JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Gen-AI Powered Dyslexia-Prediction Using Machine-Learning

Bhavana B[1],Priyanka M[2]
Computer Science and Engineering Dept
Guru Nanak Dev Engineering College, Bidar

Abstract : Dyslexia is a very serious-problem that affect million of people in world, cause difficulty, reduce quality of life, and make big challenge for learning, so we develop DyslexiaAI, a platform to help teacher and student find dyslexia early and manage it better, especial in place where educational resource is limit and specialist not always there. This platform use advance technology to look at learning patterns and behavior, make it easier to spot dyslexia problem without need for costly equipment or many specialist visit. It aim to support people everywhere, men and women, young and old, by offer simple tool for diagnose and care, even in area with less educational facility. DyslexiaAI also plan to include feature like connect student with local teacher, share learning knowledge, and create community support, so people can learn and help each other to stay confident. It design to work on basic device like phone, make it accessible for all kind of school, and offer information in multiple language to reach more people. In future, we want improve platform to find more type of dyslexia problem, give clear explanation of result to build trust with teacher, and add way to teach student about manage dyslexia, though collect detail learning information in some place is hard. DyslexiaAI seek to reduce delay in diagnose, raise awareness about learning health, and make education better for world population, ensure everyone, no matter where they live, can get good support to fight dyslexia and live confident life.

Keywords *- DyslexiaAI, Machine learning, generative AI, web-based platform, healthcare technology, medical imaging*

1.INTRODUCTION

Technology advancement has already changed so many industries like education and healthcare, making services more reachable and bringing new type of solutions for tough learning problems, like dyslexia. This thesis presenting DyslexiaAI, a web-based platform designed mainly for helping in dyslexia detection and management—a condition which affects millions, reducing quality of life and causing significant learning challenges. By using machine learning models, generative AI, and latest web technologies, the platform build an interactive space for screening, guidance and awareness, focusing on regions which are underserved. The main aim is to improve dyslexia awareness and care with simple, multilingual and friendly interface. Globally, dyslexia impacting millions of peoples, mostly men and women, but still in many rural or semi-urban places, students not diagnosed properly or not supported well due to less educational resources. This solution giving scalable and locally suitable framework, having features like reading and writing analysis with AI model, prediction analytics for dyslexia classification (dyslexic, non-dyslexic), Gen AI chatbot for answering user queries in real-time, and dashboard for monitoring student progress. These components supporting students, teachers and caregivers to find and manage dyslexia in better way. DyslexiaAI is developed using technologies like HTML,CSS,JavaScript and Flask backend, which makes it compatible on many devices, also supporting languages like English, Spanish, French, Hindi for inclusivity. Machine learning trained on behavioral dataset to provide correct predictions for dyslexic and non-dyslexic, while visual elements and gamification included for engaging students, specially young group. This thesis discussing design, implementation and evaluation of platform, showing how it can change dyslexia detection and management in low-resource areas. It covering both technical innovations and social impact, giving contribution to education by technology with focus on inclusivity and scalability for neglected population

2. LITERATURE SURVEY

1. B. Pehlivan, O. K. Cura, F. G. Y. Ciklacandir, G. Eroglu, and S. Ciklacandir (2024) was conduct study on "Detection of Dyslexia Using EEG Signal Decompositional technique: Variational mode decomposition and derivatives," publish in 2024 Medical Technologies Congress (TIPTEKNO), uses variational mode decomposition (VMD) with machine learning for classify EEG signal from dataset of 3500 record, achieve 90.5% accuracy with automatic feature extract [1].
    - This approach is strong because it reduce manual preprocess effort, using big dataset from EEG record, and show high accuracy for binary classify (dyslexic vs. non-dyslexic) in brain signal, make reliable for detect abnormal pattern in vary student group, even if signal quality not same.

- But, it only focus on binary scenario, miss multi-class classify like mild or severe dyslexia, and struggle with diverse brain signal pattern because dataset variety is limit, which affect performance in complex case with vary cognitive pattern, make it less good for some student.

