EARLY DETECTION OF DYSLEXIA USING EFFICIENT PREDICTIVE APPROCH

Aditi Dhorkule[1], Saurabh Kure[2], Saurabh Mahajan[3], Geetesh Patil[4], Rajesh Phursule[5]

[1,2,3,4] *Department of Information Technology, Pimpri Chichwad College of Engineering, Pune, Maharashtra, India*
[5]*Professor, Dept. of Information Technology Engineering, Pimpri Chichwad College of Engineering, Pune, Maharashtra, India*

Abstract:

Dyslexia may be a learning disorder that happens mostly in children during their infancy. Children who have effected by dyslexia face difficulties while writing words, spelling, and reading words despite having average or above-average intelligence. As a consequence, dyslexic children often suffer from negative feelings, like low self-esteem, frustration, and anger. Hence, it is very important to detect dyslexia in children or peoples at their early age because later it can cause serious harm. Researchers have proposed a good range of techniques to detect developmental dyslexia, which incorporates game-based techniques, reading and writing tests, facial image capture and analysis, eye tracking, Magnetic reasoning imaging (MRI), and Electroencephalography (EEG) scans. Game-based techniques provide insights into reading and writing disorders. Via game-based techniques, we can take a quiz for the dyslexic students which include all the questions which we need to detect dyslexia in students. And by analyzing all the sections in the quiz we can calculate score of the quiz. Similarly we can include survey for the students or parents which include all the questions for which dyslexia is concern. By combining quiz and survey score we can detect whether the student has dyslexia or not. By decoding this information using machine learning (ML) we can provide differential analysis. This work developed using Dyslexia ML, by applying different machine learning algorithms to train our data we can predict labels for corresponding data and also a risk for dyslexia. We try different algorithms to train our data and test for new input which we get from quiz and survey scores. Based on these results we examined that the random forest algorithm achieves high accuracy compared to other algorithms which is about 97%. Based upon this we found that game-based technique works well for obtaining results of students whether he/she has dyslexia or not.

Keywords: **Dyslexia, EEG, reading difficulty, children, eye tracking, machine learning, screening**

---****---

Introduction:

Dyslexia is a most common learning disability. According to international dyslexia association (IDA) over 15-20% world population and this includes 10-17% school aged children's who have affected by dyslexia. Also according to review of dyslexia in Belgium 4000 students or 20-30% of students population diagnosed with dyslexia in higher education. Therefore it is important to detect dyslexia in early age of students or children's. As it may cause serious problems in their future. Students with dyslexia have poor fluency in reading, writing, spelling, speech and also other disorder including short-term memory. Because of this disorders affected peoples suffer from frustration, depression, loss of self-confidence and also they misbehave with others due lack of achievements as

compare to others. A person affected with dyslexia includes more than one behavior. Also in study it is found that 75% dyslexic students have more than one type of cognitive difficulties. There are two difficulties in dyslexia peoples one is cognitive which includes reading, writing and spelling disability and second is emotional which recognize dyslexia behavior.

Dyslexia categories into two parts, first is acquired dyslexia and second is developmental dyslexia. Acquired dyslexia is a condition when a normal person with a standard reading ability having a brain damage due to various reasons such as accident or stroke. Developmental dyslexia inherited through DNA. Dyslexia people have an average or above average IQ level. Hence it is very important to detect dyslexia at early age of peoples.

There are many methods to detect dyslexia such as using game based assignments, using eye tracking or eye detection, face recognition and using manual testing by giving them assignments. All this methods have their advantages but according to our research game based assessment using machine learning is best method to get accurate result of dyslexia. This includes quiz and survey to detect dyslexia and prediction done using machine learning model.

Detection of dyslexia using machine learning involves training and testing dataset using various algorithms which gives higher accuracy. Steps involve detecting dyslexia using machine learning as follows:

1. Data collection
2. Pre-processing, feature extraction and future selection
3. System training and classification
4. Performance evaluation

Literature Survey:

1. Paper title:

A Study of Computer-Based Learning Model for Students with Dyslexia

Findings:

Computer based learning model plays an important role to detect dyslexia among the peoples who are affected by more than one type of dyslexia or more than one type of behavior. This approach includes total 11 types of learning model to detect dyslexia based on learning model features. This approach helps to detect spelling, reading and writing disability and also used to predict the disorders like short-term memory. But this model do not emphasize on students behavior also does not give accurate result to address different difficulties of students with dyslexia suffer from more than one cognitive difficulty.

