International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 7, July, 2026, pp. 2189-2200
Journal homepage: http://iieta.org/journals/isi

Design and Evaluation of a Rule-Based Expert System for Early Screening of Learning Disabilities in Educational Settings

Laberiano Andrade-Arenas*, Gustavo Villar-Mayuntupa, Ada Alvarado-Paucar, Sebastián Ramos-Cosi, Enrique Huamani-Uriarte, Alicia Alva-Mantari

Facultad de Ciencias e Ingeniería, Universidad de Ciencias y Humanidades, Lima 15304, Peru

Corresponding Author Email: landrade@uch.edu.pe

https://doi.org/10.18280/isi.310706

Received: 19 March 2026
Revised: 9 June 2026
Accepted: 21 June 2026
Available online: 31 July 2026

Keywords:
dyscalculia, dysgraphia, dyslexia, early screening, knowledge representation, learning disabilities, rule-based expert system

ABSTRACT

Learning disabilities, including dyslexia, dyscalculia, and dysgraphia, can negatively affect students' academic development when they are not identified at an early stage. This study proposes a rule-based expert system to assist the early screening of learning disabilities in educational environments. The system was developed through expert knowledge representation based on the Buchanan methodology and implemented using the Experta inference engine, Flask framework, and PostgreSQL database. The proposed system integrates pedagogical rules to evaluate learning-related indicators and generate classification outcomes for potential learning difficulties. Its performance was assessed using 90 test cases, consisting of 60 positive and 30 negative cases, through a confusion matrix and corresponding evaluation metrics. The system achieved an accuracy of 91%, precision of 100%, sensitivity of 87%, specificity of 100%, and an F1-score of 93%. The results demonstrate that the proposed approach can effectively identify potential learning difficulties while avoiding false-positive classifications. By combining artificial intelligence techniques with educational expertise, the developed system provides a practical decision-support approach for teachers and educational practitioners involved in early learning disability screening.

1. INTRODUCTION

In schools, teachers report that a significant number of students experience learning difficulties in various areas, such as writing, composition, and basic mathematics, particularly in the use of basic operations like addition, subtraction, multiplication, and division. These difficulties may be due to specific learning disorders, such as dyslexia, dyscalculia, and dysgraphia, which affect students' academic performance and social integration [1].

One of the main problems in regular basic education is that many schools lack inclusive education areas and specialized professionals, such as psychologists, educators, and teachers trained in identifying and addressing learning difficulties. This leads to problems going unnoticed or being diagnosed late, limiting opportunities for early intervention [2].

Without proper detection and support, students may experience emotional and social problems that affect their self-esteem, making them more likely to be victims of bullying or exclusion in the classroom. These difficulties can also contribute to dropping out of school, as students who do not receive support often develop a negative attitude toward learning. From a pedagogical perspective, teachers often lack access to technological tools that facilitate the early identification of learning difficulties. This hinders effective diagnosis and personalized educational intervention [3]. In this context, expert systems can play a key role in early detection and guiding appropriate teaching strategies, thereby optimizing learning processes and improving educational inclusion [4].

This research is important because the proposed expert system model can be implemented at various Bertolt Brecht School locations and other Peruvian schools. It will benefit teachers, educational psychologists, and students by facilitating the early detection of learning difficulties. The model's contribution to the education sector is significant because it will improve teaching processes, strengthen educational inclusion, and ensure that each student receives appropriate attention. Currently, many institutions lack the technological tools necessary for identifying and addressing learning disorders, which limits the ability to intervene in a timely manner and negatively impacts students' academic performance. This expert system will provide support based on artificial intelligence and inference rules, enabling more accurate diagnoses and effective responses from teachers and specialists.

This project's impact extends beyond the classroom. It will benefit parents by providing clear information about their children's learning progress and guidance on supporting them at home. Additionally, involving authorities in the project's implementation will strengthen institutional policies promoting inclusive education. Technologically, this project promotes the use of artificial intelligence in education, fostering innovation in teaching methods and psychopedagogical assessment. In the long term, continuous improvement of the model will be possible through data collection, which will optimize its accuracy and effectiveness in different educational settings [5]. This study is a significant advance in detecting and supporting learning difficulties, contributing to educational equity and developing more personalized, effective teaching methods in Peru.

The main objective is to design and implement an expert system based on inference rules that can identify learning difficulties in students, such as dyslexia, dyscalculia, and dysgraphia. The system will be designed to support students, parents, and administrators at Bertolt Brecht School by providing accurate information that facilitates early detection and timely educational decision-making.

2. LITERATURE REVIEW

This section will review the literature of various studies to address the design of an expert system to support teachers and educational psychologists in identifying learning difficulties in the classroom. It will examine several studies that highlight the importance of technology in education, particularly in the early detection of learning difficulties, and how these tools can facilitate teacher intervention, ensuring that each student receives the necessary support to overcome their educational challenges.

Nugraheni et al. [6] developed a web-based expert system aimed at supporting the diagnosis of learning disorders in children through the application of the Dempster-Shafer theory. The proposed method models uncertainty by assigning degrees of belief to symptoms and combining multiple pieces of evidence to infer the most probable diagnosis. The system enables users to input observed symptoms and generate outputs that include the identified disorder, its definition, and suggested treatment options. The evaluation results demonstrate an accuracy of 92%, highlighting the effectiveness of the approach in assisting the diagnostic process for learning disorders.

On one hand, Ana [7] explained that the pandemic has impacted all sectors of society, including the economy, social relations, and education. It has particularly altered interpersonal connectivity, as social and physical distancing became the primary means of preventing contagion. These measures necessitated that all primary and secondary school activities be conducted from home, forcing a shift from conventional classroom instruction to online learning. However, theoretical learning without practical application can hinder the development of complex cognitive skills; therefore, it is essential for educators to ensure that these practices are meaningful and contextually appropriate. Consequently, there is an urgent need for online expert systems that serve as pedagogical tools. It is, thus, necessary to develop programs that contribute to professional development within virtual and distance education frameworks. The methodology for this study involved a bibliographic review of articles and journals focusing on the development of expert systems designed to promote effective online learning during the pandemic.

Similarly, given the growing need for early detection of learning disorders such as dysgraphia, dyslexia, and dyscalculia in children, Andrade-Arenas and Yactayo-Arias [2] set out to evaluate an expert system developed in Python designed to facilitate early diagnosis. Background research underscores the importance of having effective tools to assist parents, teachers, and healthcare professionals in this timely identification process. In a sample of 21 simulated cases, the system demonstrated outstanding performance, achieving a 95% accuracy rate, 100% precision, 93% sensitivity, and 100% specificity. Furthermore, the acceptability assessment, conducted with 15 parents selected through convenience sampling, reflected a high level of satisfaction: an overall average of 4.78 out of 5 was obtained, with a standard deviation of 0.45, suggesting remarkable consistency in responses. The results support the effectiveness of the expert system as a useful tool for the early diagnosis of learning difficulties. However, although the findings are promising, the importance of further research is recognized to overcome current limitations and strengthen the external validity of the system to ensure its effective application in real clinical contexts.

Anggrawan et al. [8] proposed an expert system for the early diagnosis of learning disorders in children, considering conditions such as dyslexia, dysgraphia, dyscalculia, and dyspraxia. The study emphasizes the importance of timely identification, as lack of awareness may prevent children from receiving appropriate support and can negatively affect their cognitive and emotional development. Methodologically, the system integrates both the certainty factors (CFs) and Dempster-Shafer theory approaches to handle uncertainty in symptom-based diagnosis. The results indicate that the CF method achieves higher accuracy (90%) compared to the Dempster-Shafer method (87%), demonstrating its relative effectiveness in this context. The study contributes by comparatively implementing both methods within a single diagnostic system, enabling the identification of different types of learning disorders in children.

