ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 9, Issue 12, December 2022

Neural biomarkers for dyslexia detection using Machine Learning: A Review

[1] Mallikarjun Nuchhi,[2]Sagar Hegde,[3]Pavan Kalyan,[4]Prakruthi Prasad,[5]Rohan P
[6] Dr. Anusha Preetham

[1][2][3][4][5] Students, DSATM, Bengaluru, India.
[6] Faculty, DSATM, Bengaluru, India.
[1] mallikarjunnuchhi40@gmail.com, [2]sagarhegde01@gmail.com, [3]pavankalyan1122001@gmail.com
[4]prakruthiii13@gmail.com, [5]prohan758@gmail.com, [6]anushapreetham@dsatm.edu.in

Abstract— Dyslexia can be defined as a neurological disease that is branded by sloppy word understanding and overall deprived interpretation skills. It impacts a large number of school-aged kids, with boys being disproportionately affected, placing them at risk for poor academic achievement for the rest of their lives. Long-term, researchers want to develop a dyslexia diagnostic tool based on neural biomarkers. In this regard, a significant range of machine learning and, more recently, deep learning approaches have been deployed with above-chance classification accuracy across diverse types of data sets. In this paper, we carefully examine the latest machine learning techniques to detect this disease and its biomarkers.

I. INTRODUCTION

Dyslexia, a complex neurological brain condition, has recently attracted a lot of interest from present day neuroscientists. It's a neurological condition marked by sluggish and erroneous word understanding, as well as phonological impairment, which affects 5-17 percent of the diverse population in the world. In most situations, this illness begins in childhood and progresses to puberty, where it might have a negative impact on academic achievement. Dyslexia can also have a negative impact on a child's self-esteem and development of self- perception. Students with dyslexia frequently face bullying in the classroom, as well as detachment and isolation. Some dyslexic children have problems with higher-order processing and executive control systems along with visual attention span issues, which have a substantial impact on their ability to read. They may also have trouble remembering things and recognizing letters. As a result, dyslexic kids and youth have significant difficulties with word recognition, verbal memory, and communication speed. Dyslexia can either be acquired or developmental, based on the agony experienced by the victim's brain during development or a major injury such as stroke.

Dyslexia is caused by a neurological imbalance in the brain, which appears as behavioural and cognitive deficiencies in another three interacting aspects. The three-dimensional approach for existing dyslexia therapies includes, namely, cognitive domain, the behavioural domain and the biological domain, according to this diagram. Language impairment (phonological disorder), occurs when a person has been unable to form associations between alphabets (grapheme) and sounds (phoneme), resulting in reading and spelling difficulties, visual shortfall occurs when the magnocellular and parvocellular sub - systems of the brain are dysfunctional.

Anomalies in the magnocellular routes and cells have been linked to a variety of visual abnormalities, notably dyslexia. Auditory deficit, often known as hearing problems, occurs when the centralized auditory nerve system, which is responsible for hearing, is impaired, resulting in difficulty sensing, interpreting, and digesting audio waves.

Using datasets acquired from a variety of sources, researchers have developed many ways for identifying and diagnosing dyslexia and associated signs. Standardized psycho-educational examinations, eye movement monitoring, online games, EEG & MRI exams, PET scans, MEG scans along with video/images captured in phonological or cognitive activities are all examples of these source. Study participants' success, intellect, phonological processing, reading proficiency, and language growth are measured using standardized psycho-educational exams.

The review is undertaken on the basis of application and responses generated for each selected study, with the goal of exposing certain important issues that are preventing these strategies from being clinically acceptable. As a result, the purpose of this review paper is to test the performance, significance, constraints and donations of the latest studies that use ML methods for dyslexia detection.

II. RESOURCES AND PROCEDURES

A.Identification and collection of data

This work conducted a thorough works survey and analysis to describe a wide range of ML algorithms for its detection and categorizing dyslexia & its signs over the past ten years until 2020. (2010- 2020). To find and identify relevant papers, the selection method tracks the Preferred Reporting Items for Systematic Review & Meta- Analyses (PRISMA) criteria. Initially, acceptable publications were located using a combination of query terms from the search keywords in Google Scholar, PubMed Scopus, IEEE Xplore, Web of Science (Woos) and Science Direct's electronic literature search approach.