11 Learning models are there to detect dyslexia based on their feature as follows:

Model	Difficulty
Mylexic	• Reading • Writing
Dyslexia Baca	• Confusion
MyBaca	• Reading
Kast et al.	• Spelling
DysEggxi	• Spelling
E-Talk Pen	• Phonological Awareness
JollyMate	• Writing
IASD	• Reading • Writing
YUSR	• Phonic • Reading • Writing
Automatic Behavior	• Off-task behavior
iLearnRW	• Reading • Writing • Behavior

Table 1: *Models with difficulties of dyslexia peoples*

Above models target users as per learning models functionality and evaluation will be done.

2. Paper title:

Predictive Model for Dyslexia from Eye Fixation Events

Findings:

Dyslexia can be detected using eye fixation and saccades of peoples. Fixation is obtained using their eye movements. According to research dyslexic people have high eye fixation and saccades. Eye fixation is nothing but ability of person to focus on particular point and a rapid movement of the eye between fixation points called saccades. Dyslexic people fixation and saccades are different from non-dyslexic peoples. According our research dyslexic people has 380ms eye fixation. It is observed that people with dyslexia have large pause time while reading or writing. In this method reading of dyslexic peoples detected using eye gaze features. Also infrared video-oculography is used to calculate right and left eye pattern. Support vector machine (SVM) algorithm is used to record horizontal and vertical eye movements and recursive feature elimination-support vector machine used for feature reduction. 94% accuracy achieved using this algorithm. Dyslexic people have more unwanted eye saccades hence by using feature selection it can be reduced and also used to reduce over fitting and increase prediction accuracy. In this approach ober-2 and eye gaze patterns are used for capturing eye movement data. In this methods operations on dataset perform using machine learning algorithms such as support vector machine (SVM), K- nearest neighbor algorithms (KNN) and random forest algorithms. SVM algorithm includes polynomial functions like linear, radial basis function (RBF) and sigmoid achieves 92-93% accuracy. 95% accuracy achieved using KNN by observing eye fixation

of peoples. And random forest used to detect dyslexia using decision tree achieves 92% accuracy

	Machine Learning Algorithms for Dyslexia					
	Support Vector Machine (SVM)				KNN	RF
	Linear	RBF	Sigmoid	Hybrid		
Accuracy	0.93	0.93	0.92	0.93	0.94	0.91
Precision	0.92	0.90	0.91	0.90	0.93	0.94
Recall	0.94	0.97	0.94	0.96	0.96	0.88
Specificity	0.92	0.87	0.89	0.88	0.92	0.94
F1 score	0.93	0.93	0.92	0.93	0.95	0.90

Table 2: *Results of different algorithms for dyslexia*

3. Paper title:

A study of dyslexia using different machine learning algorithm with data mining techniques

Findings:

The method in this paper is used to identifying dyslexia based on checklist by using data mining. Checklist contains signs and symptoms of dyslexia using artificial neural network method with WEKA tool. Data mining is used with different types of algorithms for regression and classification. Data mining is well known method used for detection of various diseases. In this approach dyslexia detection is done using 10-cross validation algorithm with WEKA tool as it gives more accurate result than support

vector machine and KNN. WEKA is a collection of machine learning algorithms for data mining task. These algorithms can be applied by calling our main java code and also directly applied on dataset. WEKA tool plays an important role for dyslexia detection in training and testing dataset and also to remove unwanted data points. As WEKA is main tool for data classification, pre-processing, clustering, association rule, visualization and regression. WEKA is used for dyslexia using four features:

- Explore
- Experimenter
- Simple CLI(command line interface)
- Knowledge

This approach focuses on following points for more efficient technique of dyslexia detection to provide timely treatment:

- Use different algorithms with data mining to extract all the data points on the dataset for accuracy and prediction
- Compare performance of different data mining algorithms on dataset
- Select or identify best algorithm for prediction of problems

WEKA's statistical output helps to determine algorithms which have best performance based on data mining classification. As we discussed checklist are used to detect dyslexia using dyslexic people's characteristics. There are total 16 most beneficial characteristics of checklist which are taken from the assessment. Further unwanted checklist removed using data mining method. After that the prepared dataset by removing all unwanted elements load into ARAF format and then operation perform on this dataset using different algorithms like support vector machine(SVM), naive byes, decision tree and neural network. For better accuracy 10 cross validation algorithm is used with above algorithms.

