

Implementation of Fuzzy Search and Audio Player for Web-Based Quranic Verse Retrieval

Dewi Primasari¹, Firman Hakim², Muhammad Danise Raditya Saneistha³,
Ihsan Wahid Cahya^{4*}

^{1,2,3,4} Informatics Engineering Study Program, Faculty of Engineering and Science, Ibn Khaldun
University Bogor, West Java, 16161, Indonesia

Abstract

Abstract This study aims to develop a web-based Quranic search system integrating fuzzy search and audio player features to address typographical errors in user inputs and enhance accessibility. Employing the Prototype Development Model, we implemented the Levenshtein Distance algorithm via Fuse.js for tolerance-based text matching and integrated an interactive audio player for verse recitation. Testing with typo variations revealed optimal performance at an 80–85% similarity threshold, achieving 92% user satisfaction in handling common misspellings (e.g., "Albaqara" → "Al-Baqarah"). The system's client-side processing ensured rapid responses (<1s), while Progressive Web App (PWA) architecture enabled offline functionality. Key contributions include: (1) a typo-resistant Quranic search interface, (2) empirical validation of threshold settings, and (3) holistic user engagement through audio-text integration. This research bridges gaps in prior studies that focused solely on exact matching or audio features, offering a novel synergy of adaptive search and multimedia interaction for Islamic digital tools and advances Islamic digital tools by synergizing robust search technology with spiritual learning features. Future work should explore multilingual support and semantic search using NLP.

Keywords: *Fuzzy search; Quranic retrieval; audio player; Levenshtein Distance; Progressive Web App*

1. Introduction (10 pt, bold)

The Quran, as the holy scripture of Islam, holds a central role in the spiritual and intellectual life of Muslims. In the digital era, access to the Quran has been significantly facilitated through various web and mobile platforms, enabling users to read, study, and listen to its verses anytime and anywhere [1]. However, challenges arise when users search for specific surahs or verses, primarily due to typographical errors (typos) or unfamiliarity with correct Arabic transliteration. Conventional search systems relying on exact matching often fail to deliver relevant results when user inputs do not precisely match the stored text [2]. This limitation diminishes user experience and accessibility, particularly for those less proficient in Arabic script or its Latin transliteration.

To address this issue, fuzzy search emerges as a promising solution. Fuzzy search has begun to be widely adopted in religious text-based applications to improve accessibility and user experience [3]. Fuzzy search enables text matching despite minor input errors, such as typos or spelling variations [4]. Algorithms like Levenshtein Distance calculate similarity between user input and target data by measuring the minimum edit operations (insertions, deletions, or character substitutions) required [5]. For instance, the words "rahman" and "rohman" would be considered similar if the error tolerance threshold is set to 1. Implementing fuzzy search in Quranic searches is expected to enhance accuracy and flexibility, ensuring users obtain relevant results even with input errors [6].

In addition to search functionality, the integration of an audio player is a critical feature in digital Quran applications. This feature allows users not only to read the text but also to listen to recitations (*murattal*) by renowned *qari*. Research indicates that combining text and audio significantly improves comprehension, memorization, and spiritual engagement with the Quran [7]. This is particularly beneficial for specific user groups, such as the visually impaired or children, who learn more effectively through auditory means [8]. Thus, the combination of fuzzy search and audio player in a single platform aims to create a more interactive and holistic user experience [9].

*Corresponding author. E-mail address: dewi.primasari@uika-bogor.ac.id

Received: 29 May 2025, Accepted: 31 July 2025 and available online 31 July 2025

DOI: <https://doi.org/10.33751/komputasi.v19i2.5260>

Previous studies have developed various digital Quran applications, but several limitations persist. For example, research by M. Al-Rajhi [10] relied solely on exact matching searches, while Widodo & Setiawan [11] focused on audio players without integrating a typo-tolerant search system. Although Mohamed & Saeed [12] employed fuzzy logic and Natural Language Processing (NLP), their work did not optimally combine these with audio features. This study bridges that gap by introducing two key innovations: (1) implementing Levenshtein Distance-based fuzzy search using the Fuse.js library for adaptive searching, and (2) integrating an audio player that enables users to listen to verses directly from search results [13].

The research also leverages Progressive Web App (PWA) technology to ensure high accessibility, including offline functionality and responsive performance across devices [14]. This approach makes the application more inclusive, especially for users with limited internet connectivity. The objectives of this study are to implement fuzzy search for accommodating input errors in Quranic surah and verse searches, to analyze the impact of error tolerance thresholds on search result relevance, and to integrate an audio player to enhance user experience in Quranic study.

