



Identification of Learning Disability using Machine Learning Techniques

Shilpa B. Kodli | Akshata

Department of Computer Science and Engineering (MCA), Visvesvaraya Technological University, Kalaburagi, Karnataka, India

To Cite this Article

Shilpa B. Kodli and Akshata. Identification of Learning Disability using Machine Learning Techniques. International Journal for Modern Trends in Science and Technology 2022, 8(11), pp. 145-150. <https://doi.org/10.46501/IJMTST0811025>

Article Info

Received: 08 October 2022; Accepted: 21 October 2022; Published: 17 November 2022.

ABSTRACT

Learning disability is a neurological disorder. The children with learning disabilities are neither slow nor intellectually retarded. Learning disability, Dyslexia is a neurological condition that is detected by uncertain cognizance of words and unfortunate understanding abilities. We proposed a machine learning model to analyse EEG signals from people with learning difficulties and provide results within a minute with the highest level of accuracy. The binary classification is performed using two machine learning classification algorithms, Random Forest Classifier and Linear Regression Classifier. Random Forest classifier algorithm shows better result as compared to Linear Regression Algorithm.

Keywords—E-learning, LD, Dyslexia, ML, Random Forest classifier and Linear Regression algorithm.

INTRODUCTION

Learning disability is a general term that describes specific kinds of learning problem. The LD affected children are neither slow nor mentally weak. An affected child has difficulty playing in games or reading or letter recognition. E-learning structure made uses of Models which is an open-source Learning Management System (LMS). This system recognizes two learning profiles for children's i.e., with Learning Disability (LD) and without Learning Disability (Non-LD) using committed courses arranged in view of various pieces of a LD child [7]. The work also multi stages of our approach for informal to get the learning limits for Dyslexia children's. The mainstage for instance data collection has two techniques where the main methodology connects with a smaller age group of 8-10 years with confined limits however the

resulting to connects with the age bundle 11-13 years for instance grades 6-8 with extra limits. The assessment of these responses has been acted in python language. To recognize whether the child has LD (Dyslexia in the case) or not, Machine Learning (ML) is used. Two ML computations to be explicit are RandomForestClassifier and Linear Regression are used to perform two-fold course of action with LD (1) and non-LD (0) as the two classes of the dataset [8]. The results are shown for both the philosophies and relative assessment shows that the dataset made in the last philosophy for finding limits including NLP is better and all the more remarkable. The equal classifier is performed using two AI RandomForestClassifier and Linear Regression are used. E-Learning has been spread out as a convincing means for learning various features like "Learning at your own

speed" and "getting confirmed from home". Regardless, the thought isn't very typical based on a brain of preparing for Learning Disabled child. LD is an issue impacting the framework. To spread it out doubtlessly, it is the different way by which a singular's brain works. Youths having LDs experience issues in scrutinizing, creating, thinking, evaluating, figuring out information or seeing in isolation [9]. They are like manner to find it hard to learn at a run of the mill or typical speed like others. We are focusing in on distinguishing and trying to address a couple of parts of Dyslexia which is inconvenience in examining and is one of the more routinely known LDs.

The examination of LD revolves around perceiving the conditions that impact the student's personal growth and legitimize the plan of explicit aides or uncommon organizations, for instance, varieties to the instruments they can use for explicit cycles, for example, changes to get to, help, intervention, and learning.

Learning Disability (LD) as demonstrated by the WHO's International Classification of Diseases the LD are known as "Express developmental issues of scholarly capacities" and are classification as follows:

- Unequivocal figuring out issue.
- Express spelling issue.
- Express strife of arithmetical capacities.
- Other developmental issues of scholastic capacities.
- Developmental disturbance of scholarly capacities.
- Mixed disturbance of scholarly capacities.

Machine learning is used for the development of system for the purpose of mathematical and scientific calculation and also for the automatic prediction of disease and generating the report. ML algorithms are shorted on the basis of the type of output of the dataset. The algorithms classify the output classes and detect whether an unknow sample is LD or not on the basis of the trained model.

