



## Implementation of the Jaro-Winkler Algorithm for Automatic Grading of Short Answer Questions in the Kifayatul Azhar Elementary School Exam

Rizki Halomoan<sup>1</sup>, Benny Yoga<sup>2</sup>, Nauval Ardana<sup>3</sup>, Trisna Gelar<sup>4\*</sup>, Santi Sundari<sup>5</sup>

<sup>1,2,3,4,5</sup>Department of Computer and Informatics Engineering, Politeknik Negeri Bandung, Indonesia  
Correspondence E-mail: [trisna.gelar@polban.ac.id](mailto:trisna.gelar@polban.ac.id)

### ABSTRACT

Kifayatul Azhar Primary School faces significant obstacles in examination administration. High logistical costs and inefficient social media coordination hamper physical exam distribution. Existing digital tools, such as Google Forms, cannot adapt to minor typographical errors in short-answer evaluations. This research develops a web-based automated grading system using the Jaro-Winkler algorithm and the Rational Unified Process (RUP) framework. The RUP methodology manages iterative development through four stages: Inception, Elaboration, Construction, and Transition. The system engine performs text preprocessing tasks such as case folding and tokenization prior to calculating Jaro-Winkler similarity scores. Experimental results indicate that  $d_w \geq 0.8$  serves as the optimal threshold for differentiating acceptable typographical variants from substantive errors. The system validation process included 26 integration test cases and 76 User Acceptance Test (UAT) cases, encompassing four distinct user roles: Admin, Teacher, Mentor, and Parent. All tests demonstrated a complete functional success rate. This system provides a centralized, automated environment that reduces administrative workload while maintaining grading objectivity.

© 2026 Universitas Pendidikan Indonesia

### ARTICLE INFO

#### Article History:

Submitted/Received 05 Mar 2026

First Revised 17 Mar 2026

Accepted 29 Mar 2026

First Available online 01 Apr 2026

Publication Date 01 Apr 2026

#### Keyword:

Automatic Grading,  
Elementary Education,  
Jaro-Winkler Algorithm,  
Rational Unified Process (RUP),  
Short Answer Questions.

## 1. INTRODUCTION

Kifayatul Azhar Primary School implements a flexible learning model through online-based Home Schooling.[1] Talent Class Program. The school organizes this program to serve 157 students scattered throughout various regions in Indonesia. For instance, students outside Java Island participate in all academic activities independently via digital platforms under parental supervision. The existence of this specialized program demands a support system capable of crossing geographical boundaries efficiently.[2], [3].

Digitalization of the evaluation system becomes the primary solution to reduce school operational costs.[4], [5]. The school management has been sending physical exam papers (hardcopy) to each student's home address using expedited services. This procedure causes a surge in logistics costs and delays in question distribution during Midterm (UTS) and Final Exams (UAS). The implementation of web-based exams can cut the physical distribution chain so that the evaluation process runs faster and more affordably.[6].

Online exam execution faces technical obstacles in the automation of short-answer scoring. [7]. The Google Form technology currently used only accommodates accurate assessment for multiple-choice question types. Teachers frequently find cases where the system absolutely penalizes student answers simply due to differences in capitalization or typographical errors (typos)[8]. For example, the system will award a score of zero to the answer "matahari" if the set answer key is "Matahari", even though the substance is identical.

The limitations of said platform increase the administrative workload for educators.[4]. Teachers must manually correct thousands of short-answer entries in subjects such as Islamic Education (PAI), Science (IPA), and Social Studies (IPS) to maintain assessment objectivity.[9]. Previous researchers have explored various string-matching algorithms to handle text writing variations automatically.[10]. One study indicates that the use of the N-Gram method can improve word error detection accuracy in Indonesian text by up to 83.64%[11].