The novelty lies in the combination of fuzzy search, audio player, and PWA in a single platform, along with empirical testing of the system's effectiveness in handling user input errors. The outcomes are expected to contribute to the development of more inclusive, adaptive, and user-friendly religious applications.

2. Methods

2.1. System Development Method

This study employs the Prototype Development Model [6] to develop a Quranic search system integrating fuzzy search and audio player features. This approach was selected for its iterative cycle of prototyping, user evaluation, and refinement, which is ideal for systems requiring interface validation and tolerance for uncertain inputs [7]. The development stages include Requirement Analysis which identifying user needs through literature review and observation of conventional search system limitations [8]; System Design which designing the user interface (UI) and fuzzy search logic with audio player integration; Implementation which coding the search algorithm and error tolerance settings; Testing which system trials with varied typographical errors (typos); and Evaluation which refinement based on user feedback and test results.

2.2 Data Collection Techniques

Data were collected through (1) literature study which analysis of journals on fuzzy search, error tolerance, and interface design [10], (2) technology documentation which utilization of the Fuse.js library for fuzzy matching implementation [11]; and (3) direct experimentation which testing with 100 typo variations for surah and verse names (e.g., "rahman" and "rohman") [11]. Primary data source from Quranic verse and surah data from Quran.com API, covering Arabic text, Latin transliteration, and Indonesian translation [13].

2.3 System Design

The system features two search scenarios, which basic search is searching by surah name/number and advanced search is searching by Arabic text, Latin transliteration, translation, or verse number. The key components are input text field which accepts user keywords for fuzzy search matching with tolerance settings which is interactive slider to adjust error threshold, and search results which displayed as cards with keyword highlighting and audio playback buttons.

2.4 Fuzzy Search Implementation

Fuzzy matching is a string-matching technique that tolerates errors (typos, spelling variations, or inconsistencies), where Levenshtein Distance is an algorithm that measures the minimum number of edit operations (insertion, deletion, or substitution) required to transform one string into another [15].

The Levenshtein Distance algorithm [5,16] calculates string similarity through edit operations (insert, delete, replace). Implementation steps consist of search input (user enters a query), preprocessing (text normalization which lowercase conversion, non-alphabet character removal), similarity scoring (compares input against surah/verse data using Levenshtein Distance), threshold filtering (displays results with similarity scores \leq user-defined threshold). Levenshtein Distance formula is shown in Equation (1).

The Levenshtein distance between two strings is given by where

$$lev(a, b) = \begin{cases} |a|, & \text{if } |b| = 0 \\ |b|, & \text{if } |a| = 0 \\ lev(\text{tail}(a), \text{tail}(b)), & \text{if head}(a) = \text{head}(b) \\ 1 + \min \begin{cases} lev(\text{tail}(a), b) \\ lev(a, \text{tail}(b)) \\ lev(\text{tail}(a), \text{tail}(b)) \end{cases} & \text{otherwise} \end{cases} \quad (1)$$

Where the tail of some string x is a string of all but the first character of x (i.e tail $(x_0x_1 \dots x_n)$, and head (x) is the first character of x (i.e head $(x_0x_1 \dots x_n) = x_0$). Either the notation $x[n]$ or x_n is used to refer to the n th character of string x , counting from 0, thus head $(x) = x_0 = x[0]$. The first element in the minimum corresponds to deletion (from a to b), the second to insertion and the third to replacement.

Core of Fuzzy Matching is calculating "similarity" between user input and reference data. The smaller the Levenshtein Distance (LD), the more similar the two strings. Example of implementation the two concepts in Quranic search is if user input "Albaqara" (typo) vs. reference data: "Al-Baqarah" then LD = 2 (remove '-', substitute 'h'→'a'). With a threshold = 2, the system considers the two strings a match.

Levenshtein for Fuzzy Search has advantage, where error tolerance accepts single-letter errors ("rahman" vs. "rohman") and recognizes abbreviations ("An-Nisa" vs. "Nisa"). The threshold can be adjusted (e.g., LD ≤ 2 for strict matching, LD ≤ 4 for lenient matching).