2. LITERATURE REVIEW

Margaret Mary et. al. [1] proposed of predicting LD model which was developed using soft computing techniques. For this proposed system's performance had been improved by a hybrid approach. A combination of Markov model and Deep learning Neural Network was

done for classification and prediction. MATLAB was used to create the system. Outcome of the proposed approach showed the better efficiency of the proposed hybrid classification scheme.

De Clercq-Quaegebeur, et. al. [2] developed two methods to support the hypothesis that causes weak phonological representations affect aspects of mathematical involving the manipulation of verbal codes in individuals with dyslexia. First, dyslexia have weak phonological processing and weakness in aspects of mathematical that involve the manipulation of verbal code. Second, the neurological circuit that control verbal aspects of individuals with dyslexia. This indicates a high frequency of athematic deficits in children with dyslexia. These different pattern result from individual trajectories of development and arithmetic learning.

G Chakraborty and Vani [3] developed an evaluation of software and hardware options that can help dyslexic children. K Means, K Nearest Neighbor, Adaptive clustering, LS Algorithm, Support vector machine, and Human Markov Model are among the machine learning techniques employed. Machine learning algorithms are commonly used to extract speech features and determine accuracy and performance improvement in youngsters. As a result, the study will be focused on youngsters aged five to seven who are experiencing difficulty reading Hindi language words. To help dyslexic children, machine learning techniques has been used in the design process. Two and three letters are used here. If a youngster pronounces a word incorrectly the first time, he or she will be given another chance to read. If the same occurrence happens three times, the system will say the word out loud, accompanied with an image, so that the youngster will learn the term. The word uttered by the child. The same session will be repeated for the child for another 20 minutes with various words. Machine learning techniques will be utilized to recognize speech in this case. yields roughly 90% to 100% accuracy, and 30% accuracy if the system is tested with a new user. Rehman Ullah Khan et.al. [4] developed a framework containing information of 857 younger students scores in different tests. The information was gathered in one more principal examination of planning unique tests for dyslexia. This framework has three parts: the demonstrative module is a pre-screening application that

can be utilized by specialists, students and guardians for distinguishing the side effects of dyslexia. The subsequent module is characterization, which arranges the children into two non-dyslexia and for dyslexia. A third module is an investigation instrument for scientists. The outcomes show that 20.7% appear to be dyslexia out of 257 in the testing informational index which has affirmed by human master.

H. M. Al-Barhamtoshy et.al. [5] thought about individuals' troubles in perusing, spelling, composing and talking. Therefore, a computational investigation classifier had been accomplished utilizing dyslexia measurements strategies. Gibson trial of mind abilities will be utilized with impact of working memory, examining (hearing and discourse) and visual memory and comprehension, visual and evaluating discriminates, composing and coordinated movements, math and using time effectively, conduct, well-being, advancement and character, mental capacity in people groups with advancing exceptionally perusing troubles taken into our thought. Calculation investigation with classifiers will be utilized to examine the proposed dataset that incorporates 80 youngsters' records. This calculation model is planned and executed to assist with uncovering the basic issues that might influence figuring out how to perceive or compose as well as issues that may likewise cause issues with retains appreciation. This model is executed to assist guides and guardians with grasping the trouble and get child in the right path to schooling achievement.

Lateef et al. [6] developed a dyslexia diagnostic method that is informed by neural-biomarkers. Large numbers of machine learning method and deep learning methods have been implemented across various types of datasets with above-chance classification accuracy. Outcomes for each of the 22 selected articles using the preferred, Reporting Items for Systematic review and Meta-Analysis (RISMA) protocol, with the view to outlining some critical challenges for achieving high accuracy and reliability of the state-of-the-art machine learning methods. Therefore, higher classification relevance can be achieved using deep learning models for dyslexia and its biomarkers by addressing identified potential challenges.

3. PROPOSED SYSTEM

The aim of the proposed system work is to develop a prediction tool and thereby effectively predict LD and accurately measure its percentage and also determine the importance of preprocessing in classification. This strategy perceives two student profiles: Learning Disability (LD) and without Learning Disability (Non-LD). The work additionally incorporates a few stages of our casual testing approach for catching learning factors for Dyslexia. The main stage, information gathering, has two methodologies: the first is for a more modest age of 8-10 years with restricted boundaries, while the second is for a bigger age of 11-13 years (grades 6-8), with extra boundaries, Speech-to-Text (STT) change was performed on the sound reactions of the student utilizing Natural Language Processing (NLP).