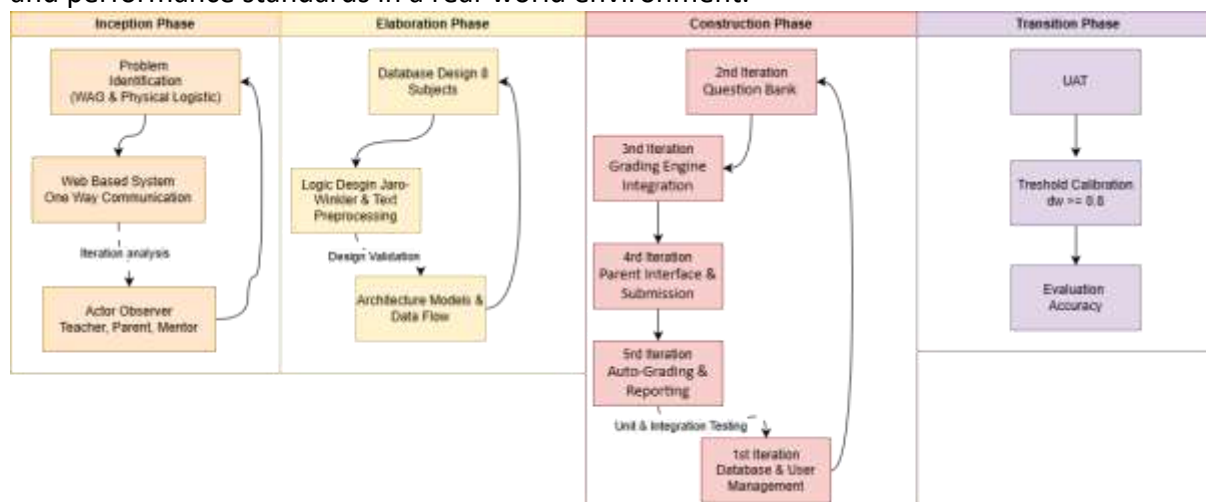
This research proposes the Jaro-Winkler algorithm as an automatic correction engine for short answers.[12]. The Jaro-Winkler algorithm calculates the similarity level between two strings by giving more weight to character matches at the beginning of the word (prefix matching). For example, the system will still provide a high similarity score for the word "Indonesia" even if the student types "Indoensia" due to the similarity of its first four characters. The researcher utilizes the Rational Unified Process (RUP) method in building this application to ensure the system's functionality meets specific field requirements.[13].

The developed application aims to automate assessments in eight core subjects, such as PAI, PPKN, and IPA. This software focuses the evaluation process on string similarity comparison and does not involve semantic or meaning analysis.[14]. The application of this algorithm is expected to significantly cut teacher correction time without ignoring tolerance for student typing errors. This step serves as a crucial strategy for SD Kifayatul Azhar in optimizing the efficiency of educational evaluation and improving corrective feedback in the digital era.[15].

## 2. METHODS

This research utilizes the Rational Unified Process (RUP) framework to manage the software development life cycle iteratively. The selection of RUP allows the researcher to align the specific needs of Kifayatul Azhar Primary School with the development of the Jaro-Winkler algorithm through four main phases. Figure 1 illustrates the iterative development lifecycle of the system following the RUP framework. The process begins with the Inception phase, focusing on stakeholder requirement analysis at SD Kifayatul Azhar. It proceeds to the

The development process begins with the Elaboration phase, where the system architecture and Jaro-Winkler algorithmic logic are meticulously designed. This architectural blueprint, which serves as the foundation for the entire development lifecycle, is illustrated in **Figure 1**. Following the design, the Construction phase is executed through five distinct iterations, ensuring each module is validated through rigorous unit and integration testing. Finally, the Transition phase involves field implementation and threshold calibration ( $d_w \geq 0.8$ ) through User Acceptance Testing (UAT) to ensure the system meets the required accuracy and performance standards in a real-world environment.



**Figure 1.** The four phases of the Rational Unified Process (RUP) framework were applied in the development of the automatic grading system.

## 2.1. Inception Phase

This phase establishes the project boundaries by observing the interactions among three main actors: Teachers, Mentors, and Parents. The learning system in the Talent Class Program at SD Kifayatul Azhar relies on mentors as an information bridge to distribute teaching modules and exam questions from teachers to parents. Current coordination activities rely entirely on WhatsApp Groups (WAG), leading to disorganized message stacks. Technical issues arise regarding the documentation of exam links and modules, which are frequently lost in conversations, making it difficult for teachers to conduct centralized grade recording at the end of the semester.