. In the field of dyslexia recognition and examination, keywords such as "classification methods", "classification problem", "biomarkers prediction", "intervention programmed", "assistive technologies", "machine learning methods", "deep learning algorithms", "neural networks", "deep neural networks", "neural biomarkers" & "MRI datasets". The articles that did not capture such conversation were filtered out of the search results. The titles, abstracts, and search keyword were used to screen the articles. At the conclusion of the search, there were 334 items found. 131 articles were shortlisted from the original lot after eliminating duplication. During the stage of screening, 2 reviewers were appointed to independently analyse the articles. After 43 papers were eliminated, 84 suitable articles were forwarded to three reviewers for a full examination. After meeting the criteria, they were moved on to the next stage of the screening process. The following were the criteria for inclusion: (1) publications published in English between 2010 and 2020; (2) articles that employed one or more machine learning methods to identify dyslexia; (3) articles that make use of the dyslexia dataset category which is mentioned in the introduction. The reviewers then compared their findings until they came to an understanding. Only 22 papers were chosen for critical assessment in the end.

B. Data collection

The gathering of dyslexia-related datasets using the methodologies indicated in the introduction is the first step towards detecting dyslexia biomarkers. As a result, early studies collected three types of datasets for the investigation of machine learning systems. The first set of data depicts the behavioural signs that dyslexic people exhibit. During the learning process, they are linked to phonologic awareness, working memory, reading/writing, and measuring student engagement level using facial information. According to studies based on machine learning methodologies, this sort of dataset is collected from students' writing styles, typing, reading, test results, and facial engagements in various languages and cultures. Brain imaging modalities, which capture distinctive behaviours and authorizations in the brains of study participants, are the second type of dataset used by ML methods for dyslexia detection. fMRI, MEG, EEG, PET, and other imaging modalities are among them. The third type of dataset is linked to patterns of eye movement during cognitive tasks. Rollo and Ballesteros gathered this type of data with the help of instruments like an eye movement tracker and EEG tests. For effective dyslexia diagnosis and analysis, various contemporary ML studies combine the collection of the very first category of datasets with the collection of either of the categories, two or three datasets. Although gathering the second and third kinds of datasets can help achieve improved accuracy in the identification of dyslexia and related biomarkers, they are very costly to obtain. In addition, their coverage is usually confined to a smaller group of people. Furthermore, throughout the scanning process, individuals, especially kids, may exhibit unusual behaviour. While there are several online platforms for collecting first-category dyslexia datasets (e.g., TestingMon.com and Pearson), second- and third-category biomarkers are openly present in multiple databases and directories such as Open Neuro, Bishop Blog, Kaggle and The Eckert Labs, among others.

C. Data research and pre-processing

The main objective of preparing datasets and pre - processing is to enable the classification to retrieve the most significant interpretable features from the dataset. How efficiently dyslexia biomarkers are categorized using typical machine learning classifiers is defined by the type of processing tasks that are performed on the transfer and sharing before they are introduced into the classification models. Pre-processing processes include data normalisation, extraction of features, tissue fragmentation, softening, aligning with based on image template, and modulation, to name a few. Classical machine learning algorithms such as ANN, SVM, KNN and others translate data into something of qualitative format as the first step.

D. Feature abstraction and selection

Because the quantity of structures can be predicted is mostly attributed to computational intricacy, an appropriate structure extraction and collection method is critical in the detection of dyslexia biomarkers. The goal of structure extraction and collection is used to create the most appropriate and useful characteristics possible from the initial data. This is performed to remove unnecessary data or sound from the tutoring database is inserted into machine learning anticipating models, also known as classifiers. Numerical or explicit features are possible. Phonological characteristics acquired manually from consistent exams, learners's scripting, quizzes, and test results have been selected and employed in many machine learning dyslexia noting investigations. MRI scans were used in many image-based machine learning experiments to obtain and implement characteristics related to brain tissue attributes. These data include volumetric information, geometric (shape-based) measurements, diffusion parameters, fractional anisotropy (FA), and activation patterns are examples of such features.

