Evaluating the Effectiveness of a Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B

J.R.B. Soriano[1], M.R.G. Ono[2], M.L.V. Agbayani[3], A.C.G. Laguisma[4], A.L. Trajano[5], K.C.C. Angeles[6]

STI College Novaliches, College of Computer Science, Novaliches, Quezon City, Metro Manila, Philippines

DOI: https://doi.org/10.47772/IJRISS.2026.1026EDU0013

Received: 19 December 2025; Accepted: 24 December 2025; Published: 05 January 2026

ABSTRACT

This research aims to develop a web-based platform that integrates artificial intelligence to make Japanese language learning more accessible and interactive. Designed for travelers, overseas workers, and self-learners, Japanese Learning Management system with Artificial Intelligence Powered Tutor utilizing Llama-3.1-8B provides JLPT N5–N4 lessons with an AI tutor that offers pronunciation feedback and captions displaying Japanese scripts with *romaji* (the Romanized representation of Japanese words) and English translations. The system was developed using Flask, Alpine.js, Tailwind CSS, and MySQL, while the AI feature was trained with Hugging Face Transformers and PyTorch for voice recognition. To achieve these objectives, the system was developed using the Agile methodology following the Scrum framework, ensuring iterative design, testing, and integration of features. Evaluation results from 219 students, three IT professionals, and one Japanese language instructor showed high satisfaction, with overall ratings of 3.85, 4.61, and 5.00 respectively. These results confirm that the system effectively enhances pronunciation, comprehension, and learner engagement through AI integration. Future improvements include expanding datasets, training the AI on stronger hardware, refining usability, and adding higher JLPT levels and other languages to make the platform more versatile and scalable. This study presents the development and evaluation of a Japanese Language Learning Management System (LMS) integrated with an AI-powered tutor utilizing LLaMA 3.1 8B, designed to support learners preparing for JLPT N5–N4 levels. The system combines structured lessons with interactive features, including real-time pronunciation feedback, voice recognition, and captioned scripts displaying Japanese characters with romaji and English translations. Built using Flask, Alpine.js, Tailwind CSS, and MySQL, the AI component was fine-tuned with Hugging Face Transformers and PyTorch for speech processing. A quantitative descriptive design guided the evaluation, involving 219 students, three IT professionals, and one certified Japanese language instructor. Assessment instruments were based on ISO/IEC 25010 software quality standards and language learning objectives. Results indicate high user satisfaction: end users rated the system 3.85 (Agree), IT professionals scored 4.61 (Agree), and the instructor gave a perfect 5.00 (Strongly Agree). Findings confirm that the platform effectively enhances pronunciation, comprehension, and learner engagement while meeting international software quality benchmarks. Future improvements include expanding datasets, optimizing AI performance, and adding higher JLPT levels and multilingual support to ensure scalability and broader applicability.

Keywords: Japanese language, learning management system, AI tutor, LLaMA 3.1 8B, pronunciation feedback, ISO/IEC 25010.

INTRODUCTION

The Japanese language has become increasingly significant for global communication, cultural exchange, and professional advancement. In 2024, the Japanese-Language Proficiency Test (JLPT) recorded approximately 1.72 million applicants worldwide, marking its highest participation to date [1]. Despite this growing interest, many learners encounter barriers such as high tuition fees, rigid schedules, and limited opportunities for interactive speaking practice. Existing online platforms often prioritize vocabulary and grammar drills, offering minimal support for pronunciation and conversational fluency. Reports on AI-based language applications reveal mixed results, particularly for Japanese, where speech accuracy and natural dialogue remain challenging [2], [3].

Recent studies on AI integration in Learning Management Systems (LMS) highlight its potential to deliver personalized learning paths, adaptive assessments, and real-time feedback, provided that ethical and technical considerations are addressed [4]. Furthermore, research on voice-activated AI assistants indicates that learners tend to modify their speech for clarity during AI interactions, suggesting that conversational AI can effectively support pronunciation improvement [5]. These insights underscore the need for language learning systems that combine structured lessons with interactive, feedback-driven oral practice.

