanced, and thus the overall diagnostic accuracy is poor. Gamification is integrating game elements into gameplay for the purpose of engagement and motivation of users [26]. The latter type of games has been developed to enable readers to be assessed [34] on the basis of bilingual content and even assess preschool children as well [32], [33] with the aim of providing a game experience.

Except for recent works by Rauschenberger et al. [35], [21] and Poole et al. [21], the only study reported to use phonological processing features with a success rate of 89.2% is Lexa. However, Lexa did not design a game to be interesting and engaged in tedious and expensive ways of collecting data. This underscores the need for gamified dyslexia detection that encompasses both very high levels of engagement and accurate neuroscientific or physiological measures.

C. Machine Learning Techniques for Dyslexia Detection

In recent years, machine learning (ML) has gained traction as an efficient approach to dyslexia prediction. For instance, Rello and Ballesteros in 2015 proposed a metric-based graphology and eye movement dyslexia screening method for Spanish speakers that yielded astonishing results. They used Support Vector Machines (SVM) for analysis on a sample of 97 subjects, out of which 48 were dyslexic [36]. In the same vein, Nilsson et al. [37] used eye tracking metrics and SVM on a sample of 185 Swedish-speaking subjects, nearly half of whom were at high risk of dyslexia. Furthermore, Lakretz et al. used a probabilistic graphical model to search for dyslexia subtypes in existing medical records in Hebrew [38]. Whereas these investigations point to the usefulness of ML technologies in dyslexia detection, many of them are based on rather small populations and a limited number of models, which affect the transferability and accuracy of the studies.

Frid and Breznitz, in a certain aspect associated with dyslexia, provided an algorithm that made use of Support Vector Machines (SVM) combined with feature extraction like maximum peak amplitude, positive area, and flat measure. The classification was performed using SVM ensembles, and the federation of the obtained results was based on majority voting. However, in such an approach, the problems resulted from the small sample size, mainly composed of young subjects with dyslexia, and the lack of adequate and accurate measurements of inter-algorithmic performance.

In a previous work, Zainuddin et al. incorporated additional features based on EEG data in the conventional K-Nearest Neighbors (KNN) algorithm, as well as in its neo-KNN variant, to differentiate dyslexic children from their normally developing peers. In another study, Mahmoodin et al. oriented disabled children to perform Support Vector Machine (SVM) on EEG rhythms related to regions in the brain associated with dyslexia. Zainuddin and others also assessed the performance of the ELM with a radial basis function kernel in distinguishing EEG recorded during handwriting tasks for three categories: normal, poor dyslexics, and capable dyslexics. This research achieved a remarkable classification accuracy of 89% in ROC analysis, and sensitivity and specificity were demonstrated.

The signs associated with dyslexia have been highlighted, and Zainuddin et al. presented an evaluation of KNN in which a correlation distance measure was utilized alongside an ELM-based classifier in radial basis function geometry targeting EEG recording during writing in dyslexics. The lower breadths of the ELM framework preferred in dealing with coded handwriting emerged in supremacy against KNN with 89% success in accuracy.

Perera et al. investigated dyslexia with the help of an SVM classifier in the course of feature extraction and classification based on EEG signals in complementary research. Additionally, they analyzed the advantages and disadvantages of EEG-based approaches and suggested measures to improve the outcomes of the evaluation. In a later study [39], a distinct EEG pattern for dyslexia was identified with the help of a Cubic Support Vector Machine, emphasizing the difficulties accompanying writing and typing evident in dyslexic individuals as compared to non-dyslexics. This study also established the existence of new writing-related brainwave patterns in dyslexic individuals localized to the anterior-frontal lobe, and typing-related brainwave patterns recorded from the frontal channels. The classification results of the SVM were estimated by some classification performance metrics such as Classification Accuracy (CA), Sensitivity, and Specificity. Although only adult right-handed subjects were described in the case study, the findings generated important data for the future study of other aspects of dyslexia, such as uncommon characteristics of brain discharges, which require more thorough focus presently.

Several other machine learning techniques, however, have also classified sub-frequency EEG bands for diagnosing learning disabilities like Autism Spectrum Disorder (ASD) using the Shannon Entropy Vector, and Artificial Neural Networks (ANN) [40]. The Multi-layer Perceptron (MLP) is often repeatedly canvassed amidst other studies on EEG of emotions linked to addiction, aggressive behavior [41]–[43]. In a study, Razi et al. [44] were particularly interested in brain-related emotions analysis based on the MLP method and the NPIA method for EEG signal classification as positive or negative emotions. Children with dyslexia were identified and managed emotionally and non-emotionally by Selvi and Saravanan [45], and machine learning algorithms like SVM, Naive Bayes, decision trees, and even neural networks were applied for the purpose. Their results suggested that SVM and neural networks were quite effective. SVM techniques are widely known supervised learning methods, and much research has shown that it has the performance of creating adequate decision boundaries in order to classify new data points [46], [47].

D. Electrooculogram (EOG)-Based Techniques

It has been reported in several studies that EOG signals can be useful in monitoring eye movements during reading, making it possible to measure reading durations, performances, and movements more efficiently. For instance, Tabkara et al. [48] employed EOG signals in the evaluation of reading duration across different font styles and sizes, while Latifoglu et al. [49] employed EOG signals to accomplish detection of movements associated with word retrieval and re-reading with a classifier accuracy of 98 percent. EOG was adopted by Banerjee et al. [50] to estimate the number of sentences comprised in the text as well as the number of words, where a word count detection accuracy rate of 87.2% was reported. The obtained findings emphasize the role of EOG in recognizing dyslexic reading patterns. However, this predictive ability can be improved when combined with complementary diagnostic information, for instance, with regard to neural or behavioral parameters.

E. Analysis of Eye Movements and Phonology of a Dyslexic Individual and A Normal Individual

There has been research aimed at identifying reading patterns and analyzing the phonological skills of dyslexics and non-dyslexic people. Eden, Stein, and others [51] assessed eye movements in children using an infrared limbus monitoring system and observed the existence of differences between people with dyslexia and healthy individuals in most areas, particularly phonological abilities. Some studies have also noted that there are differences in the pattern of eye movements between healthy and dyslexic patients during reading, where dyslexics tend to have more regressions and longer fixations in a written material [52].