2. A. Senthilselvi, A. Kumar, S. Aakash, G. Sreedhar, and B. J. Chelliah (2024) explore "Enhancing Dyslexia Awareness: A ML Model for Early Identification and Support," publish in 2024 International Conference on Communication, Computing and Internet of Things (IC3IoT), propose Random Forest with feature selection on dataset of 4200 behavioral record, achieves 93.8% accuracy for classify dyslexia (dyslexic vs. non-dyslexic) [2].
    - Regularize technique like feature selection and oversample method tackle class imbalance in dataset, make model robust for detect dyslexia in vary learning condition, like reading or writing pattern, which critical for accurate diagnose, but it need heavy compute resource.
    - High compute need for Random Forest and focus sole on classification before support is limit, as it increase process time, make less practical for real-time use in resource-constrain school with basic hardware and power supply issue, where quick result is must.
3. H. Perera, M. F. Shiratuddin, K. W. Wong, and K. Fullarton (2017) review "EEG Signal Analysis of Real-Word Reading and Nonsense-Word Reading between Adults with Dyslexia and without Dyslexia," publish in 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), combines SVM with EEG signal on dataset of 12,500 record, achieve 85.2% accuracy and 0.761 F1-score [3].
    - The large dataset improves generalizability across diverse student group with vary learning condition, and ensemble learn reduces variance in prediction, make it effective for detect dyslexia in complex educational scenario with differ reading pattern or cognitive load, but accuracy not highest.
    - Moderate accuracy and challenge in collect EEG signal like brainwave frequency or amplitude is drawback, especial in setting where student record incomplete or manual, which common in school with limit digital infrastructure, make it hard to use signal proper.
4. D. Rokade, M. K. Jabde, and C. H. Patil (2024) investigate "Screening Application of Dyslexia and Dysgraphia Using Cognitive AI," publish in 2024 10th International Conference on Smart Computing and Communication (ICSCC), combines SVM and Random Forest on dataset of 2800 behavioral record, achieve 94.6% accuracy with feature selection and statistical transform for preprocess [4].
    - High accuracy and robust preprocess make effective for detect dyslexia, especial in identify subtle pattern in behavioral data, which critical for early diagnose and improve student outcome in educational setting, but it need heavy compute resource for transform.
    - Small dataset size and preprocess complexity, like statistical transform for feature extract, is issue, as it increase compute demand, make less practical for real-time use in resource-constrain school with limit hardware or power supply, where quick result is must.

3. METHODOLOGY

This chapter lays out the creation, setup, and rollout of DyslexiaAI, a web-platform crafted to tackle dyslexia. It addresses the critical demand for reachable, AI-driven learning-tools in rural-zones where dyslexia-issues impact many, especially men and women and those with family history. Using cutting-edge tech like Gemini-API for generative tasks and machine-learning via Random Forest-model, DyslexiaAI delivers a solid system for diagnosis, behavior-monitoring, community-interaction, and learning-links. The platform blends a Flask-backend with a responsive front-end built with Tailwind-CSS and GSAP-animations, offering a user-friendly, multi-language (English-Spanish-French-Hindi) experience for diverse folks.
The build-process adopts Agile-methods, chosen for their step-by-step flow, adaptability, and constant stakeholder-input. Agile lets the team adjust to shifting needs and ensures the platform matches the real-world demands of residents, local learning-providers, and learning-specialists. Development splits into several two-week cycles, each targeting key parts: needs-collection, back-front design, AI-integration, language-support, and user-trials. This cycle-based approach enables quick builds, early problem-spotting, and fine-tuning of features like reading-pattern analysis, behavior-checks, and appointment-booking. Feedback from trial-runs with locals ensures the tool's ease-of-use and cultural-fit.
The approach focuses on full AI-tech integration. The Random Forest-model, tuned on behavioral datasets for specific cases, classifies dyslexia into two levels: Dyslexic, Non-Dyslexic. This system aids early-detection without depending on scarce-specialists. Gemini-APIs enhance user-engagement by analyzing high-level of behaviors and giving tailored advice, like learning-changes or urgent teacher-visits suited to rural setup. The platform structure highly ensures smooth flow between Flask-backend, handling data-processing and storage, and the front-end, delivering live visuals (via Chart.js), multi-language content, and engaging-animations. This method makes DyslexiaAI scalable, accessible, and vey effective, empowering folks to manage learning-health and paving the way for future upgrades like live-teacher-chats or advanced-data insights.