To address these gaps, this study introduces a Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B, a state-of-the-art large language model launched in July 2024 with extended context capabilities for sustained dialogue and instructional support [6], [7]. The system integrates voice recognition, real-time pronunciation feedback, and captioned scripts aligned with JLPT N5–N4 levels. By leveraging open-source LLM technology and LMS best practices, the platform aims to provide an affordable, flexible, and effective solution for beginner Japanese learners.

Objectives of the Study:

1. Identify gaps in AI-assisted Japanese language learning.

2. Describe the architecture and features of the proposed LMS.

3. Evaluate its effectiveness in improving pronunciation, comprehension, and learner engagement using ISO 25010-based quality metrics.

RELATED STUDIES

Foreign

Artificial Intelligence (AI) integration in Learning Management Systems (LMS) has been shown to improve personalization, adaptive assessments, and learning analytics when implemented under strong governance and ethical frameworks [8]. Studies report enhanced learner engagement and instructional efficiency, while emphasizing the need for privacy safeguards and faculty training [8]. Broader analyses of generative AI adoption in higher education highlight accelerated uptake, usefulness for study support, and the necessity to update assessment policies and lecturer competencies [9].

Research on voice-activated AI assistants demonstrates that second-language learners modify speech clarity during AI interactions, suggesting conversational AI can support pronunciation improvement [10]. These findings align with classroom observations where chatbots reduce anxiety and promote fluency, though recognition accuracy remains a challenge for non-English languages [10], [9].

Global platforms such as Duolingo have integrated AI features, but user feedback reveals mixed effectiveness for Japanese, particularly in pronunciation and speaking practice [11], [12]. This gap underscores the need for systems that prioritize oral skills and real-time feedback beyond gamified drills.

Meta's LLaMA 3.1 8B offers multilingual capabilities, instruction tuning, and a 128K-token context window, enabling sustained tutoring and context-rich feedback for language learning [13], [14]. These features provide a foundation for Japanese beginner-level oral practice when combined with voice recognition and structured prompts.

Local

The Department of Education launched E-CAIR in 2025 to integrate AI-driven tools for teaching, learning, and school administration. Initiatives include AI-assisted dashboards for leadership selection and geospatial mapping for resource allocation under a formal governance framework [15], [16], [17].

DOST developed "iTanong," a Filipino-style chatbot supporting Tagalog and Taglish, aimed at improving access to government services and expanding multilingual AI capabilities [18], [19].

The University of Southeastern Philippines deployed SePhi to enhance academic and administrative support within its LMS ecosystem, demonstrating localized AI adoption in higher education [20].

Synthesis

Foreign studies emphasize the potential of AI-driven LMS for personalization and adaptive learning, while conversational AI research validates its role in improving pronunciation and fluency. However, mainstream platforms show limitations in Japanese oral practice, creating a need for speech-focused solutions. Locally, initiatives such as DepEd's E-CAIR and DOST's chatbot projects demonstrate readiness for AI integration in education, while university-level deployments like SePhi confirm feasibility in institutional contexts.

- The proposed Japanese Language LMS with an AI-Powered Tutor Utilizing LLaMA 3.1 8B addresses these gaps by combining:

- Long-context multilingual dialogue for sustained tutoring [13], [14];
- Voice recognition and real-time pronunciation feedback aligned with JLPT N5–N4;
- Captioned scripts for linking oral and written comprehension;
- Compliance with local governance and ethical standards [15], [18].

This integrated approach aligns with global best practices and national initiatives, offering a practical, ethical, and scalable solution for Japanese language learning.

DESIGN AND METHODOLOGY

Research Design

This study adopted a quantitative descriptive research design to evaluate the effectiveness of the proposed Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B. The design was selected to provide an objective assessment of the system's performance and its impact on language learning outcomes through measurable indicators.

The research focused on three key dimensions: learning effectiveness, usability, and system quality. To achieve this, structured evaluation instruments were developed based on the ISO/IEC 25010 Software Quality Model and language learning objectives aligned with JLPT N5–N4 standards. The study involved three groups of respondents: 219 Tourism and Hospitality Management students enrolled in Japanese language courses, three IT professionals specializing in web and software development, and one certified Japanese language instructor who served as a subject matter expert.