More recently, research aimed to analyze phonology-related deficits using electroencephalography (EEG) to assess phonological awareness at varying stages of reading development [53]. This work is specifically essential because deficits in phonological awareness are considered as the most common indicators of reading and writing problems, particularly in relation to dyslexia.

Table I compares this work on dyslexia prediction with the earlier work of Poole et al. and Rello et al. However, the primary focus of earlier works during that time remained on phonological processing and eye movement analysis only, which was very traditional. The current work claims a different perspective as it uses machine learning (ML) and deep learning (DL) models to analyze EEG data acquired through videos of learning activities in the most effective way possible. The 1,280-instance dataset is the largest dataset utilized in this study. To achieve good prediction, smaller sample sizes from much smaller quantitative studies had to create a variety of voting classifiers and neural networks. In this case, classifiers included SVM and RF, while neural networks included LSTM and CNN models. These were achieved through 10-fold cross-validation with a weighted F1-score of 0.99.

In addressing the problems noted in earlier works, such as variability with data and computational

Table 1: Comparative Analysis of Dyslexia Prediction Techniques

Study	Methodology	Sample Size	Key Findings	Unique Challenges
Poole et al. [21]	Phonological processing using a non-gamified app	Not specified	Achieved 89.2% accuracy but lacked engaging game elements.	High costs and time requirements for data collection; low user engagement.
Rello et al. [20]	Eye-tracking metrics	97 participants	Identified increased fixations and regressions in dyslexic readers.	Limited to eye-tracking metrics; may not capture broader cognitive challenges.
Nilsson et al. [37]	Eye-tracking and SVM	185 participants	Used eye-movement metrics to predict dyslexia; found similar trends as Rello et al.	Similar limitations as Rello; dependency on controlled environments.
Lakretz et al. [38]	Medical records analysis	Not specified	Identified dyslexia subtypes using existing medical data.	Relies on historical data; may not accurately reflect current conditions.
Zahia et al. [22]	Functional MRI (fMRI)	Not specified	Achieved 72.73% diagnostic accuracy by identifying brain activation patterns.	Invasive method requiring specialized equipment; accessibility issues.
Cui et al. [23]	MRI and diffusion tensor imaging	Not specified	Achieved 83.61% accuracy using a linear classifier on white-matter connectivity features.	Complexity of data interpretation and high computational cost of image processing.

complexity, this study proposes a better and more efficient solution for dyslexia detection. This method can prove to be useful in the timely diagnosis of dyslexia and the implementation of conductive interventions for a person with dyslexia more effectively at the appropriate time.

From the reviewed literature, several research gaps can be identified. First, many existing dyslexia detection studies rely on phonological, reading-based, or eye-tracking features, which may not fully capture the underlying neural activity associated with dyslexia. Second, EEG-based studies remain limited and often employ only one or two traditional classifiers. Third, deep learning models have not been sufficiently explored for EEG-based dyslexia detection, particularly in comparison with classical machine learning models under the same evaluation protocol. Fourth, many previous studies lack a systematic comparison of multiple ML and DL techniques using common performance measures. These gaps motivate the present study, which systematically evaluates a broad range of classical machine learning and deep learning models for EEG-based dyslexia detection.

III. APPLIED METHODS

A. Machine Learning based Methods

These transformations allow ML and AI to become powerful tools for change from the perspective of restructuring the patterns of many aspects of everyday life. It may be noted that learning algorithms make it possible to detect new patterns in large datasets, especially when EEG data is massive, and patterns alone are insufficient for the successful completion of the task. It is implicit that these algorithms are applied in the classification of signal data through a variety of techniques.

EEG studies in recent years have not ignored some important characteristics such as energy, average valley amplitude, peak alteration, root mean square, as well as power [54]. These features are combined in order to test the capability of such features to discriminate between subjects residing in the groups ‘controlled’ and ‘uncontrolled’. These classification predictions are made by employing the Confusion Matrix (CM), which is designed in a way to summarize classification inputs and assess the performance of the classification in terms of correct and incorrect classifications. Moreover, the CM provides additional indices of performance such as sensitivity, specificity, and accuracy [54], [55].

Different machine learning techniques for classification can be employed depending on the needs of each task. Some of the classifiers include: Support Vector Machine (SVM), Optimum-Path Forest (OPF), Naive Bayes, and k-Nearest Neighbor (KNN) classifiers [56]. In the context of EEG studies, however, linear discriminant analysis, SVM, and neural networks appear to be the popular algorithms [56], [57].

This is so because linear discriminant analysis is relatively easy to utilize as it simply builds models of the probability density functions of the data and classifies new data points through some probability assignment. Still, this has limitations when dealing with complex nonlinear classification problems, as is the case with EEG, where usually unsatisfactory results are often observed [58]. On the other hand, neural networks and SVM have proved accurate in the classification of EEG. In particular, neural networks are very well suited for the separation and definition of the boundaries of nonlinear functions, while for the SVM, both linear and nonlinear classification problems can also be well performed [58]–[60].

This work, in particular, uses a broad selection of classical machine learning techniques to classify dyslexic individuals, drawing on features that have already been extracted from EEG signal data. These are standard algorithms designed specifically for structured datasets and therefore appropriate for use in the analysis of EEG data. As illustrated in Figure 2, different types of machine learning were used: Support Vector Machines, and other classifiers such as Random Forest Classifier, Logistic Regression, K-Nearest Neighbors, Gradient Boosting, AdaBoost, Gaussian Naive Bayes, and Bernoulli Naive Bayes were employed. Figure 2 illustrates the workflow followed for classical machine learning-based dyslexia detection.

EEG signals are first collected and transformed into statistical features. These features are then evaluated using 10-fold cross-validation to ensure reliable performance estimation. Multiple classifiers, including SVM, Random Forest, Logistic Regression, KNN, Decision Tree, Gradient Boosting, AdaBoost, and Naive Bayes variants, are trained and tested. Finally, the models are evaluated using accuracy, precision, recall, and F1-score to identify the most effective classical ML method.