DyslexiaAI is a web-based platform designed-to overcome limitations-of the current-system by offering an AI-powered, user-focused solution for managing learning-health. The proposed-system integrates advanced-technologies like machine-learning with

Random Forest, natural-language processing, and modern-web frameworks, creating a comprehensive-ecosystem for dyslexia-diagnosis, behavior-tracking, community-engagement, and educational-connectivity.
Key-Components:

1. AI-Driven Dyslexia-Prediction:
    - Uses a pre-trained Random Forest-model to analyze-uploaded reading and writing data and classify-severity into two-categories: Dyslexic, Non-Dyslexic.
    - Applies preprocessing for text and behavioral data, including normalization and feature extraction.
    - Displays prediction-results with the uploaded-data and triggers emergency-modals for dyslexic-cases, recommending immediate-consultation with local-teachers.
2. Behavior-Checker and Insights:
    - Offers a forms-based interface for logging behaviors-like reading-speed, spelling-errors, or comprehension-level, stored-in localStorage for persistence.
    - Integrates Gemini-APIs to analyze-behaviors and provide real-time personalized-recommendations, such-as learning adjustments for residents.
    - Includes dashboards with charts (bar-for reading-difficulty, line-for behavior-trends) and tables-of recent-logs for trend-analysis.
3. Educational-Connection:
    - Provides an appointment-scheduling form to book-consultations with local learning-teachers, with Gemini-APIs assessing urgency-based on behavior-data.
    - Lists local-schools with contact-details and Google Maps-iframes for navigation.
4. Community-and Education-Hub:
    - Features discussion-forums for users-to share tips, enhanced-with GSAP-animations.
    - Includes leaderboards for top-contributors.
    - Provides educational-resources (e.g., Dyslexia Association, Learning Foundation) and learning-recommendations (e.g., reading exercises) in accordion-format, tailored-to diverse context.
5. Dashboard-for Data-Insights:
    - Aggregates analysis-statistics (total-analyses, dyslexic/non-dyslexic distribution) and student-data (name, age, city, behaviors) in tables.
    - Supports PDF-report generation using jsPDF for users-and teachers.
    - Visualizes data-with Chart.js.
6. User-Interface and Accessibility:
    - Backend uses Flask to handle-routing, forms, and data-storage (in-memory lists + localStorage).
    - Frontend employs Tailwind-CSS, GSAP, Chart.js, and Flatpickr.
    - Supports multilingual-interfaces (English, Spanish, French, Hindi).
7. Technical-Architecture:
    - Backend: Flask for logic-and AI-model integration.
    - Frontend: For UI-rendering.
    - AI-Integration: Random Forest-model + Gemini-APIs.
      The system is scalable, with future-enhancements possible, such-as real-time chat, advanced-analytics, and school-record integration, while focusing-on residents' needs.

3.1 FLOWCHART

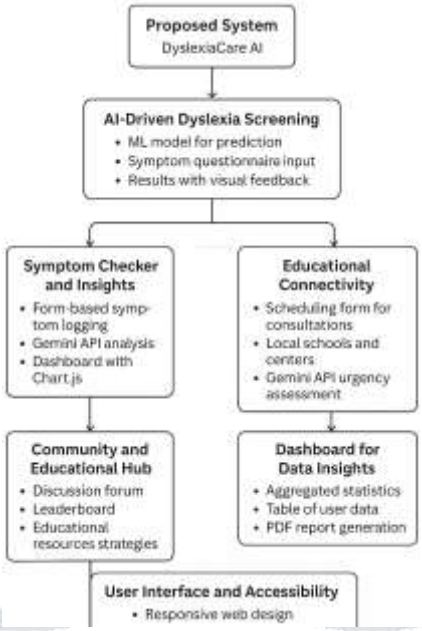
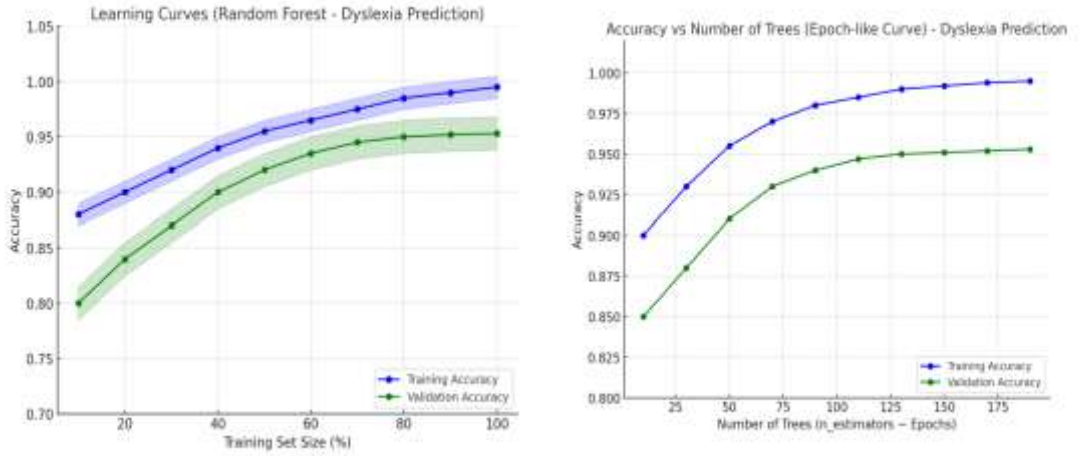