The procedure began with the deployment of the LMS in a controlled environment. Participants were oriented on system navigation and features, including the AI-powered tutor and voice recognition functionality. Students engaged with the platform by accessing modules, practicing conversational Japanese through AI interactions, and completing assessments. IT professionals evaluated technical aspects such as functionality, reliability, and maintainability, while the language instructor assessed the educational validity of the content and AI feedback.

Data collection was conducted using Likert-scale questionnaires for end users and IT professionals, complemented by expert evaluation forms for the instructor. System performance tests were also carried out to measure voice recognition accuracy and AI feedback responsiveness. The data gathered were analyzed using descriptive statistics, including mean, standard deviation, and weighted mean, to interpret user perceptions and validate the system's effectiveness in improving pronunciation, comprehension, and learner engagement.

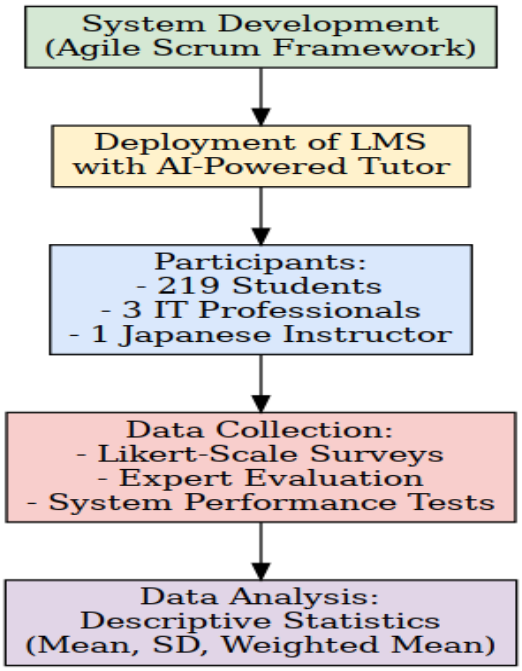


Figure 1. Research Design of the Study

Figure 1 illustrates the overall flow of the research design adopted in this study. It begins with System Development, where the Japanese Language Learning Management System with an AI-Powered Tutor utilizing LLaMA 3.1 8B was created using the Agile Scrum framework. After development, the system was deployed in a controlled environment for evaluation.

The next stage involves Participants, which include three distinct groups:

- 219 Tourism and Hospitality Management students enrolled in Japanese language courses,
- 3 IT professionals specializing in web and software development, and
- 1 certified Japanese language instructor serving as a subject matter expert.

Following participant engagement, the Data Collection phase was conducted using structured instruments such as Likert-scale questionnaires, expert evaluation forms, and system performance tests to measure voice recognition accuracy and AI feedback responsiveness.

Finally, the Data Analysis phase employed descriptive statistical methods including mean, standard deviation, and weighted mean to interpret user perceptions and validate the system's effectiveness in improving pronunciation, comprehension, and learner engagement.

A. Developmental Research Design

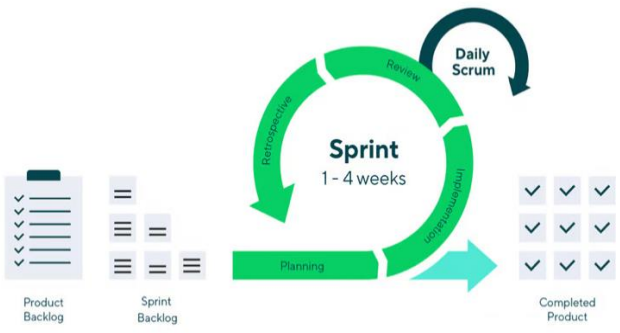


Figure 2. Agile Development Model

Figure 2 illustrates the Agile Scrum methodology applied in the developmental research design for creating the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B. The diagram shows a cyclical process that emphasizes iterative development and continuous improvement.

The process begins with the Product Backlog, which contains all the prioritized features and requirements for the system. From this backlog, a Sprint Backlog is created, representing the tasks to be completed within a specific sprint cycle. Each Sprint typically lasts between 1 to 4 weeks and includes four key phases: Planning, Implementation, Review, and Retrospective. During the sprint, the team conducts Daily Scrum meetings to monitor progress and address issues promptly.