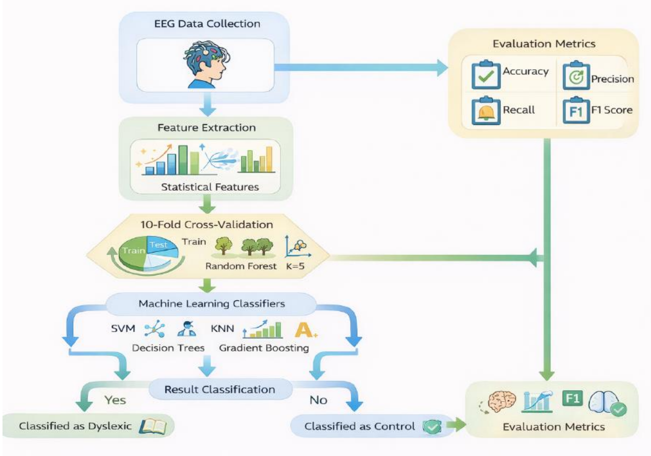

Figure 2: Classical Techniques Flow Chart.

In order to provide appropriate assessment of the model, a cross-validation approach, whereby 10-fold cross-validation is used, is applied. This approach consists of dividing the available dataset into 10 portions or folds, whereby 9 portions are used to train the model and 1 portion to test the model. This function is performed in cycles, allowing each of the ten portions to be tested once. Several metrics were adopted to appraise model robustness, including accuracy, precision, recall, and F1

UCP JOURNAL OF ENGINEERING AND TECHNOLOGY

score. Thus, using these methodologies, the current paper seeks to develop a dependable and trustworthy system capable of revealing dyslexia at the earliest.

1) Decision Tree Classifier:

Hierarchy is formed in a tree structure with the help of this algorithm, and new instances are classified based on handwritten feature values. Classification trees are interpretable and very effective for grasping non-linear dependencies [61].

2) Random Forest Classifier:

Random Forest models use a collection of decision trees to detect overfitting and improve generalization and robustness also when faced with many features [62].

3) Logistic Regression Classifier:

Logistic Regression is a modeling approach that predicts a dependent variable (binary) based on several independent variables. Logistic Regression is a less costly approach, with its coefficients fully interpretable, hence it is a good approach when assessing the contribution of features to the classification problem [63].

4) K-Nearest Neighbors (KNN) Classifier:

KNN is a non-parametric approach of classifying instances based on the class of majority of the neighbors that are nearest to the instance in the feature space. KNN is a simple technique, however, its performance depends greatly on how the parameter k and distance metric are selected [64].

5) Gradient Boosting Classifier:

In [65], boosting is described as an approach that attempts to change the performance of the model negatively by weak learners (typically decision trees) that are combined to reduce the prediction errors. Boosting is also good in handling intricate relationships in the data and is good for complex classification problems.

6) AdaBoost Classifier:

AdaBoost is an adaptive boosting technique that focuses on misclassified instances on each iteration and increases their weight when creating new classifiers.

7) Gaussian Naive Bayes Classifier:

Naive Bayes Gaussian is a probabilistic model that treats features as independent and uses Gaussian distribution to model the continuous features. Its wide area of usefulness shows a reasonable level of effectiveness even with the level of simplicity it offers, as long as there is a need for a high level of computational efficiency [66].

8) Support Vector Machine (SVM) Classifier:

Support Vector Machines are able to operate as discriminative classifiers by finding the ‘best’ or ‘optimal’ hyperplane, so that they can successfully divide the feature space of the classes to be separated. These are largely useful and applicable in the case of large dimensions. Additionally, the hyperplane can be modified to formulate more complex relationships of the dimensions by adding the appropriate kernel functions [67].

9) Bernoulli Naive Bayes Classifier:

The Bernoulli Naive Bayes Classifier is a special case of Naive Bayes Classifier developed for cases where the features of the data are represented in binary values. Under the assumptions of feature independence, it is one of the more appropriate classifiers for a dataset given in binary format [68].

This study analyzes the classifiers mentioned above on their effectiveness pertaining to dyslexia detection, specifically based on the EEG-derived signal features. Each algorithm has its own unique modeling capabilities, and enables different ways of studying the neural patterns related to dyslexia. The comparative analysis explains in detail the advantages and disadvantages of the different machine learning techniques in relation to the understanding of neurophysiological data.

B. Deep Learning based Methods

Apart from the such discussed studies, this study also aims at the potential applications of the deep learning algorithms and its combinations in EEG classification tasks. Deep learning has developed in revolutionized manner and has proven to be effective in such application areas. More importantly, convolutional neural network (CNN), recurrent neural networks (RNN), and deep belief network have also maintained better classification metrics of EEG data over stacked autoencoder and multilayer perceptron neural networks as noted in research works [69], [70]. These algorithms have been popular in a number of applications such as motor imagery [71], mental workload evaluation [72], seizure detection [73], event-related potentials [74] and sleep scoring [75]. It would be interesting in future research to look at deep learning approaches and its combinations in the detection and diagnosis of dyslexia.

This is in addition to classical machine learning algorithms in EEG based classification of dyslexia individuals where a number of neural network models are integrated to improve the classification accuracy. Some of the deep learning approaches utilize the hierarchical feature extraction properties of neural networks which enable complex more to be learnt when it comes to data patterns identification.

Figure 3 describes a pipeline for dyslexia detection using advanced deep learning techniques. Within this pipeline, it involves the processes of EEG signals feature extraction, deep generation of artificial models such as CNNs and long-short-term memory networks (LSTMs) and their bidirectional counterparts, and 10-fold cross-validation as well as performance assessment using precision, recall, F1-score, and accuracy as performance indicators for the models.

1) Convolutional Neural Network (CNN):

CNNs or Convolutional Neural Networks are a type of deep learning or neural network approach that have proven to be very effective at learning hierarchal spatial representations from various structured data (such as Images or time series data). In regard to this specific case, CNNs will be used to learn features from EEG signals by identifying spatial components, as well as local patterns, and spatial dependencies.

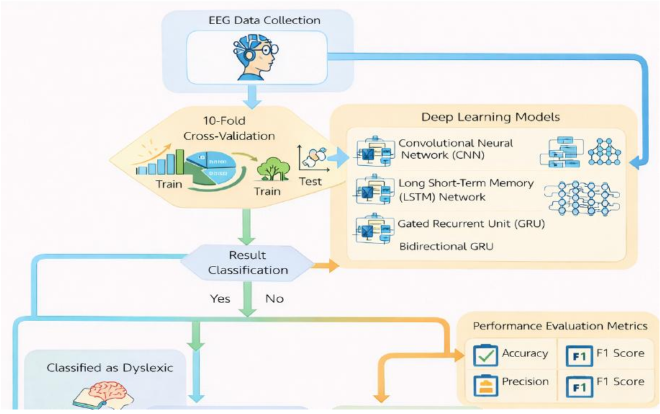

Figure 3: Deep Learning Techniques Flow Chart.