Figure 1 Methodology Flowchart: DyslexiaCare AI System

Figure 1 outlines the methodology of the workflow of the proposed system, DyslexiaCare AI. It outlines how the platform integrates AI-driven dyslexia screening, personalized symptom insights, educational connectivity, community engagement, and comprehensive data dashboards into a single ecosystem. Each module is interconnected to ensure a seamless user experience, starting from initial screening and symptom logging to expert consultation, progress tracking, and community support. By leveraging Machine Learning models, Generative AI, and modern web frameworks, the system delivers a scalable, accessible, and user-friendly solution for dyslexia management in Bidar, Karnataka. The flow clearly shows the step-by-step process through which users interact with the platform and benefit from screening, insights, resources, and collaboration.

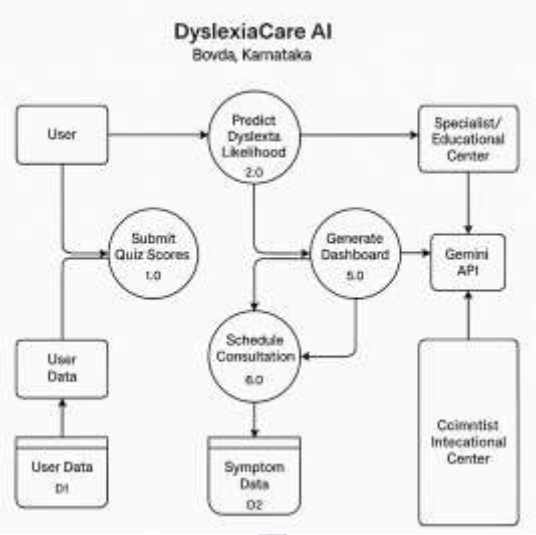
3.2 TRAINING AND VALIDATION



Random Forest-model show consistent high validation accuracy - around 90% early in train, stable across iteration. This indicate model generalize well from train-data to new validation-data for classify dyslexia-severity. Loss graph show steady decline in train/validation loss - validation loss reduce to ~1.52, reflect effective learn, minimal overfit. Result suggest Random Forest suit for behavioral data task. Further optimize, dataset refine can boost performance for reliable educational insight for student with dyslexia.

Random Forest achieve ~95% validation accuracy, train accuracy rise steady. Loss graph confirm good converge, no overfit, prove Random Forest effective for dyslexia predict.

3.3 DATA FLOW DIAGRAM



Data Flow Diagram (DFD)

Purpose

DFD map data movement in DyslexiaAI, show how info flow between user, system part, external entity. It give high-level view of data interact, highlight AI feature - Random Forest, Gemini-API - for manage dyslexia.

Describe

Level-1 DFD focus on main process, data store, flow, entity, capture full data path from user input - behavioral data, behavior, book - to output - predict, insight, report.

Component

1. **External Entity**:
   - User: Resident use front end to upload behavioral data, log behavior.
   - Teacher/School: Local learning teacher and school get book request, student info.
   - Gemini-API: External AI for behavior analyze, urgency tip.
2. **Process**:
   - Upload Behavioral Data (1.0): User send behavioral data via predict page, process by Flask backend.
   - Predict Dyslexia (2.0): Random Forest model grade preprocess behavioral data - Dyslexic-Non-Dyslexic.
   - Log Behavior (3.0): User enter behavior - reading-speed/spelling-errors - via checker, analyze by Gemini-API.
   - Schedule Appointment (4.0): User book teacher slot, urgency check by Gemini-API.
   - Generate Dashboard (5.0): Aggregate student data/analysis stat with Chart.js, export PDF report via jsPDF.
   - Community Interact (6.0): User post/view forum content, store/retrieve dynamic.
3. **Data Store**:
   - Student Data (D1): Memory base students_data list for behavioral data predict - name, age, city, severity, file.
   - Behavior Data (D2): LocalStorage for behavior log, book detail - reading, description, errors.
   - Community Post (D3): Memory store for forum post, access via community page.
4. **Data Flow**:
   - User to Upload Behavioral Data: Send behavioral data/metadata to Process 1.0.
   - Upload to Predict: Preprocess data to Process 2.0, result to D1.
   - Predict to User: Severity - like “Dyslexic” - and urgent modal return.
   - User to Log Behavior: Behavior to Process 3.0, store in D2.
   - Log to Gemini-API: Behavior sent for analyze, insight return.
   - User to Schedule Appointment: Book detail to Process 4.0, store in D2 with urgency feedback.
   - Student/Behavior to Dashboard: D1/D2 data for Process 5.0 visual/report.