An in-depth analysis was conducted on the Midterm (UTS) and Final Exam (UAS) procedures, which require the distribution of physical exam papers. Mentors must prepare printed packages and use expedition services to reach 157 students scattered throughout Indonesia. This procedure not only burdens the school with logistics costs but also demands extra workload for mentors to ensure the delivery status of each package. Furthermore, the assignment submission method via photo uploads in WAG is highly ineffective, as mentors must manually verify the validity of files one by one. This condition prolongs the verification process for task submissions in every class.

Based on these findings, the system requirements are focused on a web-based application capable of providing one-way communication from mentors to parents regarding exam schedules and thematic activities. The system must accommodate answer inputs by parents according to the duration determined by the school. Key identified features include digital

exam submission management, a question bank with various types (multiple choice, short answer, and essay), and an automatic grade recording function. This phase concludes that the digitalization of exams is not only a solution to reduce paper distribution costs but also a strategy to improve academic documentation management that was previously fragmented across social media.

## 2.2. Elaboration Phase

This phase matures the system design through the integration of string-matching logic and database architecture. Activities began with modeling the data flow that connects answer inputs from the parent interface to the automatic grading engine. The architectural structure is designed to manage question variations from eight core subjects, including Islamic Religious Education (PAI), Science (IPA), and Civics (PPKN). The output of this phase consists of database schemas and sequence diagrams that ensure every answer input undergoes a validation process before being calculated by the algorithm.

The short-answer evaluation process applies systematic Text Preprocessing stages. Student answers are processed through two main steps: case folding and tokenization.[10]. Case folding converts all characters to lowercase to eliminate case sensitivity, which was a weakness in the previous platform. Subsequently, tokenization separates words based on spaces and punctuation to facilitate the identification of text units for comparison. Stop-word removal and stemming are not implemented in this system to maintain the integrity of the original words.[16], which often serve as primary keywords in elementary school students' short answers.

The Jaro-Winkler algorithm serves as the core component in determining the similarity score between student answers and the answer key. The calculation begins by determining the Jaro Distance ( $d_j$ ), which considers the number of matching characters ( $m$ ) and the number of transpositions ( $t$ ) between two strings ( $S_1$  and  $S_2$ ):

$$j \frac{1}{3} x \left( \frac{m}{S_1} + \frac{m}{S_2} + \frac{m-t}{m} \right) \quad (1)$$

The system enhances grading accuracy by applying weighting to the initial characters (prefix matching). The Jaro-Winkler formula  $d_w$  provides an additional score based on the length of the common prefix ( $l$ ) using a constant scaling factor ( $p=0.1$ ).

$$d_w = d_j + (p(1 - d_j)) \quad (2)$$

For instance, if a student's answer contains a typographical error at the end of the word but is correct at the beginning, the algorithm still yields a high score. Algorithmic simulations were conducted on actual answer samples from Talent Class students to determine the assessment threshold. The simulation results indicate that a value of  $d_w \geq 0.8$  is the optimal threshold for the 'Almost Correct' category tolerable by the system. This stage concluded with the creation of a functional mockup to verify the grading workflow before proceeding to full-scale coding.

## 2.3. Construction Phase

The system development was carried out gradually through five main iterations. Each iteration produced functional modules that integrated academic management requirements

with the automatic grading algorithm. The development focus during this phase was divided into several activities, which are summarized in **Table 1**.

**Table 1.** Development Iterations in the Construction Phase

Iteration	Development Focus	Main Activity
1	Database & User Management	Designing the schema for eight subjects and implementing a multi-actor authentication system (Teacher, Mentor, Parent).
2	Question Bank Module	Developing features for inputting and managing various question types: multiple-choice, short answer, and essays.
3	Grading Engine Integration	Integrating the Jaro-Winkler algorithm with the text preprocessing workflow (case folding and tokenization).
4	Parent Interface & Submission	Developing the digital exam submission portal, including real-time countdown timers and answer input handling.
5	Auto-Grading & Reporting	Implementing the automated scoring logic for short answers and generating centralized academic reports.