CNNs are a part of the family of supervised learning models, and, as such, consist multiple sequential, and interconnected, layers, each of which performs a specific task (such as convolution, activation, pooling, or even 'fully connected' layers).

In this case, the convolutional layers will perform 'feature extraction' through the use of learnable filters, while the pooling layers will perform spatial dimensionality (and, hence, also) computational complexity reduction. Tuning of the various layers is also one of the popular methods of CNN model dimensionality (and solution) reduction to improve learning/representational confidence (and/or) model solution depth.

Due to their iterative learning approach, CNNs have been shown to be very accurate across many applications of biomedicine (for instance) ECG-arrhythmia detection, EEG epilepsy diagnosis, and (even) the classification of various stages of sleep via PSG signals, and classification of eye movements.

2) Long Short-Term Memory (LSTM):

Long Short-Term Memory (LSTM) Networks were developed to overcome the drawbacks of the classic Recurrent Neural Networks (RNNs).

While classic RNNs were not able to learn patterns of dependencies over time (i.e. chains of events), LSTMs were able to overcome this limitation by "remembering" (i.e. keeping) information over large time spans (i.e. over long multiples of time periods) to sufficiently learn a) chains of events and, in particular b) dependencies of events over time (i.e. temporal dependencies, in particular) to the specific context of EEG signals [76].

3) Bidirectional Long Short-Term Memory (Bi-LSTM):

Bidirectional Long Short-Term Memory Networks (or Bi-LSTMs) enhance the functionality of standard LSTMs by adding the ability to learn two (2) different sequential time series (in a forward and, also, backward direction).

Thus, Bi-LSTMs are able to learn and capture not only past dependencies, and/or temporal dependencies, but also (and) future dependencies of the series being analyzed.

4) Gated Recurrent Unit (GRU):

Gated Recurrent Unit (or GRU) is a very simplified version of an LSTM, due to the reduced complexity of architecture while allowing the ability to learn and model event dependencies in time series data.

By merging the input and forget gates into a single update gate, GRUs can learn temporal relationships more efficiently due to the lowered number of parameters and computations [77].

5) Bidirectional Gated Recurrent Unit (Bi-GRU):

Bidirectional processing models coupled with GRUs are called Bi-GRU. That is, the model can learn positive and negative temporal patterns, which explains the model's ability to learn more intricate relationships in EEG signals [78].

To conclude, the fusion of these deep learning models provides for detailed analysis of the spatial and the temporal dimensions of the EEG signals which in turn provides for the recognition of more intricate patterns and relationships, thus enhancing the precision and the reliability of dyslexia detection.

IV. EXPERIMENTAL SETUP

The task of dyslexia prediction was treated as a supervised classification problem. The classification task aimed to differentiate between two classes: (1) Dyslexia, and (2) Non-Dyslexia.

A. Dataset & Pre-processing

All techniques were evaluated on a freely available dataset [79] from Kaggle. The dataset was prepared in the following manner and was released for public use. EEG signals were collected from a sample of 10 college students as they watched MOOC video clips. Videos expected not to confuse the college students because of their basic concepts in algebra and geometry were included.

Furthermore, videos identified to potentially create confusion to the average college student if only his basic concepts are familiarized include Quantum Mechanics and Stem Cell Research. This compilation comprised 20 videos (NECTV). Each of the topics was further divided into two with 10 videos for each and found in the NECTV series and each had a running time of about 2 minutes. To boost the target's confusion, each two-minute clip mid-way from the beginning of the topic was used.

The dataset comprises 12,811 data points inclusive of: 1) Dyslexia = 6,567 and 2) Non-Dyslexia = 6,244.

The dataset used in this study consists of EEG recordings collected from participants while they watched educational video clips. The EEG signals were recorded using a single-channel wireless EEG device placed over the frontal lobe. Each recording was associated with a class label representing either the dyslexic or non-dyslexic category. The final dataset contained 12,811 EEG samples, including 6,567 dyslexic samples and 6,244 non-dyslexic samples. These samples were used to train and evaluate machine learning and deep learning models for EEG-based dyslexia classification.

Before model training, the EEG data were prepared in a structured format suitable for machine learning and deep learning analysis. The raw EEG recordings were cleaned and organized into numerical feature representations. Irrelevant or incomplete entries were removed, and the class labels were encoded into binary categories. The extracted EEG features were then used as input variables for the classification models. This preprocessing step ensured that the models received consistent and meaningful input for distinguishing dyslexic and non-dyslexic EEG patterns.

The EEG data were used to capture neural activity patterns associated with the cognitive processing of educational video content. Since dyslexia is linked with differences in neurological processing, EEG signals provide useful information for identifying possible differences between dyslexic and non-dyslexic individuals. In this study, the extracted EEG features were supplied to classical machine learning classifiers and deep learning models. The models learned patterns from the EEG-based feature space and classified each sample into one of two categories: dyslexic or non-dyslexic.

B. Techniques

The model training process was conducted using both classical machine learning and deep learning approaches. Classical models included Decision Tree, KNN, Random Forest, Gradient Boosting, Logistic Regression, AdaBoost, Gaussian Naive Bayes, Support Vector Machine, and Bernoulli Naive Bayes. Deep learning models included CNN, LSTM, Bi-LSTM, GRU, and Bi-GRU. Each model was trained and evaluated using the same dataset and evaluation protocol. Accuracy, precision, recall, and F1-score were used to compare the classification performance of all models. To reduce the risk of overfitting and obtain a more reliable estimate of model performance, 10-fold cross-validation was applied. In this procedure, the dataset was divided into ten equal parts. In each iteration, nine parts were used for training and one part was used for testing. This process was repeated ten times so that each part of the dataset was used once as the testing set. The final performance was calculated across all folds. This approach helps ensure that the models are evaluated on unseen data and improves the reliability and generalization of the reported results.