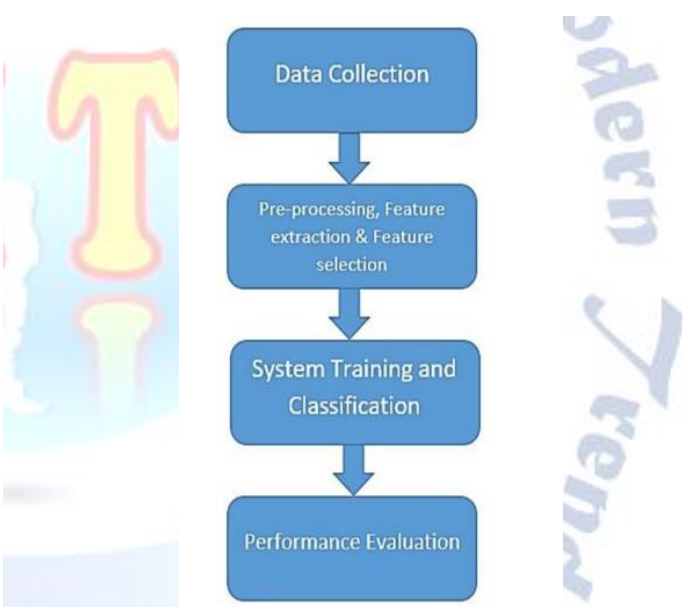


Figure 1: proposed system

The proposed system shown at figure 1 represent the classification of learning disability using machine learning techniques are used:

Data collection: The first step in dyslexia related dataset using techniques mentioned in introductory remark. Behavioral elements of participants during standardized tests, such as reading and writing, phonological awareness, and working memory, are examined by psychologists in traditional dyslexia procedures. Dyslexia people are identified by their low scores on these exams. However, because symptoms

differ between students, these procedures are generally time consuming and useless for a wide group of students. As a result, researchers adopt machine learning methodologies, which are less time intensive and often inexpensive.

Preprocessing, feature extraction and selection: feature extraction and selection process is an important task in the detection of dyslexia. The acquired information should be pre-handled and separated. This requires the information being changed over into either a quantitative (numbers) or subjective (text-based classifications) design. Preprocessing is utilized to track down applicable properties and dispense with invalid qualities. Following the pre-handling, the element extraction system should be performed, in which important highlights are found and a scope of values is designated. The following stage is to figure out which set of prevailing characteristics are most critical in characterizing the item's class.

System training and classification: for classification of dyslexia construction and training models in achieved using machine learning algorithms. The dataset trained the information such as dyslexia or non-dyslexia is contained in the training dataset. For this we have used the Random Forest Classifier.

Random forest is a sort of supervised algorithm machine learning method that is based on ensemble learning. Random forest classifier can be used to solve for regression or classification problem. The Random Forest algorithms is made up of collection of decision tree, and each tree in the ensemble is comprised of data sample drawn from a training set with replacement, called the bootstrap sample.

Performance Evaluation: python environments using different tools and libraries functions such as scikit-learn, NumPy, and TensorFlow in order to construct predictive models. Accuracy is used for evaluating the performance of dyslexia detection techniques using

4. RESULTS AND DISCUSSION

The proposed model has been implemented using python in Django framework. We have compared proposed Random Forest classifier's results with Linear Regression algorithm. Figures 2 to 6 show the results

obtained using Random Forest Classifier predicting weather child having Dyslexia or not.