All modules developed within these five iterations underwent unit testing and integration testing. This process ensured that data synchronization between student answer inputs and the teacher's grade dashboard operated without logical conflicts during mass access. The use of these iterations guarantees that every feature, from the question bank to the automatic grading engine, was validated before proceeding to the field implementation stage.

#### 2.4. Transition Phase

The Transition Phase focuses its activities on final verification and system handover to the users. The primary objective is to ensure that the application can stably and accurately replace the physical exam distribution procedures. Validation was conducted through a User Acceptance Test (UAT) mechanism involving Teachers, Mentors, and Parents as end-users. For instance, users tested the entire workflow, from the distribution of exam schedules by mentors to the submission of answers by parents, to ensure the system interface is easily understood.

Testing the accuracy of the automated grading is a crucial parameter in this validation stage. Teachers performed a comparison between the scores generated by the Jaro-Winkler algorithm and the results of manual assessments across eight core subjects. This activity aimed to calibrate the similarity threshold to ensure fair grading for students despite typographical errors. For example, the system is declared valid if it can recognize variations in short answers that contain typos but maintain the correct substance according to the answer key.

### 3. RESULTS AND DISCUSSION

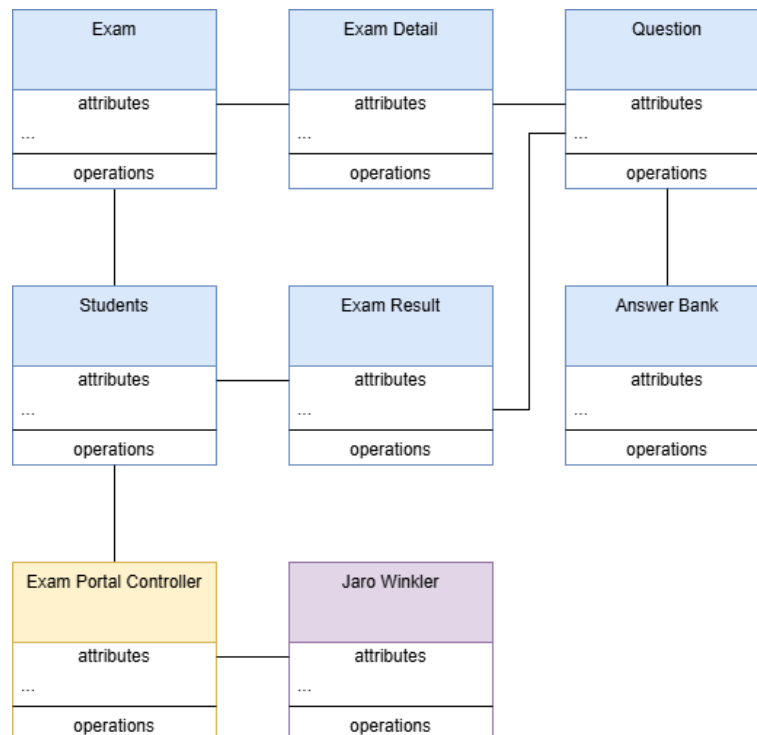
#### 3.1. Analysis of Requirements (Inception & Elaboration Outcomes)

As show in **Table 2**, The initial analysis identified that the primary challenge at SD Kifayatul Azhar was the high logistical cost of physical exams and the inefficiency of manual verification via social media. The system was designed to handle eight core subjects.

**Table 2.** Sample of Short Answer Keys for Assessment Calibration

Subject	Questions	Answer
PPKN	What is the first principle of the Pancasila?	Ketuhanan yang maha esa
PAI	What is the second source of Islamic law after the Al-Quran?	Hadist
Science (IPA)	What is the primary source of light and energy for the Earth?	Matahari
Science (IPA)	What type of mirror is used for vehicle rearview mirrors?	Cembung
Art (SBdP)	What is the term for the speed of a musical beat?	Tempo
Art (SBdP)	Art that has length, width, and height (volume) is called...	3D
Social Studies	Who is the authority responsible for regulating road traffic?	Polisi lalu lintas
PAI	What is the meaning of the word 'Zikir' in a literal sense?	Ingat
PJOK	What is the starting position for running?	Berdiri
Indonesian	What is the antonym of the word "soft" (lembut)?	Keras

The architectural design of the system was matured during the Elaboration phase to ensure seamless integration between the database and the grading engine. **Figure 2** presents the class diagram that defines the relationships between the Teacher, Mentor, and Parent actors.