Table II shows all techniques which were applied, and all the experiments were performed using the Python 3.8 framework, and the implementation can be found online.[2]

Firstly, classical machine learning models were experimented with: all the features were passed to the machine learning models and performance was reported. For every experiment, a unique name was assigned to each approach, as shown in Table II. For Experiment 2, a CNN-based model was proposed and developed with 64 filters at the convolution layer, 3 hidden layers with the ReLU activation function and sigmoid at the final layer with a pool size of 2. The network was trained using 100 epochs with the Adam optimizer using 10-fold cross-validation. For Experiments 3, 4, 5, and 6, LSTM, GRU, BiLSTM, and BiGRU were proposed and developed with the same hyperparameters and parameters as CNN.

Table 2: Experiments

Exp. No.	Machine Learning Techniques	Exp. No.	Deep Learning Techniques
Exp 1.1	Decision Tree Classifier	Exp 2	CNN
Exp 1.2	K-Nearest Neighbors Classifier	Exp 3	LSTM
Exp 1.3	Random Forest Classifier	Exp 4	BI-LSTM
Exp 1.4	Gradient Boosting Classifier	Exp 5	GRU
Exp 1.5	Logistic Regression Classifier	Exp 6	BI-GRU
Exp 1.6	AdaBoost Classifier		
Exp 1.7	Gaussian Naive Bayes Classifier		
Exp 1.8	Support Vector Classifier		
Exp 1.9	Bernoulli Naive Bayes Classifier		

C. Evaluation Measures

The most commonly used evaluation measure for classification tasks is accuracy; as the dataset is not completely balanced, other measures including weighted average precision, recall, and F1 are also reported.

D. Evaluation Methodology

This work was regarded as a supervised classification problem, with the primary focus on arranging the samples into 2 classes: (1) Dyslexic, and (2) Non-Dyslexic. For completing this task, a total of ten machine learning algorithms were employed through the Scikit-learn library. The algorithms employed were: Random Forest, Decision Tree, Naive Bayes Classifier, Gaussian Naive Bayes, Gradient Boosting Classifier, AdaBoost, Multilayer Perceptron, K-Nearest Neighbours, Support Vector Machine, and Logistic Regression. For the evaluation of these algorithms, 10-fold (K=10) cross-validation was used. Each algorithm was supplied with input sources comprising features obtained using the pre-processing techniques. The performance of the models was further evaluated according to the macro average F1 scores, and the system performance was reported through this parameter.

UCP JOURNAL OF
ENGINEERING AND TECHNOLOGY

V. RESULTS AND ANALYSIS

A. Machine Learning Performance

The results show a fairly low level of performance for all the machine learning experiments (Exp 1.1–Exp 1.9). The greatest accuracy was observed in Exp 1.8, reaching up to 60.2%, while other experiments performed within the accuracy ranges of 53.0%–59.9%. Similarly, the measures of precision, recall, and F1-scores were moderately low for these experiments, remaining in a constant range of approximately 0.53–0.60. Machine learning methods were limited by pre-defined feature engineering for the EEG data and, therefore, were unable to reveal intricate patterns related to dyslexia. Nevertheless, they set the stage for deep learning models to be compared against them.

Table 3: Results

Models	Accuracy	Precision	Recall	F1
Exp 1.1 — Decision Tree	0.599250	0.60000	0.60000	0.60000
Exp 1.2 — KNN	0.539942	0.53000	0.53000	0.53000
Exp 1.3 — Random Forest	0.536720	0.54000	0.54000	0.54000
Exp 1.4 — Gradient Boosting	0.557480	0.56000	0.56000	0.56000
Exp 1.5 — Logistic Regression	0.530013	0.54000	0.53000	0.47000
Exp 1.6 — AdaBoost	0.597611	0.60000	0.60000	0.60000
Exp 1.7 — Gaussian Naive Bayes	0.528920	0.56000	0.53000	0.49000
Exp 1.8 — Support Vector Machine	0.602000	0.60000	0.60000	0.60000
Exp 1.9 — Bernoulli Naive Bayes	0.561314	0.57000	0.56000	0.53000
Exp 2 — CNN	0.996094	0.998437	0.998437	0.998437
Exp 3 — LSTM	0.996094	0.996175	0.996184	0.996176
Exp 4 — BI-LSTM	0.996094	0.996175	0.996184	0.996176
Exp 5 — GRU	0.996094	0.996175	0.996184	0.996176
Exp 6 — BI-GRU	0.996094	0.996175	0.996184	0.996176

B. Deep Learning Performance

In all metrics, deep learning models (Exp 2–Exp 6) outperformed machine learning models. The deep learning experiments consistently achieved accuracy no lower than 99.6%, with precision, recall, and F1-scores all equally high at approximately 99.8%. This demonstrates the superiority of deep learning models when handling high-dimensional data such as EEG signals for dyslexia prediction.

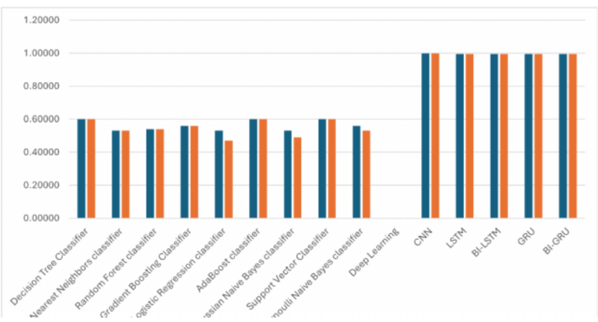

Figure 4: Performance wise Comparison.

C. Reasons for Superior Performance of Deep Learning

1. **Feature Learning from Raw Data:** Deep learning models based on neural networks can automatically extract features directly from raw EEG data. This capability enables them to identify intricate and subtle patterns within high-dimensional EEG datasets without requiring manual feature engineering.
2. **Hierarchical Representation:** Deep learning models learn hierarchical representations of data. In the case of EEG signals, this includes low-level features such as frequency and amplitude variations and high-level features such as focal dyslexia-related brainwave patterns. This hierarchical understanding enhances classification performance.
3. **Non-Linear Relationships:** EEG signals exhibit complex non-linear relationships with dyslexia. Deep learning models, particularly deep neural networks, are naturally suited to modeling such relationships, leading to higher predictive accuracy.
4. **High Capacity and Flexibility:** Deep learning models have large parameter spaces, allowing them to learn diverse patterns across different EEG channels. This flexibility enables them to handle variations in brain activity among dyslexic individuals, making them robust for dyslexia detection.
5. **End-to-End Learning:** Deep learning models operate in an end-to-end manner, integrating feature extraction and classification into a single process. This integration improves both the speed and accuracy of operations compared to traditional machine learning models.