On the other hand, Alzahrani and Algahtani [9] proposed an artificial intelligence-driven approach for the identification and diagnosis of specific learning disabilities, particularly dyslexia and dysgraphia, which significantly affect children's academic performance. The methodology is based on a weighted ensemble learning model built upon the XGBoost framework, designed to analyze data collected from personalized assessments, including performance scores and task completion times. This approach enables the system to handle imbalanced datasets and improve diagnostic reliability. The model was implemented within an application that evaluates reading and writing difficulties while also providing tailored instructional recommendations for parents and educators. Experimental results show that the proposed model outperforms other machine learning and deep learning approaches, achieving accuracy rates of 98.7% for dyslexia and 99.08% for dysgraphia, highlighting its effectiveness as a diagnostic and support tool.

Kurniawan and Tiaharyadini [10] applied machine learning techniques for early dyslexia detection in primary school children, combining a literature review with the development and evaluation of multiple models. The study employs algorithms such as decision tree, K-nearest neighbors, logistic regression, Naive Bayes, and random forest to compare their diagnostic performances. The results show that decision tree achieves high precision (92.31% for dyslexia-prone and 90.62% for diagnosed dyslexia) with strong recall values, while K-nearest neighbors attains the highest overall accuracy of 94% and perfect precision (100%) for non-dyslexia cases. Logistic regression demonstrates notable predictive capability with precision up to 95.38% and recall above 90%, whereas

Naive Bayes reaches perfect precision (100%) in certain categories but has slightly lower performance in diagnosed dyslexia. Random forest provides balanced results, with precision ranging from 91.18% to 94.23% and recall between 92.31% and 93.94%, achieving an overall accuracy of 93%. These findings confirm the effectiveness of machine learning models in enabling accurate and early identification of dyslexia.

Furthermore, Mulakaluri and Gowdra Shivappa [11] aimed to develop an expert system for diagnosing learning disorders in children using Dempster-Shafer theory to make diagnoses more accurate and timelier. They employed a technological development approach encompassing several phases: identifying the problem, constructing the knowledge base with the help of experts, developing the inference engine based on this theory, implementing the web system, and testing the system's accuracy. The system focused on three disorders-dyslexia, dysgraphia, and dyscalculia-calculating density values for 30 identified symptoms. As a result, an interactive web application was created that allows users to enter symptoms and receive diagnoses and treatment recommendations. Validation tests with 50 cases showed 92% accuracy, reflecting the system's reliability. The authors conclude that Dempster-Shafer theory is effective for diagnosing learning disorders. However, they suggest incorporating fuzzy logic into future developments to improve the system's flexibility and validity.

Raatikainen et al. [12] proposed an artificial intelligence–based approach for diagnosing dyslexia in children through a testing tool that evaluates cognitive and academic performance. The methodology involves administering quizzes and task-based assessments tailored to potential learning impairments, where the resulting data-such as performance scores and task completion time-are processed using an ensemble feature-aware model (EFAM) built on XGBoost. This EFAM-XGB model is designed to handle imbalanced datasets while improving diagnostic accuracy. The system was implemented as an integrated and user-friendly application that not only identifies reading disorders but also provides instructional recommendations for parents and teachers. The results demonstrate a high accuracy of 98.7% in dyslexia detection, outperforming existing machine learning approaches and highlighting its effectiveness as a diagnostic tool.

Similarly, Drotár and Dobeš [13] explored an automated approach for dyslexia identification based on the analysis of eye movement patterns during reading, aiming to provide an alternative to traditional expert-based assessments. The methodology combines random forest for feature selection with a support vector machine for classification, using eye-tracking data to detect individuals with low reading fluency, including those with borderline performance. This hybrid model focuses on extracting the most relevant features from reading behavior to improve diagnostic reliability in more realistic scenarios. The results show that the proposed approach achieves an accuracy of 89.7% and a recall of 84.8%, demonstrating its effectiveness in identifying dysfluent readers and supporting the feasibility of automated dyslexia detection in natural reading environments.

Finally, Bhavsar et al. [14] noted that dysgraphia, a learning disorder that affects written expression, can negatively impact children's academic performance and self-esteem. Its diagnosis is limited by subjective and inaccessible methods. The objective of this study was to develop an automated machine learning model to more accurately detect dysgraphia. Handwriting data was collected from 120 schoolchildren using a digital tablet, and over a thousand kinematic, spatial, and dynamic features were extracted. The algorithms AdaBoost, random forest, and support vector machine were applied; AdaBoost achieved an accuracy rate of 79.5%. This demonstrates that machine learning models can objectively identify dysgraphia, even in heterogeneous samples. The study concludes that this technology is an efficient tool for early dysgraphia screening in schools, optimizing pedagogical and therapeutic interventions.

In conclusion, a literature review of various studies was conducted to inform the design of an expert system. However, the review revealed that the integration of technological tools into the educational process remains inadequate, particularly in the early detection of learning difficulties. Thus, implementing an expert system to facilitate the work of teachers and educational psychologists was proposed. This system would provide more precise and timely support in identifying problems and enable more effective interventions.

3. METHOD

3.1 Methodological design

A Scrum and Buchanan-based approach was used to develop the web application, enabling iterative development, expert validation, and structured system design.

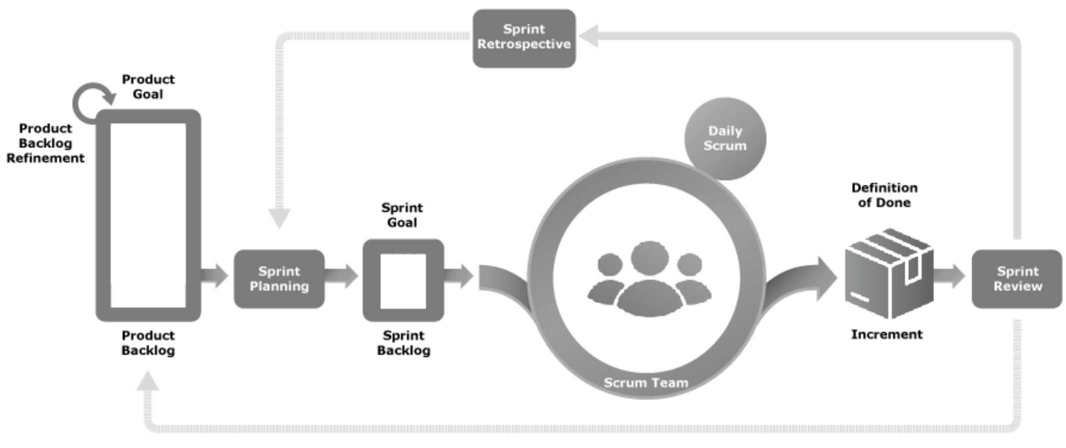


Figure 1. Phases of the Scrum methodology

3.1.1 Agile Scrum methodology

Scrum is a widely used agile project management framework designed to help teams work together and adapt quickly, especially in software development. It emphasizes iterative progress through short cycles called "sprints," regular team meetings, and defined roles, such as product owner, scrum master, and development team (see Figure 1). These roles work together to efficiently deliver high-quality products. Scrum's core principles of transparency, inspection, and adaptation foster collaboration, accountability, and continuous improvement within teams [14, 15].