**Figure 2.** Class Diagram of the Automated Grading System showing the structural relationship between users and the assessment modules.

### 3.1.1 Algorithmic Comparison and Accuracy Analysis

Based on the data provided by the partner regarding student answers during the final semester evaluation, an analysis was conducted to determine the appropriate method for correcting correct or incorrect answers. Several student answers contained typographical errors, leading to a series of experiments. The system's effectiveness was tested by comparing student inputs that contained typographical errors against the gold standard keys. Table 3 illustrates how the Jaro-Winkler algorithm maintains a high similarity score ( $d_w$ ) even when student input is slightly inaccurate.

**Table 3.** Comparison of Jaro-Winkler Scores ( $d_w$ ) for Student Answer Variations

No	Answer Key ( $S_1$ )	Student Input ( $S_2$ )	$d_w$ Score	Assessment Status
1	Matahari	matahari	1.00	Correct (Case folding applied)
2	Matahari	matahair	0.958	Correct (Minor typo)
3	Hadis	Hadis	1.00	Correct
4	Hadis	Habis	0.866	Correct (Minor typo)
5	Cembung	Cembun	0.94	Correct (Missing character)
6	Tempo	Temppo	0.93	Correct (Double character)
7	Ketuhanan yang maha esa	Ketuhanan maha esa	0.88	Incorrect (Missing word)
8	Ketuhanan yang maha esa	Kemanusiaan yang adil dan beradab	0.655	Incorrect (Substantive error)
9	keras	pelan	0.600	Incorrect (Substantive error)
10	ingat	lupa	0.483	Incorrect (Substantive error)
11	Polisi lalu lintas	Polis lalu lintas	0.904	Correct (Minor typo)
12	3D	3 dimensi	0.750	Incorrect (System)

As seen in **Table 3**, the Jaro-Winkler algorithm performs exceptionally well in handling student typos. For the answer "Hadis" instead of "Hadist," the system yields a score of 0.96, which is well above the established threshold of  $d_w \geq 0.8$ . This allows the system to grant a score of 100 to the student automatically. Conversely, for substantive errors like *pelan* (where the key is "keras"), the score drops significantly to 0.60, allowing the system to correctly mark the answer as wrong. A unique case was observed in the comparison between '3D' and '3 Dimensi'. Although both terms are semantically identical in the context of elementary school art subjects, the Jaro-Winkler algorithm yielded a low similarity score ( $d_w \approx 0.75$ ). This occurs because the algorithm calculates lexical similarity based on character sequences rather than semantic meaning. Since only the character '3' matches, the score falls below the 0.8 threshold. **Figure 3** showcases the standalone prototype developed to verify the Jaro-Winkler calculation logic.



**Figure 3.** Standalone Prototype of the Jaro-Winkler Distance Calculator used for algorithmic verification during the Elaboration phase.

### 3.1.2 Analysis of Results

The experimental results indicate that the Jaro-Winkler algorithm effectively handles most student typographical errors. By setting a threshold of  $dw \geq 0.8$ , the system successfully identifies answers that are lexically like the gold standard, such as "Hadis" and "Cembun," which would otherwise be marked incorrect by standard exact-match systems. This flexibility significantly reduces the administrative burden on mentors at SD Kifayatul Azhar, allowing for a more objective and faster grading process.

Despite its effectiveness in handling typos, this study identifies a significant limitation regarding semantic variation. As shown in Table 3 (No. 12), the pair "3D" and "3 Dimensi" yielded a low similarity score of 0.53 despite being semantically identical. This occurs because the Jaro-Winkler algorithm operates solely on character sequences (lexical matching) and does not possess natural language understanding (semantic matching).