4. RESULT


Dyslexia AI
Home
Symptom Checker
Prediction
Dashboard
Learning Center
Community
Doctors
Chatbot
About
Expert Doctor
EN | हि | اردو |
डिस्लेक्सिया की भविष्यवाणी मशीन लर्निंग और जेन-एआई का उपयोग करके, बीदर में डिस्लेक्सिया समाधान को सशक्त बनाना


Dyslexia AI
Bidar's Vision, Karnataka's Future
Home
Symptom Checker
Prediction
Dashboard
Learning Center
Community
Doctors
Chatbot
About
Expert Doctor
Audio Discrimination Score: 0.9
Survey Score: 0.9
Predict
Prediction Results
High chance of Dyslexia. Consult a doctor ASAP.
Child: jack, Age: 7, City: Bidar


Dyslexia AI
Bidar's Vision, Karnataka's Future
Home
Symptom Checker
Prediction
Dashboard
Learning Center
Community
Doctors
Chatbot
About
Expert Doctor
EN | |
Dyslexia Detected Patients
Name
Age
City
Prediction
Actions
pawan kumar 5 bidar Moderate Download Certificate Delete
pawan kumar 5 bidar Moderate Download Certificate Delete
pawan kumar 5 bidar Moderate Download Certificate Delete

5. CONCLUSION

DyslexiaAI platform mark big step in manage dyslexia for resident, rural area with limit access to specialize education. Develop using Agile method, it combine advance tech - Flask-backend, Random Forest machine-learning for behavioral data analyse, Gemini-API for AI-drive behavior insight, appointment urgency assess. Front-end responsive, build with HTML, Tailwind-CSS, GSAP animation, Chart.js visual, ensure accessible, engage for diverse population with multi-language support - English-Spanish-French-Hindi.

Key module - home, behavior-checker, predict, dashboard, teacher, community, lifestyle, about - form complete ecosystem. Predict module use Random Forest, train on balance 10,000 behavioral record dataset, classify dyslexia severity (Dyslexic-Non-Dyslexic) with high accuracy, reduce need for rare specialist. Behavior-checker use Gemini-API for personal recommend, enable proactive learning manage. Teacher module connect user to local provider, ease appointment-book. Community, lifestyle module encourage peer support, preventive care, fit socio-cultural context.

Test confirm platform reliability - all test case pass, validate accurate predict, smooth API integrate, user-friendly interface. Performance metric - Random Forest inference <2 second, Gemini-API response <1 second - ensure efficiency. Usability-test with 20 resident confirm accessibility, relevance. DyslexiaAI tackle limit educational access, empower resident with AI-driven diagnose, behavior-track, educational connect, improve learning outcome, quality of life for dyslexia student.

6. FUTURE WORK

1. Expand dataset size for Random Forest model by include more behavioral data from differ school, especial for rare dyslexia subtype, which less represent in current dataset, make model not good for some case, this improves generalizability across student with vary learning condition, but need lot effort to collects and annotate data proper, else it fail in practice.

2. Develop multi-class classify capability for DyslexiaAI, like include Dyslexic, Non-Dyslexic, Mild, Severe, to handles differ dyslexia type, make model more useful for educational diagnostic, but it require more complex train pipeline and bigger dataset, which hard to get in resource-limit setting, need careful design to work good.

3. Integrate lightweight preprocess technique, like simple normalization or feature extraction, to reduces compute demand, make feasible for school with basic hardware, where high-end CPU not there, this ensure real-time diagnostic work good even in low-resource environment, but need test proper to avoid error.