In this study, Scrum was applied through iterative sprints covering the entire system development process. Each sprint focused on specific tasks such as designing the system architecture, implementing and refining the inference engine, and constructing and adjusting the rule base. At the end of each sprint, testing and evaluation were conducted to identify inconsistencies and improve system behavior. This iterative sprint-based approach enabled continuous refinement of the expert system, particularly in improving rule accuracy, reducing logical conflicts, and enhancing overall diagnostic performance.

3.1.2 Buchanan methodology

Buchanan's methodology provides a clear, orderly approach to creating expert systems. It starts with identifying the problem and continues with defining and organizing essential components. Then, it translates that knowledge into well-structured rules and inferences. Finally, it culminates in the operational implementation and rigorous evaluation of the system. Following these sequential stages ensures that each step remains consistent with the stated objectives, resulting in a robust, validated solution (see Figure 2).

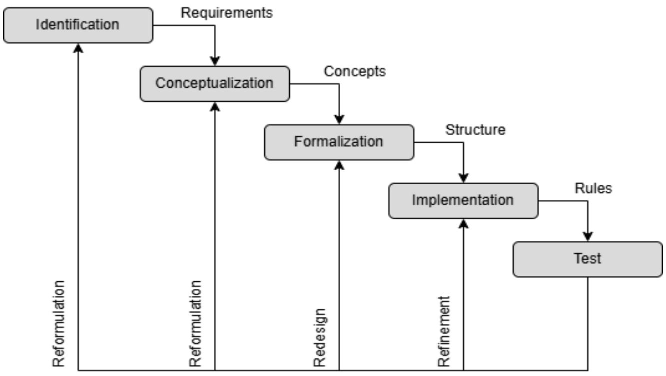


Figure 2. Phases of the Buchanan methodology

3.2 Knowledge acquisition

Semi-structured interviews were conducted with specialists in the field of education to gather information related to learning difficulties. These interviews provided detailed, contextualized knowledge about the main difficulties identified in students and the criteria professionals use to diagnose conditions such as dyslexia, dyscalculia, and dysgraphia. The information obtained through the interviews was essential for formalizing knowledge within the expert system because it enabled the definition of key symptoms, the establishment of inference rules, and the validation of real-world scenarios within the domain.

Table 1. Symptoms of learning problems

Learning Problems	Id	Symptoms	Score (1–5)
Dyslexia	dyslexia_1	Slow, syllabic, or word-by-word reading.	5
	dyslexia_2	Substitution of words with others that are visually or phonetically similar.	4
	dyslexia_3	Additions or omissions of letters and syllables.	4
	dyslexia_4	Difficulty recognizing familiar words at first sight.	4
	dyslexia_5	Confusion of visually similar letters (b/d, p/q, m/n, u/n)	4
	dyslexia_6	Letter, syllable, or word reversals.	4
	dyslexia_7	Line breaks, omissions, or repetitions when reading aloud.	4
	dyslexia_8	Persistent difficulty in the automatic recognition of frequent words (reading fluency)	5
	dyslexia_9	Problems associating grapheme-phoneme (letter-sound correspondence)	5
	dyslexia_10	Difficulty segmenting sounds (phonological awareness), not just syllables.	4
	dyslexia_11	Inconsistent reading comprehension.	3
	dyslexia_12	Deficit in verbal working memory.	3
	dyslexia_13	Confusion regarding directionality (left/right)	1
	dyslexia_14	Slow reading speed compared to peers.	4
	dyslexia_15	Avoids activities that require reading.	1
	dyslexia_16	Common spelling mistakes.	3
	dyslexia_17	Fatigue or frustration when reading.	2
Dysgraphia	dysgraphia_1	Illegible, disorganized, or inconsistent handwriting.	5
	dysgraphia_2	Poorly formed letters, irregular sizes, and poor alignment on the page.	4
	dysgraphia_3	Mixture of uppercase and lowercase letters in the same word.	3
	dysgraphia_4	Excessively slow or rushed writing with errors.	4
	dysgraphia_5	Very poor spelling.	5
	dysgraphia_6	Frequent grammatical and punctuation errors.	3
	dysgraphia_7	Difficulty planning and sequencing ideas in writing.	5
	dysgraphia_8	Problems maintaining margins or following lines.	2
	dysgraphia_9	Difficulty copying correctly from the blackboard or a book.	4
	dysgraphia_10	Pain in the hand.	2
	dysgraphia_11	Confusion when writing letters with similar shapes (e.g., m/n, p/q, d/b)	3
	dysgraphia_12	Omit, add, or repeat letters in words.	4
	dysgraphia_13	Problems remembering how to spell previously seen words.	4
	dysgraphia_14	Avoid tasks that require writing.	2

Learning Problems	Id	Symptoms	Score (1–5)
Dyscalculia	dyscalculia_1	Problems recognizing numbers and mathematical symbols.	5
	dyscalculia_2	Difficulty understanding the relationship between numbers and symbols.	5
	dyscalculia_3	Difficulty learning multiplication tables.	4
	dyscalculia_4	Confusion between mathematical signs (+, −, ×, ÷)	4
	dyscalculia_5	Common errors in basic operations.	5
	dyscalculia_6	Difficulty remembering numerical sequences.	4
	dyscalculia_7	Difficulty reading the time on analog clocks.	3
	dyscalculia_8	Problems handling money (calculating change)	3
	dyscalculia_9	Difficulty in measuring or estimating quantities.	4
	dyscalculia_10	Difficulty understanding concepts such as greater/lesser, before/after	4
	dyscalculia_11	Shows insecurity or persistent difficulty in solving mathematical problems.	4
	dyscalculia_12	Constant need to use fingers to count.	3
	dyscalculia_13	Excessive slowness in mental calculations.	4
	dyscalculia_14	Difficulty recognizing patterns or logical sequences.	4
	dyscalculia_15	Problems with spatial orientation (maps, directions, coordinates).	2
	dyscalculia_16	Avoid tasks or games that require math.	1
	dyscalculia_17	Significant difficulty in understanding which operation to apply in word problems.	5
	dyscalculia_18	Difficulty in understanding place value.	5
	dyscalculia_19	Errors in aligning figures in operations.	2

3.3 Knowledge base representation

3.3.1 Facts (symptoms)

Table 1 presents the list of symptoms that the expert system considers when diagnosing the three learning difficulties addressed. A total of 50 symptoms are included and are distributed as follows: 17 are associated with dyslexia, 14 with dysgraphia, and 19 with dyscalculia. These symptoms were selected from specialized educational sources and constitute the knowledge base used by the expert engine during the diagnostic inference process. Each symptom is weighted with a value between 1 and 5 according to its potential impact on the student's learning. A value of 1 indicates a very slight effect, while a value of 5 represents a very severe impact.

3.3.2 Rules

The expert system was implemented using the Experta library, which provides a forward-chaining inference engine based on IF-THEN production rules. In this approach, the conditions represent the symptoms observed in the student, while the actions declare new facts or derived diagnoses. To ensure the relevance and reliability of the knowledge base, the inference rules were constructed based on established literature on learning disorders and organized according to domain-specific symptom criteria associated with dyslexia, dyscalculia, and dysgraphia.