4. Paper title:

Identifying Dyslexic Students by Using Artificial Neural Networks

Findings:

This paper shows the study of how ANN (Artificial Neural Network) can be used to identify dyslexia over the conventional diagnostic techniques. Paper consist sections as Conventional diagnostic techniques, early identification using ANN (Artificial Neural Network), Proposed implementation methodology, Results of the study

i.   Conventional diagnostic techniques:

Four steps for the conventional diagnostic of dyslexia:

- Identification of the dyslexic and non-dyslexic student by parents, teachers and other.
- Identification of the dyslexic and non-dyslexic student by junior level LD (learning disability) expert.
- Identification of the dyslexic and non-dyslexic student by senior level LD (learning disability) expert.
- Final diagnosis by the special education experts.

This conventional approach have some drawbacks .First is that the above steps is time taking and many resources are required for that. It is time consuming in nature and also there is no special expert psychologist or counselor at the school level for the identification of dyslexia.

And the reason is the lack of regulated standard at this level. Thus results vary from place to place. For overcoming these problems ANN can be used.

ii.   Early identification using ANN(Artificial Neural Network) :

Here the use of Neural Networks in identifying the dyslexic students is explored. Here the aim is to categorize

the students into one of the two categories i.e. dyslexic and non-dyslexic.

The inputs to the neural network are the results of the student, which include test of the reading ability, test of rapid naming, short term memory evaluation and sequencing skills, and non-word reading for the evaluation of phonological coding skills, an IQ test to setup profile of learning strengths and weaknesses.
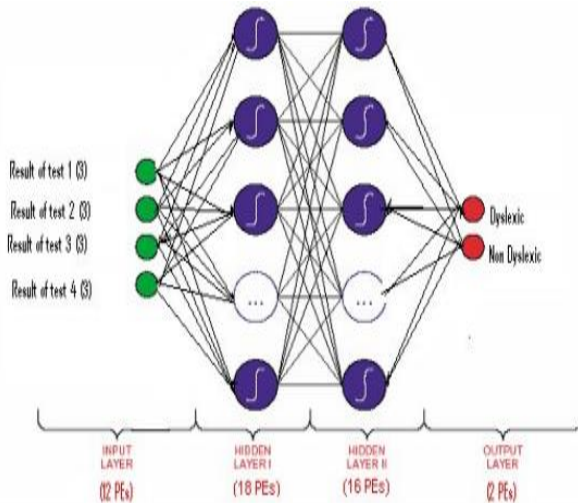
Fig 1: *Implementation of Multi layered perceptron model for dyslexic cases identification in children's*

iii.   Proposed Implementation Methodology :

First step is to collect student's school assessment data from last five years .If the student's score is continuously very poor then presume that dyslexia identification test is required .Next step is to fill the questionnaires provided with Neural Network model. Questionnaires are evaluated and result will be in the following range :

- Less than 33%.
- Between 33 and 45%
- Greater than 45%

These ranges are from the categorical independent variables. Then these are converted into 1-of-N binary representations. Thus we get of pseudo representation that increases the independent variable count from 4 to 12. All

pseudo representations of a categorical independent variable are given a value of 0 except the one that holds true for the current case, which is assigned the value 1. For e.g. by considering the test result of evolution test for mathematics deficit, if the scores of students less than 33% marks then input range for that will be set to 1 and for the remaining two to 0. This is how the classification model is obtained. This model is trained by using the supervised training mode. The output is classified in one of the two output categories Dyslexic or Non-Dyslexic.

iv.   Results :

This approach of using the neural network over the conventional method is highly beneficial for the aim of dyslexia .The test and trial basis implementation using test data evaluation gives maximum accuracy of 75 % .This model is beneficial to parents of students, teachers and school authorities, doctors/psychologists and young student

5. Page title:

Dyslexia detection using 3D convolutional Neural Networks and Functional Magnetic Resonance Imaging

Findings:

This paper is based on the approach for automatic detection of dyslexic student/child by using the 3D convolutional Neural Network. This approach is based on the classification of volumes containing brain activation areas during three different reading tasks.