Consequently, for subjects requiring technical abbreviations or synonyms, the current system still requires a "multiple-key" database approach where teachers input various acceptable variations (e.g., both "3D" and "3 Dimensi") as separate gold standards to ensure fair grading.

$$\begin{aligned} \text{True,} & \quad \text{similarity} = 1 \text{ Expert Judgment, } 0.8 < \text{similarity} < 1 \text{ False,} \\ & \quad \text{similarity} \leq 0.8 \end{aligned}$$

### 3.2. System Implementation and Iterative Development (Construction Outcome)

The Construction phase focused on the phased implementation of the system requirements through five distinct iterations. This approach ensured that the core grading engine was stable before the user interface was finalized.

#### 3.2.1 Functional Iterations

The development was divided into functional modules as detailed previously in **Table 1**. The early iterations established a robust database capable of storing diverse question types (Multiple Choice, Short Answer, and Essay) for eight core subjects. By Iteration 3, the Jaro-Winkler algorithm was integrated into the grading logic, preceded by a mandatory text preprocessing pipeline, as illustrated in **Figure 4** below. This integrated workflow ensures that the input data is properly refined before the algorithm calculates the final similarity Scores



**Figure 4.** Integration of the Jaro-Winkler Algorithm in the Admin Question Bank Module

### 3.2.2 Automated Grading Workflow

The core functionality of the system involves comparing student submissions against the stored "Gold Standard" keys. **Figure 5** illustrates the system's ability to calculate similarity scores in real-time. During this phase, unit testing confirmed that the system correctly applied case-folding (e.g., converting "cembung" and "kembang" to a common format) to avoid unnecessary grading penalties.



**Figure 5.** Real-time Jaro-Winkler Similarity Calculation during the Grading Process

### 3.2.3 Data Synchronization and Integration Testing

During the final iterations of the Construction phase, the system underwent rigorous Unit Testing and Integration Testing to ensure data synchronization across all modules. A total of 26 test cases were executed using the Black Box method, covering authentication, data management, and the core system functionalities, as summarized in **Table 4** below. This structured evaluation confirmed that each feature performs according to the specified requirements before moving into the deployment stage.

**Table 4.** Summary of Functional and Integration Testing Results

Test Code	Module / Function	Testing Objective	Status
UT-01 to UT-05	Authentication	Validating login security for Admin, Teacher, and Parents.	PASS
IT-06 to IT-15	User & Data Mgmt	Ensuring seamless data flow between student profiles and class assignments.	PASS

IT-16 to IT-21	Question Bank	Verifying the storage and retrieval of Short Answer keys for 8 subjects.	PASS
IT-22 to IT-26	Grading Engine	Testing the synchronization between student input and automatic $d_w$ calculation.	PASS

The integration tests confirmed that the data entered by Parents (student answers) was accurately captured and processed by the Jaro-Winkler engine. As shown in the test results (e.g., IT-23), modifications to the question bank were immediately reflected in the grading dashboard without logical conflicts. This synchronization is critical for the Transition phase, as it ensures that the teacher's grade recap is always based on the most recent and valid student submissions.

### 3.3. Final Validation and User Acceptance (Transition Outcome)

The Transition phase concluded with a comprehensive User Acceptance Test (UAT). The objective was to ensure that the system met the operational needs of all four primary actors at SD Kifayatul Azhar. A total of 76 test cases were executed across different user roles, covering end-to-end workflows from exam scheduling to automated grading. The results of these evaluations, as detailed in **Table 5**, demonstrate a 100% success rate in functional execution, with 90.7% of cases passing flawlessly and 9.3% passing with minor UI/UX notes (such as minor layout adjustments).

**Table 5.** Summary of User Acceptance Test (UAT) Results

No	User Role	Total Test Cases	Status: Pass
1	Teacher	24	21
2	Parent	13	13
3	Admin	12	9
4	Mentor	27	26

Specifically, the Teacher role confirmed that the automated grading correctly applied the Jaro-Winkler logic, and the Parent role successfully submitted answers without technical hurdles. The notes provided by the Admin and Mentor roles were primarily focused on decorative interface elements, which did not affect the core functionality of the exam distribution or the accuracy of the similarity scores. This phase confirms that the system is ready for full-scale implementation, effectively replacing the previous manual WhatsApp-based coordination with a structured, automated academic environment.