The process began with the symptoms entered by the user, which were evaluated against a set of 2000 rules defined in the knowledge base. When the conditions of one or more rules were met, the system automatically inferred possible diagnoses. Furthermore, an iterative refinement process was carried out to improve consistency, reduce redundancy, and minimize conflicts between rules. Figure 3 shows an example of a rule corresponding to dysgraphia; however, the same inference logic applies to other disorders such as dyslexia and dyscalculia.

```
@Rule(Symptom(id='dyslexia_9'), Symptom(id='dyslexia_10'),
    Symptom(id='dyslexia_11'), Symptom(id='dyslexia_16'),
    Symptom(id='dyslexia_1'), Symptom(id='dyslexia_4'),
    Symptom(id='dyslexia_8'))
    def dyslexia_rule396(self):
        matched = [f["id"] for f in self.facts.values()
            if isinstance(f, Symptom) and f["id"].startswith("dyslexia")]
        self.declare_result("Dislexia", matched)
```

Figure 3. Rule representation using the Experta library

Each rule is defined by a set of symptom-type indicators and is activated only when all required conditions are present. Once triggered, the system provides the diagnosis along with matching symptoms, a score, and a percentage that determines severity, allowing traceability of the decision process. Finally, validation was conducted using representative test cases, which allowed comparison between system outputs and expected diagnoses, supporting the reliability of the inference process in practical educational contexts.

3.4 Conversion of severity scores to certainty factors

The conversion is performed using a predefined scale (see Table 2). The inference engine uses these factors to calculate the level of impairment-mild, moderate, severe, or very severe-by combining and comparing the obtained values (see Table 3). Thus, the expert system can more accurately determine the probability that a student will experience a particular difficulty based on the quantity and severity of the detected symptoms.

Table 2. Conversion of severity scores to determine the severity level

Scale	Level	Certainty Factor
1	Very mild	0.10
2	Mild	0.30
3	Moderate	0.60
4	Severe	0.85
5	Very severe	1.00

Table 3. Diagnostic interpretation

Percentage Range (%)	Severity Level
0–24	Mild
25–49	Moderate
50–79	Severe
80–100	Very severe

3.5 Evaluation of the performance of the expert system

To evaluate the expert system's performance, a confusion matrix was developed. From this matrix, precision, accuracy, and recall metrics were calculated, allowing the system's ability to correctly identify each class, minimize false positives (FP), and reflect its overall reliability to be quantified.

3.5.1 Confusion matrix

The confusion matrix is a table that summarizes the performance of a classification model by comparing its predictions with the actual values of the test set. The matrix allows us to count the number of cases correctly classified as true positives (TP) or true negatives (TN), as well as the number of false positives (FP and false negatives (FN) predicted. Essential metrics can then be calculated from this information to determine the model's overall performance [16].

3.5.2 Precision

Precision is a fundamental metric for evaluating the performance of a classification model. First, accuracy indicates the proportion of correct positive predictions. Second, it allows us to estimate the reliability of the system's positive results. Finally, it helps reduce FP, which can be critical in certain contexts. Accuracy is calculated using Eq. (1).

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

3.5.3 Accuracy

Accuracy is an overall measure of a classification model's performance. It is obtained by dividing the sum of correct predictions (TP and TN) by the total number of samples evaluated. This metric, represented by Eq. (2), reflects the proportion of correctly classified cases, regardless of class. It is often used to summarize the overall results of a confusion matrix.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

3.5.4 Recall

Recall measures the system's ability to correctly identify cases that belong to a specific positive class. It represents the proportion of positive items that were correctly identified. A high sensitivity value indicates that the model effectively recognizes most positive cases. It is calculated using Eq. (3).

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

3.5.5 Specificity

It evaluates the system's ability to correctly identify cases that belong to the negative class. It represents the proportion of negative cases that were correctly classified. A high specificity value means the model effectively distinguishes negative cases and avoids misclassifying them as positive [17]. It is calculated using Eq. (4).

$$Specificity = \frac{TN}{FP + TN} \tag{4}$$

3.5.6 F1-score

It evaluates the balance between a system's ability to correctly identify positive cases and the accuracy with which it does so. It represents the harmonic mean of precision and recall, combining these two metrics into a single value. A high F1-score indicates that the model performs well in detecting positive cases and avoiding misclassifications. It is calculated using Eq. (5).

$$F1 - Score = \frac{2\, precision \cdot recall}{precision + recall} \tag{5}$$

3.6 Architecture of the proposed expert system

The proposed expert system is designed to assist professionals in identifying learning difficulties in students. Its architecture comprises several interconnected modules that enable user interaction, data management, information storage, and rule-based reasoning.

3.6.1 User interface

The user interface (UI) is the main way users interact with the system. Professionals must authenticate with a secure login to access the system. Once authenticated, they can select a student from the PostgreSQL database and choose which symptoms to analyze. The UI then retrieves the list of symptoms from a JSON file, which serves as the source of facts for the inference engine. The interface also displays diagnostic results as soon as they are generated, providing users with immediate feedback.

3.6.2 Symptoms JSON file

The JSON file contains a list of possible symptoms for the professional to select. This file serves two main functions: first, it provides the list of symptoms to the UI for selection; and second, it serves as a source of information for the inference engine, enabling the expert system to analyze the data entered by the professional.

3.6.3 Inference engine

The inference engine is at the core of the expert system. It receives selected symptoms as facts and applies a set of predefined rules to generate a diagnosis. If any rule is met, the engine returns a result indicating the identified learning problem. This module is responsible for logical reasoning and ensuring that diagnoses are consistent with the system's knowledge base.

When new symptoms are entered, they are asserted as facts into the working memory of the inference engine. The engine continuously evaluates all rules whose conditions match the current set of facts. If all conditions of a rule are satisfied, the rule is triggered, and its corresponding action is executed, generating new facts or a final diagnosis.

If any rule is met, the engine returns a result indicating the identified learning problem. This module is responsible for logical reasoning and ensuring that diagnoses are consistent with the system's knowledge base. Through this chaining process, multiple rules may be activated sequentially, allowing the system to refine or confirm diagnostic outcomes based on the available symptom set.

3.6.4 PostgreSQL database

The PostgreSQL database stores critical system information.

•Users: Data from professionals and students.

•Diagnostic history: Results obtained, which allow for monitoring and longitudinal analysis of each student's performance.

To ensure data privacy and security, appropriate protection measures were implemented. These include controlled access to the database, user authentication mechanisms, and restricted data visibility based on user roles. Additionally, sensitive information is handled following data protection principles, minimizing unnecessary data storage and ensuring confidentiality. These measures aim to align the system with educational data privacy standards.

3.6.5 Flow of information

The flow of information follows these steps:

1. The professional logs into the system and selects a student.
2. The UI retrieves the list of symptoms from the JSON file and allows the professional to select the applicable ones.
3. The inference engine analyzes the symptoms as facts and compares them with the system rules.
4. If any rule is met, a diagnostic result is generated and displayed in the UI.
5. The diagnostic result is stored in the PostgreSQL database as the student's record.

Figure 4 shows the system's overall information flow, detailing the interactions among users, the interface, the inference engine, the JSON file, and the database.