The below figures are a series of screenshots of our results and their analysis

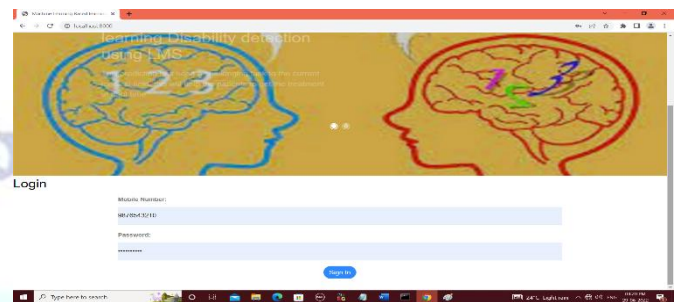


Figure: 2

This figure shows the login of Administrator of the application using mobile number and password to open the home page for the users.

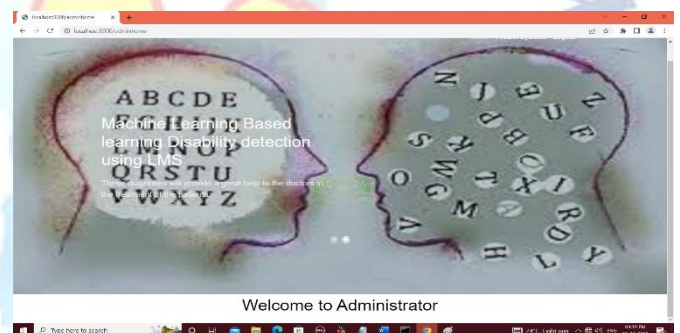


Figure: 3

This figure indicates the Administrator home page after admin login. Where only admin can have Access to this page

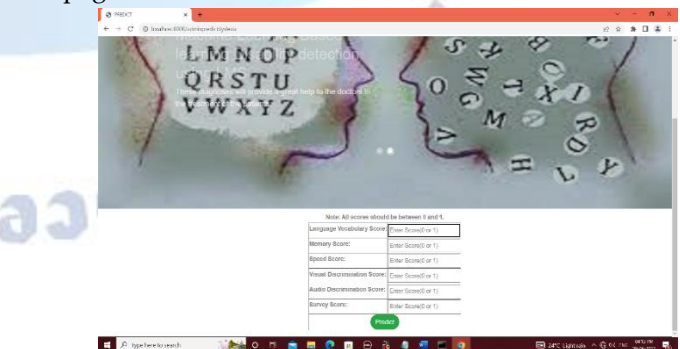


Figure: 4

This figure displays different test for different types of learning disabilities. The test case

includes language vocabulary Score, memory Score and speed Score and visual Discrimination score, Audio Discrimination and last is survey score with reference to dataset.

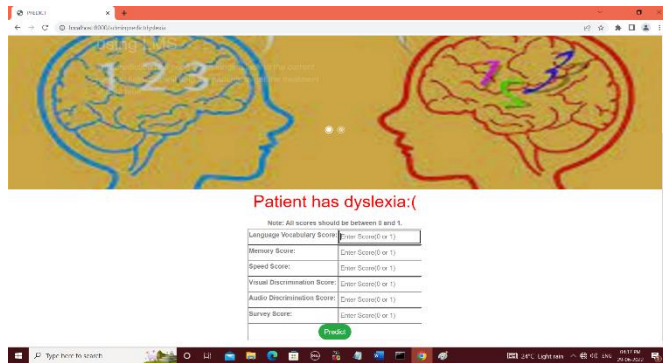


Figure: 5

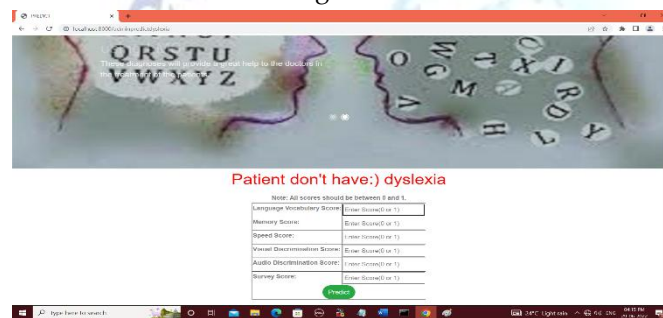


Figure: 6

Fig 5 and 6 shows the test results of all the cases. Here we can see a result in the form of 0s and 1s pattern, where 1 is used to identify dyslexic child and 0 is used to predict non dyslexic child



Figure 7: The Above Graph Shows the Using This

In fig.7, we have Compared prediction performance between Linear Regression and Random Forest Classifier whose accuracy is 43.89% and 93.66% respectively. From the graph it is clear that Random Forest Classifier achieves better results as compared to Linear Regression classifier.