- Lexical Decision Task
- Lexical/Orthographic Matching Task
- Semantic categorization Task

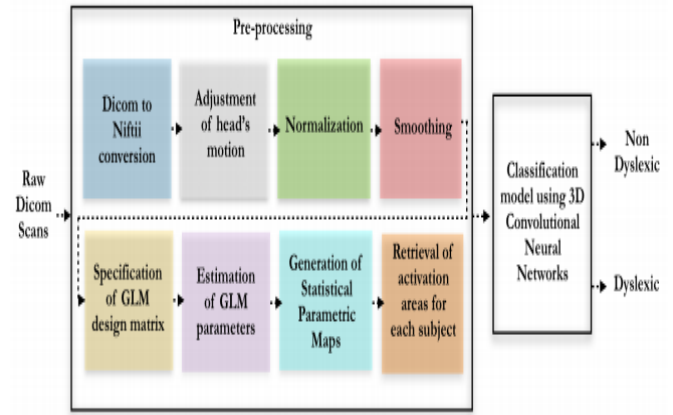For the study the students were divided into three groups as follow:

- Control Group (TDR)
- Dyslexia Group (DXR)

- Monocular Vision Group (MVR)

These three groups are based on inclusion and exclusion criteria listed below:

1. *Inclusion criteria:* The students were taken in the study only when they had diagnosis of dyslexia and did not receive the treatment. The students in the monocular vision group were those student who did not show any reading disabilities and did not show any literacy weakness due to ocular pathology.

2. *Exclusion criteria:* The student were not allowed to participate in the study if they had previously diseases like neurological disease ,head trauma , psychiatric illness ,chronic drug treatment , social deprivation ,inadequate schooling and intolerance to MRI scanning .
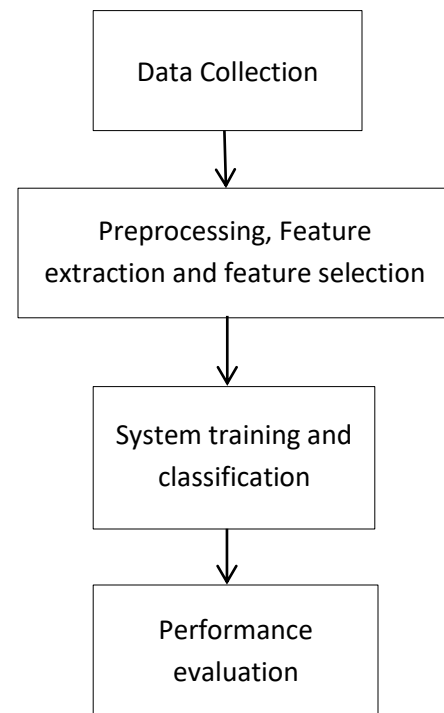
Proposed Framework:

1. The preprocessing step is the important step. The Preprocessing step depends on a series of transformation to bring all the subject brains into one single model and this step is conducted using SPM12 software with Mat lab 2018b environment.

2. After the preprocessing the 3D volumes containing activation areas retrieved and they are passed to 3D convolutional neural network which was trained to distinguish among the different activation areas such as dyslexic(DXR) or Non dyslexic(TDR and MVR).

Following are the preprocessing steps which were performed on the FMRI scan to bring all the subject brains into one single model:

- Dicom to Niftii conversion
- Adjustment of head motion
- Normalization
- Smoothing

- GLM design matrix and parameters
- Generation of statistical parametric maps
- Retrieval of activation areas
- 3D convolutional neural networks

Fig2: *FRI scan pre-processing steps*

Steps that are used for dyslexia detection:

Some methods from reference papers:

