**DYSLEXIA PREDICTION USING VARIOUS MACHINE LEARNING ALGORITHMS IN CLOUD**

*R.V. Keerthivasan*

*BSc (Hons) Cloud Computing and Big Data*

*School of Computer Science and Applications*

*REVA University*

*Email:* [*keerthivasanVA@gmail.com*](mailto:keerthivasanVA@gmail.com)

*R. Nathish*

*BSc (Hons) Cloud Computing and Big Data*

*School of Computer Science and Applications*

*REVA University*

*Email: nathishrajendran@gmail.com*

**Abstract:** Reading disability (known as Dyslexia) is a common problem faced by many children and younger people around the globe. There are various ways used to predict Dyslexia symptoms and habits with the help of using Machine Learning algorithms and Artificial Intelligence. The data generated to predict the Dyslexia problem from various people are stored using Cloud storage services. By applying various Machine Learning algorithms, the accuracy rate varies.

**Keywords:** Dyslexia, Decision-Tree, KNN, Structured, Unstructured.

I. Introduction:

In the Medical Industry, Reading Disability (Dyslexia) identification is a major problem to detect, because there are multiple reasons which causes this reading disability to the people.

The reading disability leads to lack of reading words, color co-ordination complexity, mismatching and assumptions of words. Dyslexia is also associated with the human eyes (vision power), majority of the people with low vision power are affected mainly and majority of the school Students from the age group of 5-12 are affected by reading disability, which can lead to lack of interest in studies and unhealthy lifestyle among the people.

In this digital era with the help of Machine Learning, predicting the accuracy of Dyslexia in the people has become possible. Identifying the causes and symptoms which lead to dyslexia helps medical professionals and people to predict and provide the necessary medications and treatments in the earlier stage and helps them to lead a better life.

Various algorithms are used in machine learning to fetch accuracy to predict the person affected by reading disability. Data are collected in different forms like structured, unstructured and semi structured data, these data can be used for training and testing the data. To process, analyze, preprocess and get the output from the collected data and there are different types of algorithms applied to these data, while using these algorithms there are time complexity and space complexity involved, which involves the time and memory that is taken by the algorithm to produce the result in a faster way and store the result.

There is various research under process to get a higher accuracy by reducing the time and space complexity with the various algorithms. But there is still a lack of best results among the different algorithms and data.

By opting the Decision Tree and KNN algorithms to the different types of data, the highest accuracy can be achieved for prediction. To store the data which is collected will be stored in the cloud to reduce the physical storage space and to maximize the availability of the resources.

II. Literature Review

There is much research that is conducted, and various results have been obtained by different research patterns. The varying results include the time and space complexity of prediction the Dyslexia. The research papers are got from various research and journal portals such as Scopus Index, ResearchGate, Academia and Google Scholar.

The research methodologies conducted by various research practitioners include major machine learning algorithms Decision Tree, KNN, SVM, Random Forest and logistic regression, which are used for getting better output. The first step is collecting data from various resources for testing the data, then the actual data is gathered from the medical industry, Google Forms and from various other resources to compare with the testing data to get the better results with less time and space complexity.

Eye co-ordination data is used to analyze the speed of the people’s view toward the words and pictures of various formats. The patterns which they can identify easier and the difficult patterns which take a long time for analyzing the data. Both sides angles of the eyes (left and right) are included in the study for getting better results.

While collecting the data, noisy data was also present, so the data has been cleaned and preprocessed into a particular format for the next steps. There was mostly positive data that was also presented with less accuracy for dyslexia.

The algorithms are implemented using Python with different development environments by analyzing the CSV files and responses from the people to predict the Dyslexia levels. There are algorithms which have performed well by providing the best accuracy rates with lower time and space complexity and vice versa.

III. Methodology

The data is collected from a variety of resources for testing the Pattern of Eye movements, reading speed, writing speed, word recognition, image recognition and speech recognition data are utilized for getting better accuracy. Developing the Python code using multiple built-in libraries which includes the Decision-Tree and KNN algorithms to analysis and produce a better output from the data.

Scikit-learn is a popular machine learning library used for performing the analysis and prediction from testing and training data.

**v) References**

|  |  |  |
| --- | --- | --- |
| 1 | 2022 | PREDICTION OF DYSLEXIA USING MACHINE LEARNING ALGORITHMS, [M. Mahalakshmi], [Dr. K. Merriliance 2] |
| 2 | 2022 | Dyslexia Prediction using Machine Learning [Isaac Punith Kumar], [ Hemanth Kumar B] |
| 3 | 2022 | Dyslexia Prediction [Pallavi Akhade 1], [Prof Sana Shaikh 2], [Aditi Ghaitade 3], [Tanuja Mukane 4], [Dipti Patil 5] |
| 4 | 2023 | PROJECT ON DYSLEXIA PREDICTION USING MACHINE LEARNING ALGORITHMS [Abhishek M\*1],[Hemanth Kumar BN\*2] |
| 5 | 2023 | Dyslexia Detection: Using Decision Tree Algorithm [Rishi Jain], [Pranav Patel], [Tanaya Jadhav], [Kimaya Joshi], [Ishita Jasuja] |
| 6 | 2023 | Dyslexia Prediction Using Machine Learning Algorithms - A Review [Kasthuri Stephen] |
| 7 | 2023 | Predicting Dyslexia with Machine Learning: A Comprehensive Review of Feature Selection, Algorithms,  and Evaluation Metrics [ Velmurugan S] |
| 8 | 2022 | Integrating oversampling and ensemble-based machine learning techniques for an imbalanced dataset in dyslexia screening tests [Shahriar Kaisar], [Abdullahi Chowdhury] |
| 9 | 2024 | An ensemble features aware machine learning model for detection and staging of dyslexia [Sailaja Mulakaluri1],  [Girisha Gowdra Shivappa2] |
| 10 | 2023 | A machine learning approach for dyslexia detection using Turkish audio records [TUĞBERK TAŞ], [MUHAMMED ABDULLAH], [BÜLBÜL ABAS], [HAŞİMOĞLU YAVUZ], [MERAL YASİNÇALIŞKAN] |
| 11 | 2022 | An Effective Feature Selection and Classification Technique Based on Ensemble Learning for Dyslexia Detection [Tabassum GullSajad], [Mohammad Khan] |
| 12 | 2021 | Fast Algorithm for Dyslexia Detection [Juraj Skunda's Lab], [Boris NerusilBoris Nerusil], [J. Polec], [Juraj Skunda] |
| 13 | 2021 | An Efficient Machine Learning Model for Prediction of Dyslexia from Eye Fixation Events [Jothi Prabha Appadurai], [R. Bhargavi], [B. Harish] |
| 14 | 2022 | Machine learning and deep learning performance in classifying dyslexic children’s electroencephalogram during writing [Ahmad Zuber], [Ahmad Zainuddin], [Wahidah Mansor], [Khuan Yoot Lee],[ Zulkifli Mahmoodin4] |
| 15 | 2021 | A Study on different Classification Models for predicting Dyslexia [Vani Chakraborty] |
| 16 | 2024 | Optimized k-nearest neighbours classifier based prediction of epileptic seizures [Himayavardhini Jagath Prasad 1], [Roji Marjorie S. 2] |
| 17 | 2017 | Medical Health Big Data Classification Based on KNNClassification Algorithm [Wenchao Xing], [Yilin Bei] |