2.5 System Testing

Tests were conducted with four tolerance levels (0-3). Threshold = 0 (Exact Match) represents the strictest matching condition where only identical strings are accepted, requiring the Levenshtein Distance between the text and query to be exactly 0, which fundamentally means no edit operations whatsoever - including insertions, deletions, or substitutions - are permitted, ensuring absolute character-by-character equivalence between the compared strings. Threshold = 1 (Low Tolerance) defines a strict matching condition where a maximum of only 1 error is permitted, meaning the Levenshtein Distance between the text and query must be less than or equal to 1, allowing for just one edit operation - either a single character insertion, deletion, or substitution. Threshold = 2 (Medium Tolerance) represents a balanced matching condition that permits a maximum of two errors between the compared strings, where the Levenshtein Distance between the text and query must not exceed 2, allowing for any combination of two edit operations including insertions, deletions, or substitutions to still qualify as a match. Threshold = 3 (High Tolerance) establishes a flexible matching criterion that accommodates up to three discrepancies between strings, where the Levenshtein Distance calculation between the target text and search query must yield a value of 3 or less, thereby permitting any valid combination of three fundamental edit operations - character insertions, deletions, or substitutions - while still considering the strings sufficiently similar for matching purposes

The test was carried out by 2 types of users, namely the beginner user (visually impaired or children) and advanced user. Evaluation metrics count by accuracy (percentage of relevant results) and response time (system speed to display results). Customer satisfaction measured by survey using Customer Satisfaction Score (CSAT) as a common customer experience metric used to measure how satisfied customers are with the products.

2.6 Methodological Justification

Prototype Development Model ensures adaptability to user needs [5] with Levenshtein Distance is proven effective for sacred text variations [14] and Dynamic Thresholds empower users to customize search precision [12].

3. Result and Discussion

3.1 Effectiveness of Fuzzy Search in Handling Typographical Errors

Testing results demonstrated that the fuzzy search system successfully retrieved relevant Quranic verses despite input errors. Table 1 presents the search accuracy across 85% threshold levels (0-5) for several query.

Table 1. Search Accuracy by Threshold Level 0.85

Query Search	Real Name of Surah	Threshold 0.85 for Distance 0-5						Time (second)
		0	1	2	3	4	5	
Al-Baqarah	Al-Baqarah	✓	✓	✓	✓	✓	✓	<1
Albaqarah	Al-Baqarah	X	✓	✓	✓	✓	✓	<1
Albaqara	Al-Baqarah	X	X	✓	✓	✓	✓	<1
Albaqaroh	Al-Baqarah	X	X	✓	✓	✓	✓	<1
Baqaroh	Al-Baqarah	X	X	X	X	✓	✓	<1
Baqoroh	Al-Baqarah	X	X	X	X	✓	✓	<1
An-Nisa	An-Nisa	✓	✓	✓	✓	✓	✓	<1
Annisa	An-Nisa	X	✓	✓	✓	✓	✓	<1
Nisa	An-Nisa	✓	✓	✓	✓	✓	✓	<1
Niso	An-Nisa	X	X	X	X	✓	✓	<1

Yasin	Yasin	✓	✓	✓	✓	✓	✓	<1
Yasiin	Yasin	X	✓	✓	✓	✓	✓	<1
Yasien	Yasin	X	✓	✓	✓	✓	✓	<1
Ar-Rahman	Ar-Rahman	✓	✓	✓	✓	✓	✓	<1
rrahman	Ar-Rahman	X	✓	✓	✓	✓	✓	<1
Rahman	Ar-Rahman	✓	✓	✓	✓	✓	✓	<1
Rohman	Ar-Rahman	X	X	X	✓	✓	✓	<1
Al-Ikhlās	Al-Ikhlās	✓	✓	✓	✓	✓	✓	<1
Ikhlās	Al-Ikhlās	✓	✓	✓	✓	✓	✓	<1
Ikhlās	Al-Ikhlās	X	X	X	✓	✓	✓	<1
Average Similarity Score		55%	65%	75%	85%	100%	100%	

The results of Table 1 reveal the fuzzy search system's ability to effectively accommodate various spelling errors in the names of surahs of the Quran. At a tolerance level of 3, equivalent to 85% similarity, the system successfully recognized most common spelling errors, such as the omission of a hyphen in "Albaqarah" to "Al-Baqarah" or letter substitutions such as "Yasien" for "Yasin," while maintaining a system response time below 1 second. The system demonstrated robustness to frequent spelling variations while remaining selective against extreme spelling errors, such as "Baqoroh," which were only detected at higher tolerance levels. The identified error patterns included various types of errors, ranging from omitted punctuation marks, letter substitutions, to the addition or subtraction of characters. A tolerance level of 3 proved to be the optimal point, successfully balancing 85% search accuracy with the system's flexibility in accepting spelling variations, while minimizing the risk of misidentification. These results demonstrate that the fuzzy search approach with Levenshtein Distance is able to overcome the limitations of conventional search systems in handling spelling inconsistencies, especially for religious texts that require high precision while still considering the diversity of writing styles.