D. Limitations of Machine Learning Models

1. **Dependence on Feature Engineering:** Machine learning models rely on fixed feature sets, which are often hand-crafted and may be inadequate for capturing the complexity of EEG data.
2. **Limited Capacity:** The simple architectures of machine learning models limit their ability to model hierarchical features and complex non-linear relationships, reducing their predictive performance.

VI. CONCLUSION AND FUTURE WORK

Dyslexia is a learning disorder of children that causes reading, spelling, and writing difficulties during development despite normal or above-average intelligence. These difficulties may result in low self-esteem, anger, and distress. Thus, there is a great need to identify them early so that necessary help and interventions may be provided to reduce and ultimately improve their impacts.

In detecting dyslexia, a number of approaches have been employed including games, reading and writing, face recognition, eye-tracking, and brain scans like MRI and EEG. Of these approaches, EEG is especially useful for investigating brain activity. All these approaches enable the understanding of differences in neurological functioning in people with dyslexia and provide a plausible method for the diagnosis of this learning disorder.

The prompt recognition and accurate diagnosis of neuropsychological disorders such as dyslexia enable the imposition of appropriate therapies resulting in better prognosis and lower costs in the future. In the presented study, the authors examined the EEG data of individuals with learning difficulties employing machine learning algorithms. The aim of this study was to discover the most effective models for the diagnosis of dyslexia from a range of models which were trained and tested using various machine learning and deep learning approaches, including Support Vector Machine, Random Forests, Logistic Regression, K-Nearest Neighbors, Decision Tree, AdaBoost, Naive Bayes, Multilayer Perceptron, Gradient Boost Classifier, Long Short-Term Memory networks, Bidirectional Long Short-Term Memory networks, Convolutional Neural Networks, Bidirectional Convolutional Neural Networks, and Bidirectional Gated Recurrent Units.

The accuracy score of 0.99 achieved by the best model along with 10-fold cross-validation is remarkable. These results suggest that progressive ML and DL techniques can be used to address the needs of defining dyslexia more accurately and in less time. Such approaches may lead to considerable progress in the area of learning disability diagnostics and assist in early assessment and intervention.

Although these findings are promising, there is still an exploration gap of state-of-the-art neural networks to better the prediction of dyslexia. Future works will need to explore the integration of EEG data with other bio-physiological modalities or behavioral measures to enhance the understanding of diagnostics. Also, to broaden the external validation of these models, samples such as left-handers and children can be allowed in the future. Furthermore, attention-based contemporary neural models could be looked at, as they may provide ways of increasing prediction performance and making a difference in the lives of dyslexics. These perspectives emphasize the current opportunities for improving diagnostic tools for learning disabilities and the contribution of science to their enhancement.

Data Availability Statement

The dataset utilized in this study, titled “Confused Student EEG Brainwave Data,” is publicly accessible via Kaggle at https://www.kaggle.com/datasets/wanghaohan/confused-eeg. This dataset comprises EEG recordings from 10 college students as they watched Massive Open Online Course (MOOC) video clips, including EEG signals, demographic information, and self-reported confusion levels. No new data were generated or analyzed during this study.

Competing Interests

The authors declare that they have no competing interests in relation to the publication of this paper.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

[1] C. Christo, J. M. Davis, and S. E. Brock, Identifying, Assessing, and Treating Dyslexia at School. Springer Science & Business Media, 2009.

[2] P. Ott, Teaching Children with Dyslexia: A Practical Guide. London: Routledge, 2007.

[3] G. Reid, "Identification and assessment of dyslexia and planning for learning," in The Routledge Companion to Dyslexia. London: Routledge, 2012, pp. 102–115.

[4] B. Boets et al., "Intact but less accessible phonetic representations in adults with dyslexia," Science, vol. 342, no. 6163, pp. 1251–1254, 2013.

[5] K. A. Stage and A. S. Davis, Overcoming Dyslexia: A New and Complete Science-Based Program for Reading Problems at Any Level, 2005.

[6] M. Iwabuchi, R. Hirabayashi, K. Nakamura, and N. K. Dim, "Machine learning based evaluation of reading and writing difficulties," in Harnessing the Power of Technology to Improve Lives, IOS Press, 2017, pp. 1001–1004.

[7] C. Billard and F. Delteil-Pinton, "Clinique de la dyslexie," Archives de Pediatrie, vol. 17, no. 12, pp. 1734–1743, 2010.

[8] J. Pan, M. Yan, J. Laubrock, H. Shu, and R. Kliegl, "Saccade-target selection of dyslexic children when reading Chinese," Vision Research, vol. 97, pp. 24–30, 2014.

[9] B. Fischer, M. Biscaldi, and P. Otto, "Saccadic eye movements of dyslexic adult subjects,"

Neuropsychologia, vol. 31, no. 9, pp. 887–906, 1993.

[10] C. Galdi, M. Nappi, D. Riccio, and H. Wechsler, "Eye movement analysis for human authentication: A critical survey," Pattern Recognition Letters, vol. 84, pp. 272–283, 2016.

[11] S. Heim and A. Keil, "Large-scale neural correlates of developmental dyslexia," European Child & Adolescent Psychiatry, vol. 13, pp. 125–140, 2004.

[12] J. Dushanova and M. Christov, "The effect of aging on EEG brain oscillations related to sensory and sensorimotor functions," Advances in Medical Sciences, vol. 59, no. 1, pp. 61–67, 2014.

[13] R. Bradley, L. Danielson, and D. P. Hallahan, Identification of Learning Disabilities: Research to Practice. London: Routledge, 2002.

[14] M. F. C. A. Rani, R. Rohizan, and N. A. Abd Rahman, "Web-based learning tool for primary school student with dyscalculia," in Proc. 6th Int. Conf. Information Technology and Multimedia, IEEE, 2014, pp. 157–162.

[15] S. Mohamad, W. Mansor, and K. Y. Lee, "Review of neurological techniques of diagnosing dyslexia in children," in Proc. 2013 IEEE 3rd Int. Conf. on System Engineering and Technology, IEEE, 2013, pp. 389–393.