After each sprint, a Completed Product Increment is delivered, which is a functional part of the system that meets the defined requirements. This iterative approach ensures that feedback is incorporated continuously, allowing for flexibility and refinement throughout the development process.

In the context of this study, the Agile Scrum framework facilitated the systematic development of the LMS, enabling the integration of AI features such as voice recognition and conversational tutoring while maintaining adaptability to user feedback and technical evaluations..

B. Conceptual Framework

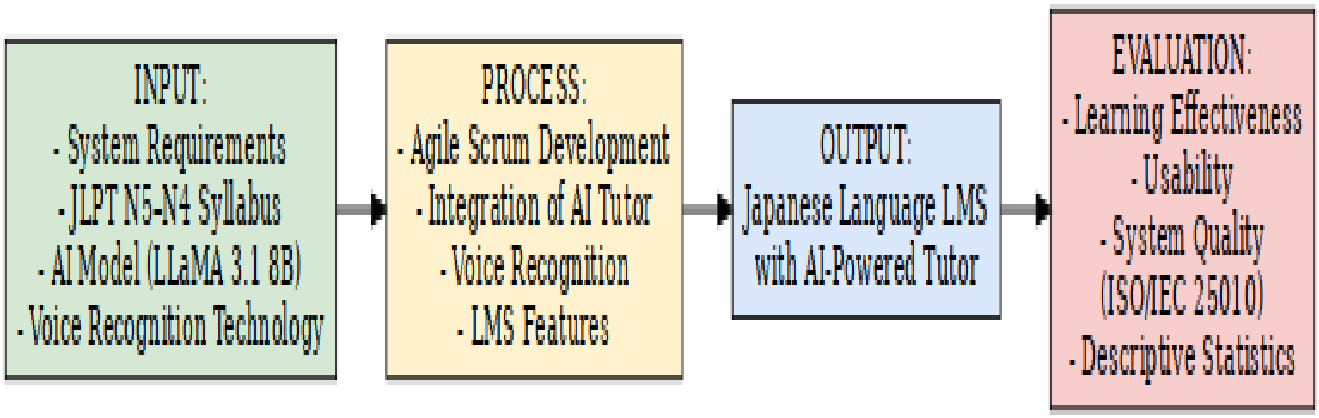


Figure 3. Conceptual Framework of the Study

The conceptual framework of this study is structured using the Input–Process–Output–Evaluation (I-P-O-E) model, which illustrates the logical flow from system design to assessment. The Input stage consists of the essential components required for development, including system requirements, the JLPT N5–N4 syllabus for language content, the AI model LLaMA 3.1 8B for conversational tutoring, and voice recognition technology to enable pronunciation feedback.

The Process stage involves the development of the Japanese Language Learning Management System using the Agile Scrum methodology. This iterative approach ensures continuous improvement and integration of key features such as the AI-powered tutor, voice recognition, and LMS functionalities like modules, assessments, and progress tracking.

The Output of this process is the completed Japanese Language LMS with an AI-Powered Tutor, designed to provide interactive learning experiences, real-time feedback, and structured lessons aligned with JLPT standards.

Finally, the Evaluation stage measures the system's effectiveness through learning outcomes, usability, and system quality based on ISO/IEC 25010 standards. Data analysis employs descriptive statistics, including mean, standard deviation, and weighted mean, to interpret user feedback and validate the system's impact on pronunciation, comprehension, and learner engagement.

This framework ensures that technological inputs and development processes are aligned with pedagogical objectives, resulting in a structured evaluation of the system's performance and educational value.

C. Artificial Intelligence Training

The Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B employs a fully independent, in-house AI model development approach tailored for Japanese language processing. The system uses open-source tools and libraries but does not rely on third-party APIs during development. Hugging Face resources were consulted only as architectural references to understand transformer-based models like GPT. The system leverages the Hugging Face Transformers library in an offline, self-contained manner, enabling developers to define, implement, and train models locally.

Custom datasets were created using Japanese language reference books aligned with JLPT N5 standards. These datasets guided the generation of example sentences and grammar lessons suitable for beginner-level proficiency. The AI was trained to provide accurate responses, evaluate vocabulary and grammar usage, and deliver meaningful feedback to learners.