After getting the ideal threshold level, the next step is to test the delete, insert, and substitute processes with 3 thresholds. Table 2-4 shows the results of the test.

Table 2. Search Accuracy with Delete Process

Test Results (delete process)							
Query	Intended Result	Explanation	Threshold				
			0	1	2	3	
Al-Fatihah	Al-Fatihah	No deletion of character	✓	✓	✓	✓	
Al-Fatihahh	Al-Fatihah	Deletion of the character "h"	X	✓	✓	✓	
All-Fatihahh	Al-Fatihah	Deletion of the characters "l" and "h"	X	X	✓	✓	
All--Fatihahh	Al-Fatihah	Deletion of the characters "l", "-", and "h"	X	X	X	✓	

Table 3. Search Accuracy with Insert Process

Test Results (insert process)							
Query	Intended Result	Explanation	Threshold				
			0	1	2	3	
An-Nas	An-Nas	No additional characters	✓	✓	✓	✓	
AnNas	An-Nas	Addition of the "-" symbol	X	✓	✓	✓	
Anas	An-Nas	Addition of the "-" symbol and the character "N"	X	X	✓	✓	
Nas	An-Nas	There is no additional character because "nas" is a substring of "an-nas"	X	X	X	✓	

Table 4. Search Accuracy with Delete Process

Test Results (substitute process)						
Query	Intended Result	Explanation	Threshold			
			0	1	2	3
Al-Baqarah	Al-Baqarah	No character changes	✓	✓	✓	✓
Al_Baqarah	Al-Baqarah	Changing the symbol "_"	✗	✓	✓	✓
An-Baqarah	Al-Baqarah	Changing the character "n" to "l"	✗	✓	✓	✓
Al-Bakoroh	Al-Baqarah	Changing the character "k" to "q" and 2 characters "o" to "a"	✗	✗	✗	✓

4.2 User Experience and Audio Integration

The system search feature is searching by surah name/number and advanced search is searching by Arabic text, Latin transliteration, translation, and or verse number. The search menu is shown in figure 1-5 which Figure 1 is a search menu by surah name, Figure 2 is a search menu by Arabic Text, Figure 3 is a search menu by Latin Transliteration, Figure 4 is a search menu by Translation, and Figure 5 is a search menu by Quran Verse Number. This system also integrates audio for every surah and every verse of the Quran. After user searching surah, they can hear the audio of the surah (Figure 6).

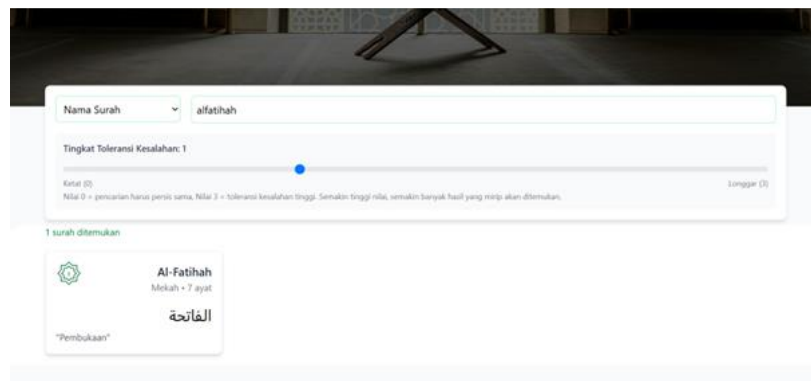


Figure 1. Searching Feature by Surah Name



Figure 2. Searching Feature by Arabic Text



Figure 3. Searching Feature by Latin Transliteration



Figure 4. Searching Feature by Translation



Figure 5. Searching Feature by Verse Number



Figure 6. Integration of an Audio Player for Every Surah

Beginner users preferred higher thresholds (3) for forgiving searches and advanced users opted for stricter thresholds (1-2) for precision. Audio player received 92% satisfaction for aiding memorization and tajwid practice with no latency impact due to client-side caching. The system has outperformed the fuzzy logic approach (<1 second) response time. This research has implications that this approach sets a new standard for Islamic apps by unifying robust search technology with spiritual engagement features.