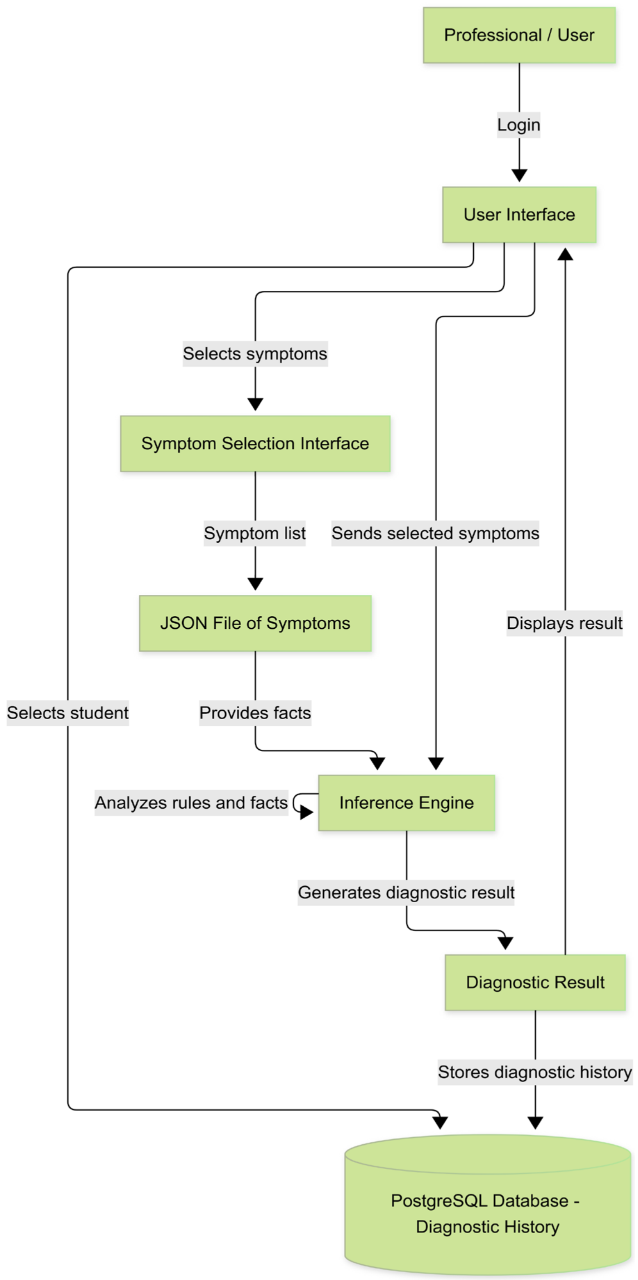


Figure 4. Information flow of the proposed system

4. RESULTS

4.1 User interface of the application

Figure 5 shows the login interface of the web-based expert system. The screen has two main fields for user access: email address and password. There is also a profile or role selector that allows users to identify themselves as an administrator, educational psychologist, teacher, or parent.

After entering the information and validating access, the system redirects the user to the main dashboard corresponding to their role. Each dashboard offers specific functions and restrictions according to profile type, ensuring the available options align with assigned responsibilities and permissions within the system.

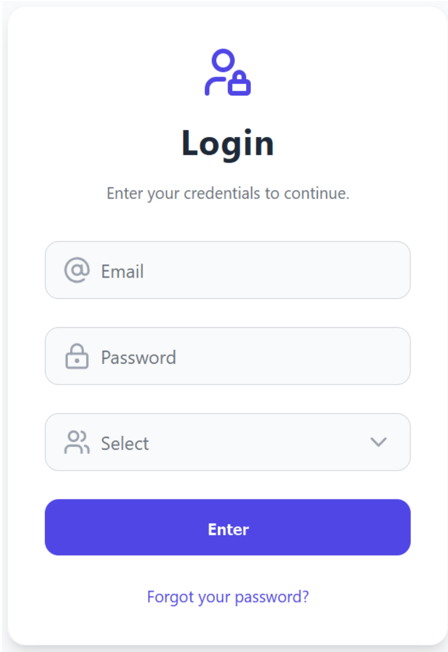


Figure 5. Login

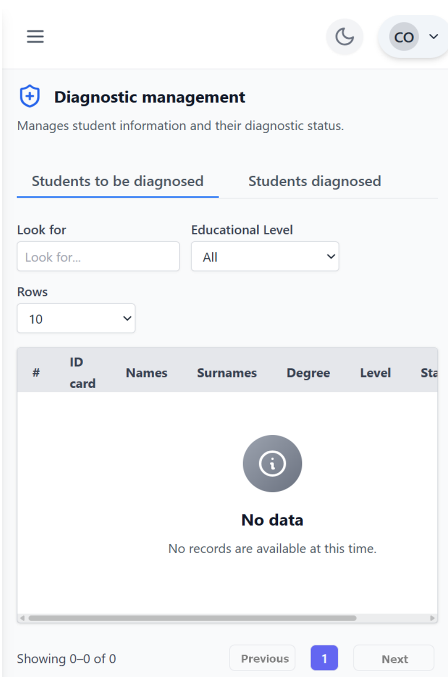
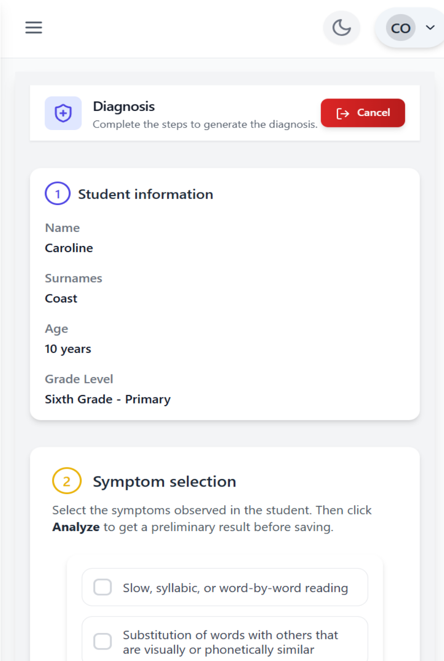


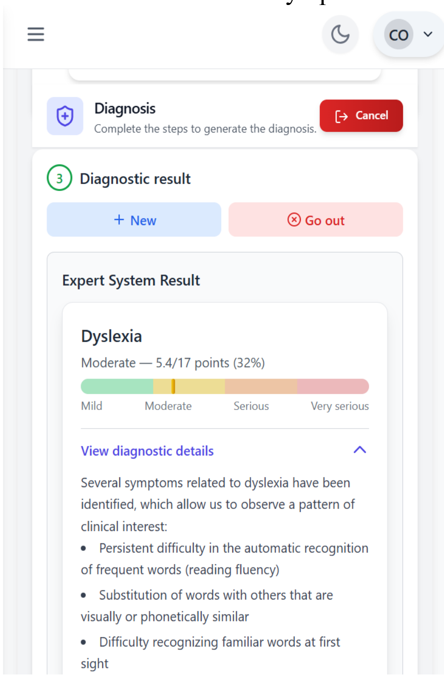
Figure 6. Main view of the diagnostic module

Figure 6 shows the student diagnostic management interface within the web-based expert system. This screen displays a list of students who have been diagnosed and those who are awaiting diagnosis. The list is organized so that the educational psychologist or teacher can quickly identify the status of each case.

The system only displays students who have been assigned by the administrator to the corresponding educational psychologist or teacher. This ensures that each professional only has access to their assigned students, maintaining a controlled workflow in accordance with their defined permissions. Furthermore, the interface includes status indicators for diagnosed and pending diagnoses. Thus, the platform facilitates orderly management of the diagnostic process, ensuring that each student is seen by the assigned specialist.



(a) Student information and symptom selection



(b) Diagnostic result

Figure 7. Diagnostic process

Figure 7(a) shows the diagnostic process interface where essential student information and the list of available assessment symptoms are presented. Each symptom is displayed as a checkbox, enabling the educational psychologist or teacher to select the symptoms they have identified. The system requires a minimum of three selected symptoms to ensure the inference is based on sufficient evidence.

After confirming the selection, the system validates compliance with this requirement and compares the information with the rules of the inference engine. If there is a match with any rule of the expert system, a preliminary diagnostic result is generated and displayed to the user in a clear, structured manner. This provides the educational psychologist or teacher with initial guidance on the type of difficulty the student may be experiencing based on the system's logic and the selected evidence.

