WEB BASED ML/DL SYSTEM FOR EARLY DETECTION OF LEARNING DISABILITIES: DYSLEXIA, DYSCALCULIA,

Dr. Dayanand J, Arham Fatima, Md Nashaj Omer, Md Muzammil and *Saniya Khan*

Dr. Dayanand J / Head of Department AIML GNDEC Bidar

Arham Fatima / Student Dept AIML GNDEC Bidar

Md Nashaj Omer / Student Dept AIML GNDEC Bidar

Md Muzammil / Student Dept AIML GNDEC Bidar

Saniya Khan / Student Dept AIML GNDEC Bidar

Abstract— This research addresses the critical need for scalable, early-intervention tools for child development by presenting an intelligent web-based system for detecting dyslexia, dyscalculia, and dysgraphia. The platform utilizes a hybrid architecture: a Random Forest Classifier for analyzing numerical reasoning through quiz-based inputs, and two Convolutional Neural Network (CNN) models for evaluating handwriting samples. By integrating OpenCV for image preprocessing and leveraging datasets from Kaggle and Mendeley, the system provides high-accuracy predictions and personalized subject recommendations. The results demonstrate a significant reduction in the time and cost associated with traditional clinical assessments.

Index Terms— Machine Learning, Deep Learning, Dyslexia, Dyscalculia, Dysgraphia, CNN, Random Forest, OpenCV.

I. INTRODUCTION

Learning disabilities such as dyslexia (reading), dyscalculia (mathematics), and dysgraphia (writing) are neurological disorders that can severely impact a child's academic and emotional growth if left undiagnosed. Traditional diagnostic methods often require expensive, specialized clinical evaluations that are inaccessible to many, particularly in rural or under-resourced areas. This project introduces an AI-driven solution that automates the screening process through a Flask web framework. By combining Machine Learning (ML) for behavioral data and Deep Learning (DL) for visual pattern recognition, the system offers an accessible, real-time diagnostic tool for parents and educators.

SYSTEM OVERVIEW

The system is designed as a modular web application that streamlines the diagnostic workflow. Key components include:

• **Frontend:** A responsive interface built with HTML, CSS, and Bootstrap to facilitate user interaction and data submission.

• **Backend:** A Flask server that handles request routing, user authentication, and coordinates the prediction logic.

• **Preprocessing Module:** Utilizes OpenCV to enhance handwriting images and Pandas for normalizing quiz scores.

• **Model Handler:** Manages pre-trained models stored in .pkl and .h5 formats to ensure fast inference times.

II. LITERATURE SURVEY

The application of intelligent systems for diagnosing learning disabilities has gained momentum due to the limitations of clinical settings. Previous research by Desai et al. highlighted the effectiveness of Random Forest in handling non-linear cognitive traits for dyscalculia. In the visual domain, Patil and Kulkarni demonstrated that CNNs could achieve high accuracy in distinguishing normal writing from the "reversal" patterns characteristic of dyslexia. Furthermore, studies by Gupta et al. emphasized the necessity of OpenCV preprocessing to improve edge clarity for dysgraphia analysis. This project builds upon these findings by offering a unified, multi-modal platform that addresses all three major disabilities simultaneously.

III. EXISTING SYSTEM

Current methods are largely static and rule-based, often relying on simple thresholding of quiz scores without adaptive learning. These systems generally lack integration between visual analysis and predictive analytics, require physical presence for evaluations, and typically focus on only a single disorder. These constraints reduce scalability and increase administrative effort in educational environments.

IV. PROPOSED SYSTEM

The proposed solution is an interactive, data-driven platform that offers several advantages:

• **Multi-disorder Analysis:** Supports the diagnosis of three distinct disabilities in one location.

• **Deep Learning Integration:** Uses CNNs to classify handwriting patterns (Normal, Reversal, Corrected) and potential levels of dysgraphia.

• **Intelligent Recommendations:** Supports DMIT report uploads to suggest tailored subject choices based on a child's cognitive strengths.

• **Real-Time Dashboard:** Provides immediate feedback with confidence scores for every prediction.

V. SYSTEM DESIGN

The system architecture emphasizes modularity and scalability through a multi-stage process.

A. Use Case Design

Defines two primary actors: the User (uploads data/views results) and the Admin (manages models/dashboard).

B. Sequence Flow

When a user submits data, the Flask backend validates the input, invokes the specific ML/DL model, and renders the result page without needing a persistent database for immediate predictions.

C. Deployment Node

The system is designed to run on a web server with models stored locally for fast access, accessible via any modern web browser.

VI. METHODOLOGY

The development follows a structured pipeline:

- **Environment Setup:** Development using Python 3.7+ with libraries including TensorFlow, Keras, and scikit-learn.
- **Model Training – Dyscalculia:** A Random Forest Classifier trained on cognitive test scores.
- **Model Training – Dyslexia/Dysgraphia:** CNN architectures featuring Conv2D, MaxPooling, and Dense layers for feature extraction from grayscale handwriting images.
- **Preprocessing:** Handwriting images are resized to 128 × 128 pixels and normalized to a 0–1 range to ensure model consistency.
- **Inference:** User inputs are processed on-the-fly, with results categorized based on patterns like “Mirror Writing” or “Low Potential”.

VII. RESULTS AND DISCUSSION

Experimental results indicate that the system is both efficient and accurate. The interface was validated through functional testing, ensuring that valid credentials lead to the admin dashboard and that image uploads trigger the correct CNN prediction flow.

- Dyscalculia Detection: Achieved 92% accuracy with an inference time of 50ms.
- Dysgraphia Detection: Achieved 88% accuracy with an inference time of 200ms.
- Dyslexia Detection: Achieved 85% accuracy with an inference time of 200ms.

VIII. CONCLUSION

The developed system provides a robust, AI-powered diagnostic framework that bridges the gap between educational needs and technological solutions. By successfully integrating Random Forest and CNN models, the project offers a scalable tool for early intervention. Future scope includes migrating to cloud hosting (AWS/Azure) for global scalability, developing mobile applications, and incorporating multilingual support to assist non-English speaking students.