## 4. CONCLUSION

This study employed a web-based automated grading system at Kifayatul Azhar Primary School, utilizing the Rational Unified Process (RUP) and the Jaro-Winkler algorithm. The system was designed to mitigate the significant logistical expenses and disorganized coordination that had previously been managed through WhatsApp Groups. The resulting platform effectively centralized examination management and enhanced the precision of documentation. Experimental findings validated that a similarity threshold of  $d_w \geq 0.8$  offered the most effective equilibrium for identifying student typographical errors. Furthermore, validation through 76 User Acceptance Test (UAT) cases revealed a 100% functional success rate across all user roles. Although the Jaro-Winkler algorithm facilitates objective lexical assessment, it is constrained by its inability to recognize semantic synonyms. Consequently,

future iterations should incorporate natural language processing to improve the system's comprehension of contextual variations within student responses.

## 5. ACKNOWLEDGMENT

The authors would like to express their gratitude to Politeknik Negeri Bandung (POLBAN) for the institutional support provided during this research. We also extend our sincere appreciation to the faculty and staff of SD Kifayatul Azhar for their collaboration as research partners and for providing the necessary data and environment for the system's implementation and testing.

## 6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

## 7. REFERENCES

- [1] Nor Annisa, Noor Padilah, Reni Rulita, Renny Yuniar, and Nita Priyanti, "Model Pembelajaran Homeschooling Pendidikan Anak Usia Dini," *EDUKASIA J. Pendidik. dan Pembelajaran*, vol. 4, no. 1, pp. 89–100, Aug. 2022, doi: 10.62775/edukasia.v4i1.106. Available: <https://jurnaledukasia.org/index.php/edukasia/article/view/106>
- [2] M. Munawarah, I. Ikhsan, L. N. W. Fajzrina, R. Aulia, and C. N. Fadillah, "Strategy for Accepting New Students Post-COVID-19: A Case Study," *Al-Athfal J. Pendidik. Anak*, vol. 8, no. 2, pp. 105–116, Mar. 2023, doi: 10.14421/al-athfal.2022.82-03. Available: <https://ejournal.uin-suka.ac.id/tarbiyah/alathfal/article/view/6330>
- [3] M. Lutviatiani, N. C. Shofiyah, C. F. Wijaksono, and A. A. Zawawi, "The Implementation of an Online Distance Learning Program with an Independent Curriculum-Based on Improving Child-Parent Relations," *Educ. J. Ilm. Pendidik.*, vol. 1, no. 2, pp. 51–62, Jul. 2023, doi: 10.37985/educative.v1i2.56. Available: <https://publikasi.abidan.org/index.php/educative/article/view/56>
- [4] H. W. Dhany and S. Sapriadi, "Pengembangan Perangkat Lunak Penilaian Otomatis Ujian Pilihan Ganda Menggunakan Algoritma String Matching," *J. Minfo Polgan*, vol. 14, no. 2, pp. 2768–2774, Nov. 2025, doi: 10.33395/jmp.v14i2.15452. Available: <https://jurnal.polgan.ac.id/index.php/jmp/article/view/15452>
- [5] M. Sabri and A. Wais, "Artificial Intelligence dalam Pengukuran dan Penilaian Pendidikan," *J. Eval. Pendidik*, vol. 16, no. 2, pp. 95–107, Oct. 2025, doi: 10.21009/jep.v16i2.60738. Available: <https://journal.unj.ac.id/unj/index.php/jep/article/view/60738>
- [6] D. P. Sidik, A. Rozak, F. Fatqurhohman, and H. Fatkurochman, "Literature Review of Artificial Intelligence in Learning: Trends and Opportunities," *RESET Rev. Educ. Sci. Technol.*, vol. 1, no. 1, pp. 43–54, Aug. 2025, doi: 10.66031/reset.v1i1.19. Available: <https://jurnal.ihsancahayapustaka.id/index.php/reset/article/view/19>
- [7] K. Harefa and A. J. Lubis, "Aplikasi Sistem Automated Essay Scoring untuk Jawaban Soal Ujian dengan Menerapkan Algoritma Jaro-Winkler," *J. Ilmu Komput. Dan Sist. Komput. Terap.*, vol. 4, no. 2, pp. 109–119, 2022, doi: <https://doi.org/10.35447/jikstra.v4i2.611>
- [8] A. R. Amri, C. Charlina, and M. Sinaga, "Kesalahan Pemakaian Ejaan Bahasa Indonesia