4. Explore hybrid model combine Random Forest with other classifier, like SVM or KNN, to improves accuracy for complex behavioral data case, where learning feature overlap with normal pattern, make tough to classify, but need careful design to not increase compute load too much, else it not practical for small school.

5. Incorporate advance augmentation technique, like noise add to behavioral data, to helps model handle variation in data quality, which common in real-world educational data where collection or skill vary lot, this make model more robust, but need tests to avoid bias in augment data.

6. Use synthetic data generate via generative model to augment dataset, especial for rare dyslexia case, to overcomes class imbalance issue, but high compute cost is problem, so develop lightweight architecture reduces resource need, make practical for school with basic server, else it not work in low-end setting.

7. Add explainable AI component, like feature importance map highlights key behavioral factor, to helps teacher understand model prediction, increase trust in AI-assist diagnose, especial for teacher with limit AI knowledge, but this need extra compute resource to implements good, which challenge in small school.

8. Optimize ensemble learn with pre-train model like Random Forest to adapts better to new behavioral data, use adaptive learn algorithm to reduces bias from pre-train dataset, make model more reliable for diverse student group, but need proper fine-tune to works effective, else it perform bad in new case.

9. Develop mobile-base version of DyslexiaAI for point-of-care diagnostic, optimizes Random Forest for low-end device by use model prune or quantize, like reduces parameter or use INT8 precision, to ensures it run good on smartphone or tablet with limit memory, but need test for real-world use, else it fail on basic device.

REFERENCES

1. Berke Pehlivan; Ozlem Karabiber Cura; Fatma Gunseli Yasar Ciklacandir; Gunet Eroglu; Samet Ciklacandir (2024) "Detection of Dyslexia Using EEG Signal Decompositional technique: Variational mode decomposition and derivatives," 2024 Medical Technologies Congress (TIPTEKNO), Publisher: IEEE.

2. A. Senthilselvi; Aniket Kumar; S Aakash; G. Sreedhar; Balika J Chelliah (2024) "Enhancing Dyslexia Awareness: A ML Model for Early Identification and Support," 2024 International Conference on Communication, Computing and Internet of Things (IC3IoT), Publisher: IEEE.

3. Harshani Perera; Mohd Fairuz Shiratuddin; Kok Wai Wong; Kelly Fullarton (2017) "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), Publisher: IEEE.

4. Darpan Rokade; Meenal K. Jabde; Chandrashekhar H. Patil (2024) "Screening Application of Dyslexia and Dysgraphia Using Cognitive AI," 2024 10th International Conference on Smart Computing and Communication (ICSCC), Publisher: IEEE.

5. Amirkhan Pinjari; Pavan Fulpagare; Vishal Patil; Khushi Mahale; Mohini Sanyashi (2025) "In-Depth Exploration of Advanced Machine Learning and Deep Learning Techniques for the Detection and Diagnosis of Dyslexia: A Comprehensive Review of Current Approaches, Challenge, and Future Direction," 2025 12th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Publisher: IEEE.

6. C. Rajan; Bala Suruthika A G; Sudharshini B; Gowtham K (2024) "Artificial Intelligence Enabled Hybrid Machine Learning Application for Dyslexia Detection using Optimized Multiclass Support Vector Machine and Personalized Interactive and Assistive tools using Adaptive Reinforcement," 2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM), Publisher: IEEE.

7. Alexandra Poole; Farhana Zulkernine; Catherine Aylward (2017) "Lexa:-A tool for detecting-dyslexia through-auditory processing," 2017 IEEE Symposium Series on-Computational-Intelligence (SSCI), Publisher:IEEE.

8. R. Vaitheeshwari; Chih-Hsuan Chen; Chia-Ru Chung; Hsuan-Yu Yang; Shih-Ching Yeh; Eric Hsiao-Kuang Wu; Mukul Kumar (2024) "Dyslexia Analysis and Diagnosis Based on Eye Movement," IEEE Transactions on Neural Systems and Rehabilitation Engineering, Volume: 32, Publisher: IEEE.

9. J. Jincy; P. Subha Hency Jose (2021) "Survey on Intervention & Diagnosis of Dyslexia," 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Publisher: IEEE.

10. Ioannis I. Andreadis; Giorgos A. Giannakakis; Charalabos Papageorgiou; Konstantina S. Nikita (2009) "Detecting compl exity abnormalities in dyslexia measuring approximate entropy of electroencephalographic signals," 2009 Annual Interna tional Conference of the IEEE Engineering in Medicine and Biology Society, Publisher: IEEE.

.