[16] A. J. Parker, T. J. Slattery, and J. A. Kirkby, "Return-sweep saccades during reading in adults and children," Vision Research, vol. 155, pp. 35–43, 2019.

[17] M. Biscaldi, B. Fischer, and F. Aiple, "Saccadic eye movements of dyslexic and normal reading children," Perception, vol. 23, no. 1, pp. 45–64, 1994.

[18] R. Schleicher, N. Galley, S. Briest, and L. Galley, "Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired," Ergonomics, vol. 51, no. 7, pp. 982–1010, 2008.

[19] C. H. Morimoto and M. R. Mimica, "Eye gaze tracking techniques for interactive applications," Computer Vision and Image Understanding, vol. 98, no. 1, pp. 4–24, 2005.

[20] L. Rello et al., "Dyslexia assessment through gamified tests: Reducing user stress during evaluations," Behaviour & Information Technology, vol. 37, no. 3, pp. 286–296, 2018.

[21] A. Poole, F. Zulkernine, and C. Aylward, "Lexa: A tool for detecting dyslexia through auditory processing," in Proc. 2017 IEEE Symp. Series on Computational Intelligence (SSCI), IEEE, 2017, pp. 1–5.

[22] S. Zahia, B. Garcia-Zapirain, I. Saralegui, and B. Fernandez-Ruanova, "Dyslexia detection using 3D convolutional neural networks and functional magnetic resonance imaging," Computer Methods and Programs in Biomedicine, vol. 197, p. 105726, 2020.

[23] Z. Cui, Z. Xia, M. Su, H. Shu, and G. Gong, "Disrupted white matter connectivity underlying developmental dyslexia: A machine learning approach," Human Brain Mapping, vol. 37, no. 4, pp. 1443–1458, 2016.

[24] I. I. Andreadis, G. A. Giannakakis, C. Papageorgiou, and K. S. Nikita, "Detecting complexity abnormalities in dyslexia measuring approximate entropy of electroencephalographic signals," in Proc. 2009 Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, IEEE, 2009, pp. 6292–6295.

[25] D. M. Eden, J. Stein, H. Wood, and F. Wood, "Differences in eye movements and reading problems in dyslexic and normal children," Vision Research, vol. 34, no. 10, pp. 1345–1358, 1994.

[26] A. Rauschenberger, R. Baeza-Yates, and L. Rello, "Technologies for dyslexia," in Web Accessibility: A Foundation for Research, 2019, pp. 603–627.

[27] J. Hamari, J. Koivisto, and H. Sarsa, "Does gamification work? A literature review of empirical studies on gamification," in Proc. 2014 47th Hawaii Int. Conf. on System Sciences, IEEE, 2014, pp. 3025–3034.

[28] A. Ritzhaupt, N. Poling, C. Frey, and M. Johnson, "A synthesis on digital games in education: What the research literature says from 2000 to 2010," Journal of Interactive Learning Research, vol. 25, no. 2, pp. 261–280, 2014.

[29] A. Mora, D. Riera, C. Gonzalez, and J. Arnedo-Moreno, "A literature review of gamification design frameworks," in Proc. 2015 7th Int. Conf. on Games and Virtual Worlds for Serious Applications (VS-Games), IEEE, 2015, pp. 1–8.

[30] K. Seaborn and D. I. Fels, "Gamification in theory and action: A survey," International Journal of Human-Computer Studies, vol. 74, pp. 14–31, 2015.

[31] A. Thomas, F. Bader, J. Thomaschewski, and M. Rauschenberger, "Integrating gamification: The human-centered gamification process," in Proc. WEBIST, 2021, pp. 430–435.

[32] O. Gaggi, C. E. Palazzi, M. Ciman, G. Galiazzo, S. Franceschini, M. Ruffino, S. Gori, and A. Facoetti, "Serious games for early identification of developmental dyslexia," Computers in Entertainment, vol. 15, no. 2, article 4, 2017.

[33] H. M. Geurts et al., "Game-based screening for early detection of dyslexia: Experiences from the Diesel-X project," International Journal of Child-Computer Interaction, vol. 6, pp. 60–67, 2015.

[34] L. Rello, R. Baeza-Yates, and J. Llisterri, "A resource of errors written in Spanish by people with dyslexia and its linguistic, phonetic and visual analysis," Language Resources and Evaluation, vol. 51, pp. 379–408, 2017.

[35] A. Rauschenberger and R. Baeza-Yates, "Recommendations to handle health-related

small imbalanced data in machine learning," 2020.

[36] L. Rello and M. Ballesteros, "Detecting readers with dyslexia using machine learning with eye tracking measures," in Proc. 12th Int. Web for All Conf., 2015, pp. 1–8.

[37] M. Nilsson Benfatto, G. Oqvist Seimyr, J. Ygge, T. Pansell, A. Rydberg, and C. Jacobson, "Screening for dyslexia using eye tracking during reading," PLoS One, vol. 11, no. 12, p. e0165508, 2016.

[38] Y. Lakretz, G. Chechik, N. Friedmann, and M. Rosen-Zvi, "Probabilistic graphical models of dyslexia," in Proc. 21st ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, 2015, pp. 1919–1928.

[39] H. Perera, M. F. Shiratuddin, K. Wong, and K. Fullarton, "EEG signal analysis of writing and typing between adults with dyslexia and normal controls," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, p. 62, 2018.

[40] R. Djemal, K. Alsharabi, S. Ibrahim, and A. Alsuwailem, "EEG-based computer aided diagnosis of autism spectrum disorder using wavelet, entropy, and ANN," BioMed Research International, vol. 2017, pp. 1–9, 2017.

[41] X. Kang, D. Handayani, and H. Yaacob, "Comparison between Butterworth bandpass and stationary wavelet transform filter for electroencephalography signal," IOP Conf. Series: Materials Science and Engineering, vol. 1077, p. 012024, 2021.

[42] A. Wahab and N. Kamaruddin, "Brain developmental disorders' modelling based on preschoolers neuro-physiological profiling," Indonesian Journal of Electrical Engineering and Computer Science, vol. 12, pp. 542–547, 2018.

[43] O. Bazgir, Z. Mohammadi, and S. A. H. Habibi, "Emotion recognition with machine learning using EEG signals," in Proc. 2018 25th National and 3rd Int. Iranian Conf. on Biomedical Engineering (ICBME), IEEE, 2018, pp. 1–5.