5. CONCLUSION

On the basis of the work done so far, we can conclude that the child in between the age group 11-13 having Learning Disability can be solved by conducting the informal testing on the E-learning platform. It is a hassle free, simplified process as it saves the time and money spent on formal testing done in specific hospitals. We have Compared Prediction Performance between Random Forest Classifier and Linear Regression algorithm. From the results it is clear that Random Forest Classifier which predicts with an accuracy of 93.6% is better as compare to Linear Regression whose prediction accuracy is of 43.8%. So, if any unknown child has to be classified as LD (i.e., Dyslexic in our case) or non-LD, their characteristics are mapped with the trained model and predicted as Dyslexic or not. This system can be further expanded to the responses of the LD child through Optical Character Recognition (OCR) as many of them have low self-esteem because of which they do not have the confidence to express themselves freely or properly via speech. And this study can be further expanded to differentiate other learning disabilities such as Non-Verbal Learning Disabilities.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

- [1] Margaret Mary, T., and M. Hanumanth Appa." Hybrid classification approach hdlmm for learning disability prediction in school going children using data mining technique." Journal of Theoretical and Applied Information Technology 95.13 (2017): 2989- 2998.
- [2] De Clercq-Quaeghebeur, Maryse, et al. "Arithmetic abilities in children with developmental dyslexia: Performance on French ZAREKI-R test." Journal of learning disabilities 51.3 (2018): 236-249.
- [3] G Chakraborty, Vani. (2020). A survey paper on learning disability prediction using machine learning.
- [4] M. Modak, O. Warade, G. Sai prasad and S. Shekhar," Machine Learning based Learning Disability Detection using LMS," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), 2020, pp. 414-419, doi: 10.1109/ICCCA49541.2020.9250761.
- [5] Rehman Ullah Khan, Julia Lee Ai Chang, Oon Yin Bee, Machine learning and Dyslexia: Diagnostic and classification system (DCS) for kids with

learning disabilities, International Journal of Engineering and Technology 3.18(2018) 97- 10

- [6] H. M. Al-Barhamtoshy and D. M. Motaweh, "Diagnosis of Dyslexia using computation analysis," 2017 International Conference on Informatics, Health and Technology (ICIHT), 2017, pp. 1-7, doi: 10.1109/ICIHT.2017.7899141.
- [7] Lateef, Usman & Muniyandi, Ravie & Omar, Khairuddin & Mohamad, Mazlyfarina. (2021). Advance Machine Learning Methods for Dyslexia Biomarker Detection: A Review of Implementation Details and Challenges. IEEE Access. PP. 1-1. 10.1109/ACCESS.2021.3062709.
- [8] H. Perera, M. F. Shiratuddin, K. W. Wong and K. Fullarton, "EEG Signal Analysis of Real- Word Reading and Nonsense-Word Reading between Adults with Dyslexia and without Dyslexia," 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), 2017, pp. 73-78, doi: 10.1109/CBMS.2017.108.
- [9] AM, Dombrowski E, Shefcyk A, et al. Learning Disabilities Screening and Evaluation Guide for Low and Middle-Income Countries [Internet]. Research Triangle Park (NC): RTI Press; 2018 Apr. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK545498/> doi: 10.3768/rtipress.2018. op.0052.1804
- [10] David, Julie and Balakrishnan, Kannan. (2010). Machine Learning Approach for Prediction of Learning Disabilities in School-Age Children. International Journal of Computer Applications
- [11] David, Julie M, and Kannan Balakrishnan, "Learning disability prediction tool using ANN and ANFIS", soft computing 18.6(2014) 1093-1112
- [12] S. D, A. Hasan, A. Singh, H. Mansoor and R. N. Shaw, "Fault Prediction of Transformer Using Machine Learning and DGA," 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2020, pp. 1-5. doi: 10.1109/GUCON48875.2020.9231086.