E. ML categorization and coaching

Methods of deep structured learning and ML are used to build and train models for dyslexia and associated biomarkers. Earlier research depended on a specific ML approach, but later studies explored hybrid dyslexia detection methods. Deep structured learning algorithms have been applied in the latest dyslexia & biomarker detection studies.

The kind of information utilized to train models that resulted, ultimately determines which machine learning algorithms are appropriate for dyslexia detection. As a result, rather than focusing on the performance of a single approach, possible research should compare the performance of several advanced ML techniques so as to investigate the results of various categories.

F. Performance evaluation

Several Toolboxes and frameworks functions, such as TensorFlow, Keras, train test split(), scikit-learn, along with LIBSVM, have been utilised in previous works to create predictive models using machine learning or deep learning approaches in Python, MATLAB, and WEKA environments. Different machine learning measures for dyslexia diagnosis are used to evaluate these models. This includes terms such as correctness, responsiveness, particular, sharpness, memorize, area under the ROC curve (AUC), mean square error (MSE), F- scores, positive predictive value (PPV), receiver operating characteristic (ROC) curve and negative predictive value (NPV). The proportion of dyslexic topics accurately categorized into being affirmative is known as accuracy. Sensitivity/responsiveness (TPR-True Positive Rate), refers to the percentage of dyslexic patients who are correctly identified. The true negative rate (TNR), also known as specificity, is the proportion of the data that is negative.

III. CONSIDERATIONS AND POSSIBLE OBSTACLES

ML algorithms are AI approaches that can learn patterns from information without having to specify it beforehand and have been progressively utilised to identify dyslexia over the past few years. The usual ML procedure for detection of dyslexia consists of: gathering of data, arranging data and pre-processing, feature abstraction and assortment, model coaching and duplication, and evaluation of performance. Regardless of the fact that it is indeed in its initial development stage, deep learning algorithms are being used in the field of dyslexia research due to the difficulties of traditional machine learning algorithms to analyse data in their native raw form, as well as other restrictions. Gathering dyslexia datasets from standardised psycho-educational exams and students' writing styles is quite cheap, according to the evidence from the evaluated studies. However, until the youngster enters school age, they may be difficult to obtain. The developing brain structure remains unchanged from childhood to adolescence, unless the child's brain is gravely harmed or critically ill, even though the sequence of stimulation changes as the child grows up due to the rapid synthesis of synapses.

IV. CONCLUSION

This research looks at current deep learning and ML techniques that are used to diagnose and predict dyslexia and its related biomarkers. Dyslexia, a complex brain developmental condition, has sparked a lot of attention in the field of modern neuroscience and ML in contemporary years. Despite the fact that significant ML approaches are deployed in this field in the past 20 years, deep structured learning algorithms are still in their infancy. Also, based on our research, SVM is the one of the most often used ML technique for detection of dyslexia and its prediction. Information for detecting and analysing this disorder has been gathered from a variety of sources.

REFERENCES

[1] A. Elnakib, A. Soliman, M. Nitzken, M. F. Casanova, G. Gimel'farb, and A. El-Baz,``Magnetic resonance imaging _ndings for dyslexia: A review,'' J. Biomed. Nanotechnol., vol. 10, no. 10, pp. 2778_2805, Oct. 2014, doi:10.1166/jbn.2014.1895.

[2] J. M. Fletcher et al., ``Classi_cation of learning disabilities: An evidencebased evaluation,'' in Identi_cation of Learning Disabilities: Research to Practice, R. Bradley, L. Danielson, and D. P. Hallahan, Eds. Washington, DC, USA: Erlbaum Associates Publishers, Jan. 2002, pp. 185_250.

[3] P. Tamboer, H. C. M. Vorst, S. Ghebreab, and H. S. Scholte, ``Machine learning and dyslexia: Classi_cation of individual structural neuroimaging scans of students with and without dyslexia,'' NeuroImage: Clin., vol. 11, pp. 508_514, 2016, doi: 10.1016/j.nicl.2016.03.014.

[4] S. O. Wajuihian, ``Neurobiology of developmental dyslexia Part 1: A review of evidence from autopsy and structural neuro- imaging studies,'' Optometry Vis. Develop., vol. 43, no. 3, pp. 121_131, 2012.