Figure 7(b) shows the diagnostic results interface. It displays the identified problem, estimated severity level, score obtained, and corresponding percentage used to determine the severity level. The screen also includes a detailed diagnostic section that explains how the expert system reached its conclusion by considering the selected symptoms and the rules activated during the inference process.

Before recording the results in the student's diagnostic record, the professional reviews and validates the generated information. This verification is essential because the expert system provides preliminary or supporting results that must be compared with the educational psychologist's clinical judgment before confirmation and final storage.

4.2 System performance evaluation

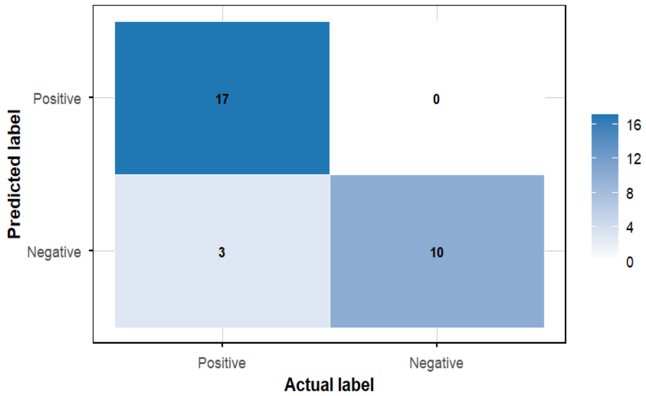
Previously, the expert system's performance was evaluated using 90 test cases, which were distributed as 20 positive and 10 negative cases for each of the three types of learning difficulties: dyslexia, dyscalculia, and dysgraphia. This distribution enabled analysis of the system's behavior with both cases that met the diagnostic criteria and cases that did not present the corresponding problem.

Performance was analyzed using confusion matrices developed for each type of difficulty. The following derived metrics were then calculated: accuracy, precision, sensitivity, specificity, and F1-score. Subsequently, the individual results were integrated to obtain an overall evaluation of the system's performance, providing an overview of its capacity to support the diagnostic process. The results for each evaluated problem are presented below.

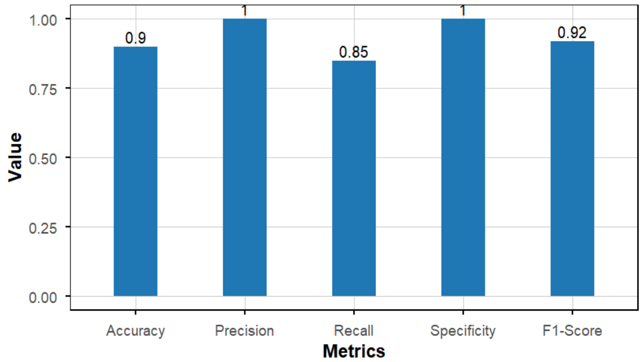
4.2.1 System performance in identifying dyslexia

The results for identifying dyslexia are presented in Figures 8(a) and 8(b). The confusion matrix shows that the system had 17 TP and 10 TN, with zero FP and three FN. These values demonstrate that the system did not generate erroneous diagnoses in the absence of the condition. However, in some cases, it failed to detect the condition when it was present.

Based on these data, the derived metrics reflect robust performance. Precision was 1.0 (100%), meaning all cases classified as positive corresponded to dyslexia. Accuracy reached 0.90, demonstrating adequate overall performance. Sensitivity (recall) was 0.85, showing that the system correctly identified most positive cases, though some were missed. Specificity, with a value of 1.0 (100%), confirms the system's perfect ability to correctly recognize cases without the condition. Finally, the F1-score of 0.92 reflects a solid balance between accuracy and sensitivity, supporting the system's effectiveness in diagnosing dyslexia.

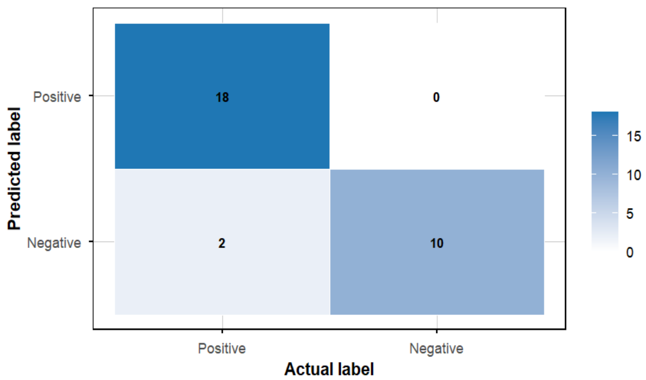


(a) Confusion matrix of the expert system for dyslexia

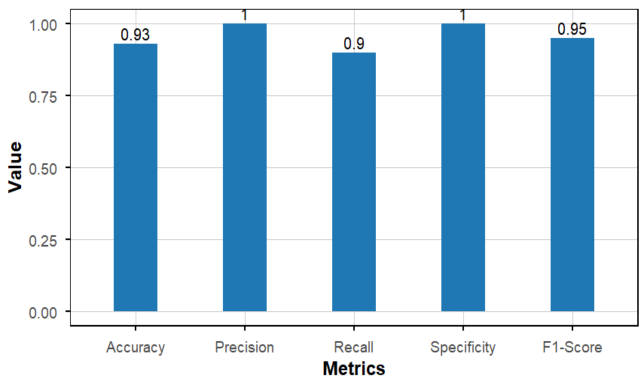


(b) Expert system assessment metrics for dyslexia

Figure 8. Performance of the expert system in identifying dyslexia



(a) Confusion matrix of the expert system for dysgraphia



(b) Expert system evaluation metrics for dysgraphia

Figure 9. Performance of the expert system in the identification of dysgraphia

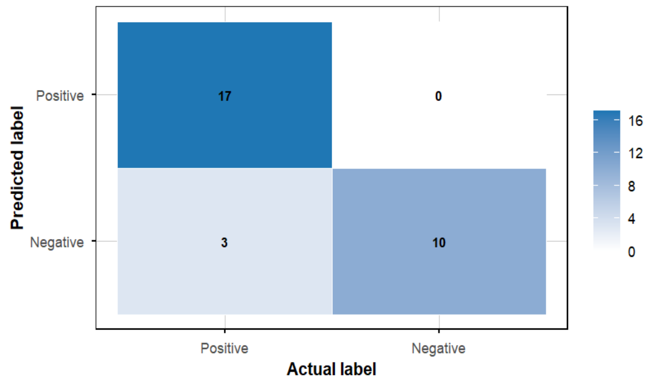
4.2.2 System performance in identifying dysgraphia

Figures 9(a) and 9(b) show the results for dysgraphia detection and demonstrate the solid performance of the expert system. The confusion matrix shows 18 TP and 10 TN, with zero FP and only two FN. These results indicate that the system did not misclassify any students without the condition and that only a few cases of dysgraphia were not identified.