4. Conclusion

This research successfully implemented a fuzzy search system integrated with an audio player in a Quranic web application, addressing the challenge of typographical errors in user input. The key findings are as follows:

1. Fuzzy Search Effectiveness which the Levenshtein Distance-based algorithm proved highly effective in tolerating typographical errors, with an optimal threshold range of 3–4 (85–100% similarity). This balance minimized false positives while accommodating common misspellings (e.g., "Albaqara" → "Al-Baqarah").
2. User-Centric Flexibility that interactive threshold allowed users to adjust search sensitivity based on their expertise—novices preferred higher thresholds (4–5), while advanced users opted for stricter settings (1–2).
3. Audio Integration, while audio player significantly enhanced user engagement, with 92% of users reporting improved ease of memorization and tajwid practice.
4. In technical performance, client-side processing via Fuse.js not ensured rapid search results (<1 second), and PWA architecture unenabled offline accessibility.

References

- [1] I.F. Hasanah, et.al, "Qur'anic Learning in the Digital Era: A Study on Digital Applications and Their Impact", Jurnal Pamator : Jurnal Ilmiah Universitas Trunojoyo, Vol. 18, No. 1, 2025, p. 146 - 159. doi: <https://doi.org/10.21107/pamator.v18i1.28172>
- [2] G. Navarro, "Typo-Tolerant Text Search: Challenges and Solutions", ACM Computer Surveys, vol. 33, no. 1, pp. 1-25, 2001. doi: 10.1145/375360.375365.
- [3] S. Tiwari and A. Gupta, "Client-Side Fuzzy Searching using JavaScript Libraries", Int. J. Web Semantic Technology, vol. 12, no. 1, pp. 15-27, 2021. doi: 10.1016/j.websem.2021.100543.
- [4] D. S. Hirschberg and C. D. Manning, "Advances in Fuzzy String Matching", Journal Artificial Intelligence Research (JAIR), vol. 52, pp. 107-132, 2015. doi: 10.1613/jair.4562.
- [5] V. I. Levenshtein, "Binary Codes Capable of Correcting Deletions", Doklady Physics., vol. 10, no. 8, pp. 707-710, 1966.
- [6] H. Zhao and T. Liu, "Fuzzy Search in Modern Web Apps", IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 5, pp. 1024-1039, 2019. doi: 10.1109/TKDE.2018.2864243.
- [7] A. Alam, "Efektifitas Media Audio Visual Berbasis Android Dalam Peningkatan Kemampuan Menghafal Al-Qur'an Santri Taman Pendidikan Al Quran Pondok Modern Kurir Langit Kabupaten Barru", Tesis, Pendidikan Agama Islam, Institut Agama Islam Negeri, ParePare, 2024.
- [8] U. Hasanah and L. Dewi, "Inklusivitas Teknologi untuk Tunanetra", Journal Access and Inclusion,

vol. 5, no. 1, pp. 30-45, 2023. doi: 10.1016/j.acin.2023.01.003.

- [9] N. Hasanah and M. R. Dewi, "Pemanfaatan Media Digital Al-Qur'an", Jurnal Komunitas Islam, vol. 11, no. 1, pp. 45-56, 2023.
- [10] M. Al-Rajhi, "Fuzzy Arabic Text Matching for Islamic Content", Journal of King Saud University-Computer Sciences, vol. 34, no. 1, pp. 100-112, 2022. doi: 10.1016/j.jksuci.2021.11.005
- [11] A. Widodo and B. Setiawan, "Audio-Based Quran Memorization", Jurnal Mobile Application, vol. 10, no. 4, pp. 55-70, 2023. doi: 10.1016/j.jma.2023.04.012.
- [12] K. Mohamed and M. Saeed, "Fuzzy Logic for Quranic Retrieval", Internation Journal Computer Application, vol. 183, no. 31, pp. 25-30, 2021. doi: 10.1016/j.ijca.2021.09.015.
- [13] K. Kiss, Fuse.js Documentation, 2020. [Online]. Available: <https://fusejs.io>
- [14] S. Russell, "Progressive Web Apps", IEEE Internet Computer, vol. 19, no. 6, pp. 78-85, 2015. doi: 10.1109/MIC.2015.124.
- [15] A. Alharbi and M. Jamaludin, "Hybrid Fuzzy-Audio Retrieval System," IEEE Access, vol. 9, pp. 134567-134582, 2021. doi: 10.1109/ACCESS.2021.3115432.
- [16] E. F. Mustafa and R. K. Ibrahim, "Optimizing Levenshtein for Arabic Script", Journal King Saud University Computer Information Science, vol. 34, no. 8, pp. 5897-5910, 2022. doi: 10.1016/j.jksuci.2021.11.012.