[5] S. O. Wajuihian and K. S. Naidoo,``Dyslexia: An overview,'' Afr. Vis. Eye Health, vol. 70, no. 2, pp. 89_98, Dec. 2011, doi: 10.4102/aveh.v70i2.102.

[6] B. Sarah, C. Nicole, H. Ardag, M. Madelyn, S.K. Holland, and H. Tzipi, ``An fMRI Study of a Dyslexia biomarker,'' J. Young Investigators, vol. 26, no. 1, pp. 1_4, 2014.

[7] 7. Association International Dyslexia, Baltimore, MD, USA. (2017). Dyslexia in the Classroom_What Every Teacher Needs to Know. [Online]. Available: www.DyslexiaIDA.org

[8] N. A. M. Yuzaidey, N. C. Din, M. Ahmad, N. Ibrahim, R. A. Razak, and D. Harun, ``Interventions for children with dyslexia: A review on current intervention methods,'' Med. J. Malaysia, vol. 73, no. 5, pp. 311_320, Oct. 2018.

[9] S. Kaisar, ``Developmental dyslexia detection using machine learning techniques: A survey,'' ICT Exp., vol. 6, no. 3, pp. 181_184, 2020, doi: 10.1016/j.icte.2020.05.006.

[10] F. Ramus, I. Altarelli, K. Jednoróg, J. Zhao, and L. Scotto di Covella, ``Neuroanatomy of developmental dyslexia: Pitfalls and promise,'' Neurosci. Biobehav. Rev., vol. 84, pp. 434_452, Jan. 2018, doi: 10.1016/j.neubiorev.2017.08.001.

[11] T. R. Vidyasagar, ``Visuo-spatial attention,'' Clin. Exp. Optom., vol. 87, no. 1, pp. 4_10, 2004.

[12] L. Danelli, M. Berlingeri, G. Bottini, F. Ferri,L. Vacchi, M. Sberna, and E. Paulesu,

[13] ``Neural intersections of the phonological, visual magnocellular and motor/cerebellar systems in normal readers: Implications for imaging studies on dyslexia,'' Hum. Brain Mapping, vol. 34, no. 10, pp. 2669_2687, Oct. 2013, doi: 10.1002/hbm.22098.

[14] U. Kuhl, N. E. Neef, I. Kraft, G. Schaadt, L. Dörr, J. Brauer, I. Czepezauer, B. Müller, A. Wilcke, H. Kirsten, F. Emmrich, J. Boltze, A. D. Friederici, and M. A. Skeide, ``The emergence of dyslexia in the developing brain,'' NeuroImage, vol. 211, May 2020, Art. no.116633, doi: 10.1016/j.neuroimage.2020.116633.

[15] S. M. Irby and R. G. Floyd, ``Test review: Wechsler abbreviated scale of intelligence, second edition,'' Can. J. School Psychol., vol.

28, no. 3, pp. 295_299, Sep. 2013, doi: 10.1177/0829573513493982.

[16] C. A. Dell, B. Harrold, and T. Dell, ``Test Review: Wilkinson, G. S., & Robertson, G. J. (2006). wide range achievement Test- Fourth

[17] Edition. Lutz, FL: Psychological Assessment Resources. WRAT4 Introductory Kit (includes manual, 25 test/response forms [blue and green], and accompanying test mat," Rehabil. Couns. Bull., vol. 52, no. 1, pp. 57_60, 2008, doi: 10.1177/0034355208320076.

[18] B. J.Wending, F. A. Schrank, and A. J. Scmitt, ``Educational interventions related to the Woodcock-Johnson III test of achievement," in Woodcock- Johnson III Assessment Service Bulletin Number, vol. 8, no. 8. Rolling Meadows IL, USA: Riverside Publishing, A Houghton Mif_in Company, 2007, pp. 1_30.

[19] J.-J. Mitchell, ``Comprehensive test of phonological processing," Assessment Effective Intervent, vol. 26, no. 3, pp. 57_63, Apr. 2001, doi: 10.1177/073724770102600305.

[20] R. M. Brunoand S. C. Walker, ``Comprehensive test of phonological processing (CTOPP)," Diagnostique, vol. 24, nos. 1_4, pp. 69_82, Mar. 1999, doi: 10.1177/153450849902401-408.