

Machine Learning-Based Dyslexia Early Detection System

Karle Snehal Gokul, Nichit Komal Shivaji, Shirsath Sakshi Shankar

Department of Computer Science Modern College of Arts, Science and Commerce

Pune, India

ABSTRACT—Dyslexia is a learning disability where there are impairments in reading, writing, and spelling skills in children. This paper presents the development of a web application that performs dyslexia diagnosis using artificial intelligence techniques. In the developed system, behavioral measurements including reading speed, spelling accuracy, phonemic ability, attention, and reaction time are measured to determine dyslexia risk using supervised machine learning models. This web application is intended to be cost-effective and applicable in educational environments without any expensive infrastructure.

INDEX TERMS—Dyslexia, Machine Learning, Educational Technology, Diagnosis, Artificial Intelligence, Learning Disabilities.

I. INTRODUCTION

Dyslexia is considered to be one of the most common learning problems amongst students. Dyslexic individuals find it difficult to read with ease, spell words, and distinguish sounds from words. The dyslexic problem may hinder learning in schools since they lower academic achievements and affects confidence among children.

According to various research studies, a majority of children face dyslexic problems that hinder their ability to learn effectively. It is vital to identify dyslexia among children as it allows for effective learning intervention approaches.

The conventional diagnosis techniques of dyslexia rely on qualified psychologists and costly infrastructure in schools. Schools are not able to employ all of the necessary equipment and experts due to limited resources.

However, recent developments in machine learning technology have made it possible for designing computer systems that monitor students' behavior when learning. The current research work aims at developing an artificial intelligence system for screening dyslexic students.

II. LITERATURE REVIEW

Various scientists have mentioned the importance of intervention at an early stage of dyslexia.

Snowing described the strong relation between dyslexia and phonological processing issues. Shivwits showed how structured literacy instruction helps dyslexic children in improving their reading abilities.

Machine learning models have proved to be highly efficient in developing education prediction systems. The use of pattern recognition algorithms in order to analyze behavioral data was described by Bishop. Mitchell introduced the idea of systems capable of increasing prediction accuracy through learning.

The latest research works by Rollo and Zhou showed that digital behavioral and cognitive data can help to implement dyslexia prediction using machine learning methods. Still, most existing systems require sophisticated hardware and experimental environments. Thus, it is necessary to design browser-based applications which will be useful for school environments.

III. METHODOLOGY

The proposed system uses supervised machine learning algorithms to predict dyslexia risk. Behavioral data is collected using browser-based educational tasks focused on reading speed, spelling accuracy, phonemic awareness, response time, and attention consistency.

Students perform activities such as reading exercises, spelling tests, and sound-recognition tasks. The collected data undergoes preprocessing steps including normalization, feature scaling, and missing-value handling.

Machine learning algorithms such as Random Forest and Support Vector Machine (SVM) are used for classification. K-fold cross-validation is applied during model training to improve prediction reliability and reduce overfitting.

IV. SYSTEM ARCHITECTURE

The system architecture includes a frontend user interface, backend processing module, machine learning prediction engine, and database management system.

The frontend provides interactive activities for students, while the backend processes behavioral data and generates predictions. The machine learning module analyses educational indicators and estimates dyslexia risk levels in real time.

The proposed platform is lightweight and accessible through standard web browsers, making it suitable for schools with limited technical infrastructure.

V. RESULTS AND ANALYSIS

Experimental evaluation shows that the information obtained from behavioral markers via web-based educational tasks is indeed useful for predicting dyslexia.

The Random Forest classifier gave about 92% of accuracy during testing. The performance evaluation was done by employing well-known criteria like precision, recall, and F1 score.

Based on the results, we can conclude that ML-powered educational screening tools can help diagnose dyslexia in a cost-effective way.

VI. CONCLUSION

This paper outlines the development of a dyslexia early detection framework using machine learning, aimed at educational screening by way of behavioral study. The proposed web-based application offers a cost-effective

tool that helps identify potential cases of dyslexia among students requiring extra educational support.

The results emphasize the increasing use of artificial intelligence in educational technology.

VII. FUTURE WORK

Possible improvements could be the inclusion of deep learning systems, larger datasets for training, the development of mobile applications, and a multilingual environment for education. Collaborating with educational psychologists will help improve the accuracy of predictions made by the system.

REFERENCES

- [1] M. J. Snowling, *Dyslexia*, 2nd ed. Oxford, U.K.: Blackwell, 2000.
- [2] S. E. Shaywitz, *Overcoming Dyslexia*. New York, USA: Alfred A. Knopf, 2003.
- [3] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [4] T. M. Mitchell, *Machine Learning*. McGraw-Hill, 1997.
- [5] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [6] L. Rello et al., "Detecting Dyslexia Using Eye-Tracking Measures," *Proc. ACM SIGACCESS*, 2015.
- [7] L. Zhou and C. Yu, "Early Prediction of Developmental Dyslexia Using Machine Learning," *Frontiers in Psychology*, vol. 12, 2021.