[44] N. Mat Razi, A. W. Abdul Rahman, and N. Kamaruddin, "Detecting learning disabilities based on neuro-physiological interface of affect (NPIOA)," Indonesian Journal of Electrical Engineering and Computer Science, vol. 19, p. 164, 2020.

[45] H. Selvi and M. Saravanan, "A study of dyslexia using different machine learning algorithm with data mining techniques," International Journal of Engineering and Technology (UAE), vol. 7, pp. 3406–3411, 2018.

[46] V. Chakraborty, "Study on different classification models for predicting dyslexia," vol. 1, pp. 29–35, 2020.

[47] P. Raatikainen, J. Hautala, O. Loberg, T. Kärkkäinen, P. Leppanen, and P. Nieminen, "Detection of developmental dyslexia with machine learning using eye movement data," Array, vol. 12, p. 100087, 2021.

[48] M. Tabkara, F. Latifoglu, K. Ceran, S. H. Cavga, M. Yaprak, and C. G. Altintop, "Electrooculography based font and punto analysis," in Proc. 2017 Medical Technologies National Congress (TIPTEKNO), IEEE, 2017, pp. 1–4.

[49] F. Latifoglu, R. Ileri, E. Demirci, and C. G. Altintop, "Detection of reading movement from EOG signals," in Proc. 2020 IEEE Int. Symp. on Medical Measurements and Applications (MeMeA), IEEE, 2020, pp. 1–5.

[50] A. Banerjee, A. Rakshit, and D. Tibarewala, "Application of electrooculography to estimate word count while reading text," in Proc. 2016 Int. Conf. on Systems in Medicine and Biology (ICSMB), IEEE, 2016, pp. 174–177.

[51] G. Eden, J. Stein, H. Wood, and F. Wood, "Differences in eye movements and reading problems in dyslexic and normal children," Vision Research, vol. 34, no. 10, pp. 1345–1358, 1994.

[52] H. Perera, M. F. Shiratuddin, K. Wong, and K. Fullarton, "EEG signal analysis of writing and typing between adults with dyslexia and normal controls," International Journal of Interactive Multimedia and Artificial Intelligence, vol. 5, p. 62, 2018.

[53] S. M. Anwar, S. M. U. Saeed, M. Majid, S. Usman, C. A. Mehmood, and W. Liu, "A game player expertise level classification system using electroencephalography (EEG)," Applied Sciences, vol. 8, no. 1, 2018.

[54] J. Rodrigues, P. P. Filho, E. Jr, A. N, and V. Albuquerque, "Classification of EEG signals to detect alcoholism using machine learning techniques," Pattern Recognition Letters, vol. 125, 2019.

[55] I. Ben Slimen, L. Boubchir, Z. Mbarki, and H. Seddik, "EEG epileptic seizure detection and classification based on dual-tree complex wavelet transform and machine learning algorithms," Journal of Biomedical Research, vol. 34, 2020.

[56] F. Lotte, M. Congedo, A. Lecuyer, and L. Fabrice, "A review of classification algorithms for EEG-based brain–computer interfaces," Journal of Neural Engineering, vol. 4, 2007.

[57] A. Rodionov and A. L'vov, "Comparison of linear, nonlinear and feature selection methods for EEG signal classification," in Proc. Int. Conf. on Actual Problems of Electron Devices Engineering (APEDE), 2004, pp. 436–439.

[58] G. Honke, I. Higgins, N. Thigpen, V. Miskovic, K. Link, S. Duan, P. Gupta, J. Klawohn, and G. Hajcak, "Representation learning for improved interpretability and classification accuracy of clinical factors from EEG," 2020.

[59] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees. Wadsworth, 1986.

[60] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[61] D. W. Hosmer, S. Lemeshow, and R. L. Sturdivant, Applied Logistic Regression. Wiley, 2013.

[62] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967.

[63] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001.

[64] K. P. Murphy, Machine Learning: A Probabilistic Perspective. MIT Press, 2012.

[65] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[66] A. McCallum and S. Nigam, "Conditional models of identity uncertainty in information retrieval," in Proc. 22nd Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval, 1998, pp. 333–340.

[67] S. Nakagome, A. Craik, A. Sujatha Ravindran, Y. He, J. G. Cruz-Garza, and J. L. Contreras-Vidal, "Deep learning methods for EEG neural classification," in Neuroergonomics. Singapore: Springer, 2020, pp. 1–39.

[68] K. Zhang, N. Robinson, S. W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network," Neural Networks, vol. 136, 2020.

[69] Z. Tayeb, J. Fedjaev, N. Ghaboosi, C. Richter, L. Everding, X. Qu, Y. Wu, G. Cheng, and J. Conradt, "Validating deep neural networks for online decoding of motor imagery movements from EEG signals," Sensors, vol. 19, p. 210, 2019.

[70] K. Ho, J. Gwak, C. M. Park, and J. I. Song, "Discrimination of mental workload levels from multi-channel fNIRS using deep learning-based approaches," IEEE Access, vol. PP, pp. 1–1, 2019.

[71] M. S. Hossain, S. U. Amin, M. Alsulaiman, and G. Muhammad, "Applying deep learning for epilepsy seizure detection and brain mapping visualization," ACM Trans. Multimedia Comput. Commun. Appl., vol. 15, no. 1s, 2019.

[72] B. Ehteshami Bejnordi et al., "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer," JAMA, 2017.

[73] S. Biswal, H. Sun, B. Goparaju, M. B. Westover, J. Sun, and M. Bianchi, "Expert-level sleep scoring with deep neural networks," Journal of the American Medical Informatics Association (JAMIA), vol. 25, 2018.

[74] T. S. Holmstrom, "LSTM networks for EEG-based temporal pattern recognition," 2020.

[75] K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, 2014.

[76] Z. Yang, Y. Yang, C. Dyer, X. He, E. Hovy, and R. Socher, "Hierarchical attention networks for document classification," in Proc. 2016 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016, pp. 148–157.

[77] H. Wang, Y. Li, X. Hu, Y. Yang, Z. Meng, and K. M. Chang, "Using EEG to improve massive open online courses feedback interaction," CEUR Workshop Proceedings, vol. 1009, pp. 59–66, 2013.