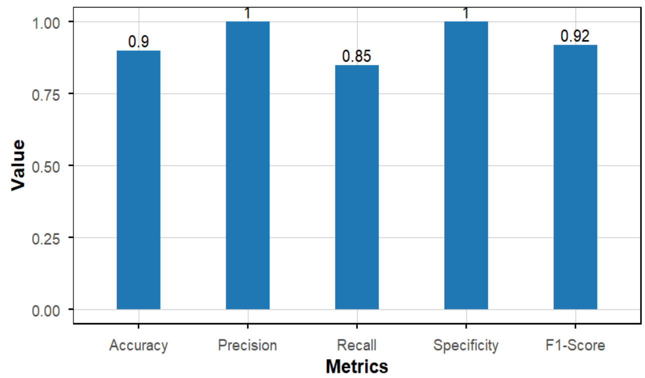
The derived metrics confirm this favorable performance. Precision reached a value of 1.0, meaning all positive predictions were correct. Accuracy, at 0.93, shows high overall performance. Sensitivity (recall) was 0.90, indicating that the system accurately identified most students with dysgraphia. Specificity was also 1.0, revealing a perfect ability to discriminate between students without dysgraphia. Finally, an F1-score of 0.95 demonstrates an outstanding balance between accuracy and sensitivity, supporting the system's effectiveness in diagnosing dysgraphia.

4.2.3 System performance in identifying dyscalculia

The results for dyscalculia detection, presented in Figures 10(a) and 10(b), show favorable performance across the board. The confusion matrix reports 17 TP and 10 TN, with zero FP and three FN. These values demonstrate that the system did not misclassify cases without the condition, though it failed to identify some actual cases of dyscalculia.



(a) Confusion matrix of the expert system for dyscalculia



(b) Metrics for evaluating the expert system for dyscalculia

Figure 10. Performance of the expert system in identifying dyscalculia

The derived metrics reinforce this trend. Precision reached a value of 1.0, indicating that all positive predictions were correct. Accuracy, at 0.90, indicates adequate overall performance. Sensitivity (recall), at 0.85, shows that the system detects most positive cases; however, there is room for improvement in completely identifying the difficulty. Specificity, at 1.0, confirms the system's ability to perfectly classify negative cases. Finally, the F1-score of 0.92 shows a solid balance between accuracy and sensitivity, validating the system's usefulness as a support tool for dyscalculia detection.

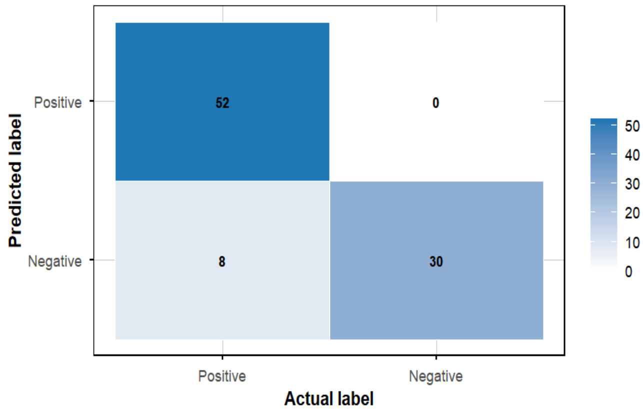
4.2.4 Overall system performance

Figures 11(a) and 11(b) show the overall results, which integrate the cases corresponding to dyslexia, dysgraphia, and dyscalculia. This allows for an evaluation of the expert system's overall performance. The confusion matrix shows 52 TP and 30 TN, with zero FP and eight FN. While this behavior confirms the system's ability to avoid misdiagnosing students without difficulties, it also reveals some real cases that went undetected.

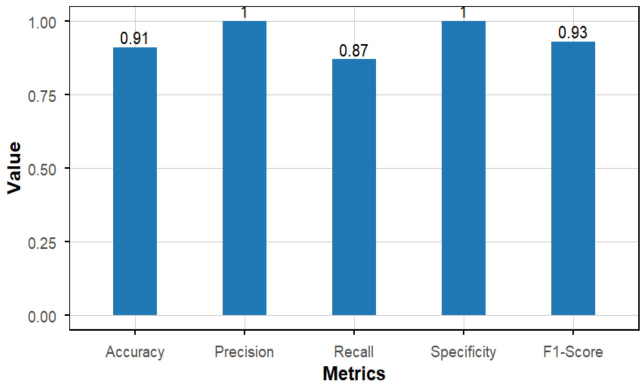
The derived metrics reflect robust overall performance. Precision yielded a value of 1.0, indicating that all positive classifications issued by the system were correct. Accuracy, at 0.91, demonstrates high performance considering the three types of learning difficulties. Sensitivity (recall) reached 0.87, showing that the system effectively identifies most positive cases, though some omissions persist. Specificity reached a value of 1.0, demonstrating perfect recognition of negative cases. Finally, the F1-score of 0.93 reveals a solid balance between accuracy and sensitivity. This confirms that the expert system reliably and consistently supports the comprehensive identification of learning difficulties.

4.2.5 Comparative performance analysis

Table 4 presents a comparative summary of the performance metrics (accuracy, precision, recall, specificity, and F1-score) of the proposed system and those reported in related studies for learning disability diagnosis, along with the corresponding methods and evaluated conditions.



(a) General confusion matrix of the expert system



(b) Metrics for the overall evaluation of the expert system

Figure 11. Performance of the expert system in identifying learning disorders

Table 4. Comparative performance of the proposed expert system and related approaches in learning disability diagnosis

Study	Learning Disability	Method	Accuracy	Precision	Recall	Specificity	F1-Score
[2]	Dyslexia, Dysgraphia, Dyscalculia	Rule-based expert system	95%	100%	93%	100%	-
[8]	Dyslexia, Dysgraphia, Dyscalculia, Dyspraxia	Certainty factors (CFs)	90%	-	-	-	-
[9]	Dyslexia, Dysgraphia	WEL-XGBoost	98.7%, 99.08%	-	-	-	-
[11]	Dyslexia, Dysgraphia, Dyscalculia	Dempster-Shafer	92%	-	-	-	-
[12]	Dyslexia	EFAM-XGB	98.7%	98.66%	98.65%	98.67%	98.67%
[13]	Dyslexia	Hybrid model	84.8%	-	84.8%	-	-
This study	Dyslexia, Dysgraphia, Dyscalculia	Expert System (Experta)	93%	100%	90%	100%	95%

5. DISCUSSION

The developed expert system identifies dyslexia, dysgraphia, and dyscalculia through an inference engine constructed from 50 symptoms and implemented as 2,000 rules using the Experta library, forming a knowledge structure derived from expert analysis that captures relevant relationships between symptoms and learning difficulties. This rule-based design supports a controlled and reliable diagnostic process, along with the requirement of professional validation before storing any diagnosis in the PostgreSQL database. The system's performance metrics, including precision (1.00), recall (0.90), specificity (1.00), accuracy (0.93), and F1-score (0.95), reflect a strong ability to identify cases while maintaining a conservative approach to positive classifications, which is particularly relevant in educational settings where minimizing incorrect labeling is essential.

When compared with related studies, expert system approaches such as those in references [8, 11] report accuracy values of 90% and 92%, respectively, using uncertainty-handling techniques such as CF and Dempster-Shafer theory. These approaches similarly rely on structured knowledge representation and demonstrate that rule-based or logic-driven systems can achieve consistent diagnostic performance, although their effectiveness depends on the design of the knowledge base and the quality of symptom modeling.

In contrast, machine learning–based approaches show different performance characteristics. Studies [9, 12] report accuracy values of 98.7% and 99.08% using ensemble learning models based on XGBoost, while traditional models evaluated in reference [10] achieve accuracy values around 93–94%, depending on the algorithm used. These methods benefit from data-driven learning and feature optimization, allowing them to capture complex patterns in diagnostic data. However, their performance is strongly influenced by dataset size, balance, and representativeness.