- dalam Soal Try Out Ujian Sekolah Berstandar Nasional Sekolah Dasar Tahun Pelajaran 2017/2018," *J. TUAH Pendidik. dan Pengajaran Bhs.*, vol. 2, no. 1, pp. 1–9, 2020, doi: <https://dx.doi.org/10.31258/jtuah.2.1.p.1-9>
- [9] N. S. Lagutina and K. V. Lagutina, "A Survey of Models for Automatic Assessment of Similarity of Students' Answer to the Reference Answer," *Autom. Control Comput. Sci.*, vol. 59, no. 7, pp. 1152–1169, Dec. 2025, doi: 10.3103/S0146411625700427. Available: <https://link.springer.com/10.3103/S0146411625700427>
- [10] Y. Rochmawati and R. Kusumaningrum, "Studi Perbandingan Algoritma Pencarian String dalam Metode Approximate String Matching untuk Identifikasi Kesalahan Pengetikan Teks," *J. Buana Inform.*, vol. 7, no. 2, Jan. 2016, doi: 10.24002/jbi.v7i2.491. Available: <https://ojs.uajy.ac.id/index.php/jbi/article/view/491>
- [11] H. Jayadianti, B. Santosa, J. Cahyaning, S. Saifullah, and R. Drezewski, "Essay Auto-Scoring Using N-Gram And Jaro Winkler Based Indonesian Typos," *MATRIK J. Manajemen, Tek. Inform. Dan Rekayasa Komput.*, vol. 22, no. 2, pp. 325–338, Mar. 2023, doi: 10.30812/matrik.v22i2.2473. Available: <https://journal.universitاسbumigora.ac.id/index.php/matrik/article/view/2473>
- [12] W. E. Winkler, "Record Linkage," 2009, pp. 351–380. doi: 10.1016/S0169-7161(08)00014-X. Available: <https://linkinghub.elsevier.com/retrieve/pii/S016971610800014X>
- [13] T. K. Tia, I. Nuryasin, and M. Maskur, "Model Simulasi Rational Unified Process (RUP) Pada Pengembangan Perangkat Lunak," *J. Repos.*, vol. 2, no. 4, Jan. 2024, doi: 10.22219/repositor.v2i4.30511. Available: <https://ejournal.umm.ac.id/index.php/repositor/article/view/30511>
- [14] A. Ayaan and K.-W. Ng, "Automated Grading Using Natural Language Processing And Semantic Analysis," *MethodsX*, vol. 14, p. 103395, Jun. 2025, doi: 10.1016/j.mex.2025.103395. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2215016125002419>
- [15] I. M. A. Purwadi, "Pengaruh Pemberian Umpan Balik Dari Teman Sejawat Terhadap Kemampuan Berpikir Kritis Dan Komunikasi Matematis Siswa Smp Laboratorium Undiksha Singaraja," *J. Pendidik. Mat. Undiksha*, vol. 11, no. 1, pp. 63–72, 2020, doi: <https://doi.org/10.23887/jjpm.v11i1.24382>
- [16] T. Mardiana, T. Bharata Adji, and I. Hidayah, "Stemming Influence on Similarity Detection of Abstract Written in Indonesia," *TELKOMNIKA (Telecommunication Comput. Electron. Control)*, vol. 14, no. 1, p. 219, Mar. 2016, doi: 10.12928/telkomnika.v14i1.1926. Available: <http://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/view/1926>