Efficient Data Analysis Using Computational Intelligence Techniques	Detection of dyslexia done by using computational tool such as artificial neural network (ARN). The intelligence technique used to detect disability in dyslexic peoples like reading and writing. ARN is disability tool for learning and it is a single layer perceptron. The detection is done successfully by training ARN model and result is achieved with approx. 90% accuracy.
Improved data capturing	Improved data capturing contain interactive multimedia method for analysis for dyslexia. Interactive multimedia is a screening based method. It reduce the use of paper based screening method or approach for dyslexia detection because of its efficient performance. This method is more beneficial and efficient than manual testing.
Eye movement Pattern	This method is more efficient for dyslexia detection. It captures the eye pattern of dyslexia people during reading for analysis of dyslexia or to detect dyslexic people. The basic approach involved in this method is the eye pattern is recognized by using eye fixation and saccade of eye movement by using different algorithms such as SVM and random forest. As it is found that dyslexic people have larger eye fixation that normal people. Eye fixation of dyslexia people is about approx. 350ms.
Computational Intelligence and Pattern Recognition Techniques	In this approach by using videography technique the capturing of eye movement pattern is done by performing unsupervised learning classification. This is done by using one verbal and two non-verbal task extraction from subjects. This method is highly efficient to capture the difference between dyslexia and non-dyslexic people and also analyzes the disabilities using self-organizing maps. This technique is one of the effective methods.

Electroencephalogram (EEG)	Electroencephalography, or EEG, is the physiological method of choice to record the electric activity generated by the brain via electrodes placed on the scalp surface. EEG is one of the fastest imaging techniques available as it is often has a high sampling rate. This technique is used for detection of different diseases and so for dyslexia.

Conclusion:

In this paper we have discussed various methods and techniques that we can use for prediction of dyslexia in early stage for the betterment of their feature. All different methods discussed above have their own functions and different methods/techniques or ways to successful accomplish the result that is dyslexia detection. Also all the methods discussed above have their features, advantages and disadvantages. From this survey we analyze best methods/techniques/algorithms that we can use for dyslexia detection with high accuracy.

References:

[1]. S. S. Abdul Hamid, N. Admodisastro and A. Kamaruddin, "A study of computer-based learning model for students with dyslexia," *2015 9th Malaysian Software Engineering Conference (MySEC)*, Kuala Lumpur, 2015, pp. 284-289, doi: 10.1109/MySEC.2015.7475234

[2]. A Jothi Prabha, R Bhargavi, Predictive Model for Dyslexia from Fixations and Saccadic Eye Movement Events, Computer Methods and Programs in Biomedicine, Volume 195, 2020, 105538, ISSN 0169-2607, https://doi.org/10.1016/j.cmpb.2020.105538.

[3]. H, Selvi; S, Saravanan M.. A Study of dyslexia using different machine learning algorithm with data mining techniques. International Journal of Engineering & Technology, [S.l.], v. 7, n. 4, p. 3406-3411, sep. 2018. ISSN 2227-524X. Available at: <https://www.sciencepubco.com/index.php/ijet/article/view/21691>. Date accessed: 26 feb. 2021. doi: http://dx.doi.org/10.14419/ijet.v7i4.21691.

[4]. Maitrei, Kohli & Prasad, T.. (2010). Identifying Dyslexic Students by Using Artificial Neural Networks. Lecture Notes in Engineering and Computer Science. 2183. 118.

[5]. Nilsson Benfatto M, Öqvist Seimyr G, Ygge J, Pansell T, Rydberg A, Jacobson C. Screening for Dyslexia Using Eye Tracking during Reading. *PLoS One*. 2016;11(12):e0165508. Published 2016 Dec 9. doi:10.1371/journal.pone.0165508

[6]. Appadurai, Jothi & Bhargavi, R.. (2019). Prediction of Dyslexia from Eye Movements Using Machine Learning. IETE Journal of Research. 1-10. 10.1080/03772063.2019.1622461.

[7]. Sofia Zahia, Begonya Garcia-Zapirain, Ibone Saralegui, Begoña Fernandez-Ruanova, Dyslexia detection using 3D convolutional neural networks and functional magnetic resonance imaging, Computer Methods and Programs in Biomedicine, Volume 197, 2020, 105726, ISSN 0169-2607, https://doi.org/10.1016/j.cmpb.2020.105726. (https://www.sciencedirect.com/science/article/pii/S0169260720315595)