Additionally, behavior-based approaches such as the one presented in reference [13], which uses eye-tracking data and classification models, report an accuracy of 89.7% and recall of 84.8%. While this approach provides an alternative perspective based on real reading behavior, it also reflects the variability and complexity of interpreting behavioral signals in natural environments.

Overall, the reviewed studies report accuracy values ranging approximately from 87% to 99%, depending on the methodology and dataset characteristics. Within this context, the proposed system (accuracy 93%, F1-score 95%) falls within the performance range reported in the literature for diagnostic systems in this domain. The evaluation also indicates a small number of FN (FN = 8), which impacts sensitivity and highlights an area for further improvement. In contrast, the absence of FP is an important advantage, as it reduces the risk of false-positive classifications. This aspect is particularly relevant in educational and psychological settings, where overdiagnosis may lead to unnecessary interventions with potential academic and emotional consequences. Consequently, the system's high specificity supports a conservative diagnostic approach that prioritizes reliable identifications and helps maintain the integrity of the decision-making process.

In terms of practical application, the system is intended to be used as a decision-support tool in school environments, assisting teachers and psychologists in the early identification of learning disabilities. Its simple interface allows non-technical users to input observed student symptoms and obtain interpretable diagnostic suggestions. To ensure effective adoption, basic training would be required for educational staff on symptom recognition, system usage, and interpretation of results.

Finally, the results suggest that the system operates as a robust diagnostic support tool with performance comparable to existing approaches. Nevertheless, the presence of FN indicates that certain cases may not be detected, which could limit early identification and underscores the need for further refinement. Future work should focus on optimizing the rule base and expanding the symptom set to improve sensitivity and reduce missed cases. In addition, the evaluation was conducted on a limited sample, which may restrict the generalizability of the findings across different educational contexts. Variations in language, curriculum, and diagnostic criteria may influence system performance; however, the rule-based structure allows for adaptability through updates to the knowledge base for different populations or regions. As a direction for future work, it is recommended to validate the system in educational environments. Likewise, larger and more diverse samples should be considered, including students from different geographic regions, age groups, and demographic contexts, to improve the robustness, generalizability, and external validity of the proposed approach.

6. CONCLUSIONS

This research demonstrates that the developed expert system provides a robust and reliable framework for identifying dyslexia, dysgraphia, and dyscalculia. By integrating 50 specific symptoms and a comprehensive knowledge base of 2,000 rules via the Experta library, the system effectively models complex diagnostic patterns. The results offer a consistent alternative to other methodologies found in current literature, providing valuable decision support for professionals who maintain final oversight by validating each diagnosis before it is committed to the database. While the current evaluation was based on a specific set of cases, the findings underscore the significant potential of rule-based architectures in the early detection of learning disabilities. Moving forward, research should focus on expanding the scale and diversity of the study samples. Furthermore, exploring the integration of machine learning and deep learning techniques could complement expert inference, thereby enhancing the system's generalizability and effectiveness across a broader range of educational and clinical environments.

ACKNOWLEDGMENT

The authors gratefully acknowledge the Universidad de Ciencias y Humanidades for the institutional support provided for the development of this article. Appreciation is also extended to colleagues and experts whose suggestions and recommendations significantly contributed to improving the quality of the work.

REFERENCES

[1] American Psychiatric Association. (2022). Diagnostic and Statistical Manual of Mental Disorders. Fifth Edition, Text Revision (DSM-5-TR). American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787

[2] Andrade-Arenas, L., Yactayo-Arias, C. (2024). Expert system for diagnosing learning disorders in children. International Journal of Electrical and Computer Engineering, 14(3): 2965-2975. https://doi.org/10.11591/ijece.v14i3.pp2965-2975

[3] Adjiovski, B., Bogatinoska, D.C., Ismajloska, M., Malekian, R. (2024). Enhancing educational technology in lectures for school students with learning disabilities: A comprehensive analysis. SN Computer Science, 5(6): 716. https://doi.org/10.1007/s42979-024-03049-z

[4] Santhiya, S., KanimozhiSelvi, C.S. (2023). A study on dyslexia detection using machine learning techniques for checklist, questionnaire and online game based datasets. Applied and Computational Engineering, 5(1): 837-842. https://doi.org/10.54254/2755-2721/5/20230722

[5] Phudinawala, H., Batatawala, I., Samnani, R. (2024). Early identification of learning and developmental disorders in school-age children. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 10(5): 247-253. https://doi.org/10.32628/CSEIT241051022

[6] Nugraheni, M., Nuraini, R., Tonggiroh, M., Nurhayati, S. (2023). Expert system for diagnosing learning disorders in children using the Dempster-Shafer theory approach. Sinkron, 7(4): 2513-2525. https://doi.org/10.33395/sinkron.v8i4.12960

[7] Ana, A. (2020). Trends in expert system development: A

practicum content analysis in vocational education for over grow pandemic learning problems. Indonesian Journal of Science and Technology, 5(2): 246-260. https://doi.org/10.17509/ijost.v5i2.24616
[8] Anggrawan, A., Hairani, H., Satria, C., Dayani, A.D. (2023). Diagnosing learning disorders in children: A comparison of certainty factor and Dempster-Shafer methods. International Journal of Information and Education Technology, 13(9): 1422-1429. https://doi.org/10.18178/ijiet.2023.13.9.1945
[9] Alzahrani, S., Algahtani, F. (2025). Implications and identification of specific learning disability using weighted ensemble learning model. Child Care Health and Development, 51(1). https://doi.org/10.1111/cch.70026
[10] Kurniawan, Z., Tiaharyadini, R. (2024). Machine learning approach for early diagnosis of dyslexia among primary school children: A scoping review and model development. Indonesian Journal of Artificial Intelligence and Data Mining, 7(2): 425. https://doi.org/10.24014/ijaidm.v7i2.30614
[11] Mulakaluri, S., Gowdra Shivappa, G. (2024). An ensemble features aware machine learning model for detection and staging of dyslexia. IAES International Journal of Artificial Intelligence, 13(3): 3147-3156. https://doi.org/10.11591/ijai.v13.i3.pp3147-3156
[12] Raatikainen, P., Hautala, J., Loberg, O., Kärkkäinen, T., Leppänen, P., Nieminen, P. (2021). Detection of developmental dyslexia with machine learning using eye movement data. Array, 12: 100087. https://doi.org/10.1016/j.array.2021.100087
[13] Drotár, P., Dobeš, M. (2020). Dysgraphia detection through machine learning. Scientific Reports, 10(1): 21541. https://doi.org/10.1038/s41598-020-78611-9
[14] Bhavsar, K., Shah, V., Gopalan, S. (2020). Scrum: An agile process reengineering in software engineering. International Journal of Innovative Technology and Exploring Engineering, 9(3): 840-848. https://doi.org/10.35940/ijitee.C8545.019320
[15] Ekechi, C.C., Okeke, C.D., Adama, H.E. (2024). Enhancing agile product development with scrum methodologies: A detailed exploration of implementation practices and benefits. Engineering Science & Technology Journal, 5(5): 1542-1570. https://doi.org/10.51594/estj.v5i5.1108
[16] Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., Doulamis, N. (2021). Multiclass confusion matrix reduction method and its application on net promoter score classification problem. Technologies (Basel), 9(4): 81. https://doi.org/10.3390/technologies9040081
[17] Akula, U.S., Marimuthu, K., Devendra, N. (2024). Understanding the sensitivity, specificity, and predictive values used in diagnostic tests. Asian Journal of Pharmaceutical and Clinical Research, 17(8): 5-8. https://doi.org/10.22159/ajpcr.2024v17i8.51243