The training process involved:

- Data Collection – Loading custom datasets specifically designed for Japanese language learning.
- Data Preprocessing – Formatting samples into structured instruction prompts.
- Parameter Configuration – Setting training parameters such as token length, batch size, and epochs.
- Model Training – Fine-tuning the GPT-based model using Hugging Face's Trainer class.
- Dynamic Adjustment – Applying mixed precision and gradient accumulation for resource optimization.
- Forecast Generation – Enabling the model to predict suitable words or sentences for learner queries.

The PyTorch framework was used to connect datasets with the model, while Flask served as the backend for API communication between the AI and the LMS interface. This design ensures transparency, data security, and performance optimized for the system's educational objectives, delivering features such as pronunciation assistance, sentence suggestions, and interactive responses.

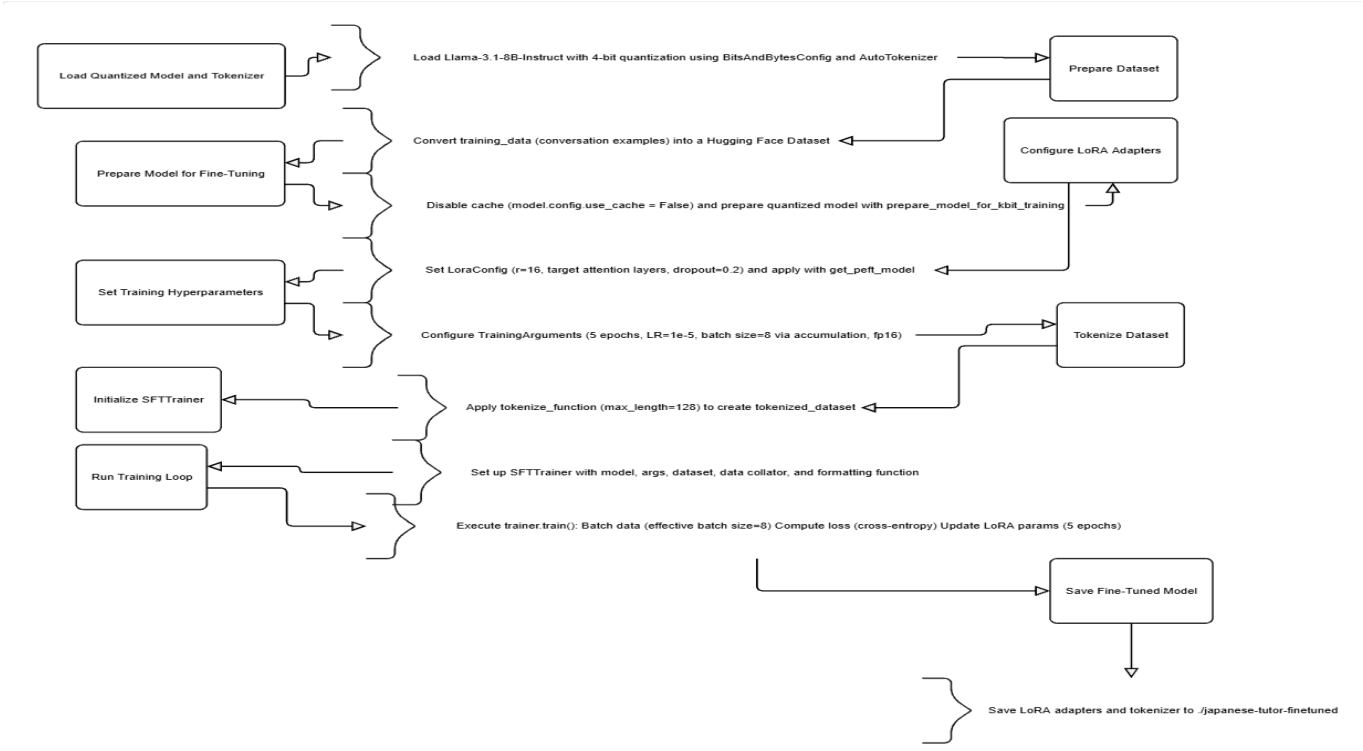


Figure 4. Artificial Intelligence Training

D. System Flow

Figure 4 below presents the learner flow within the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B. The process begins when a learner registers by submitting their name, email, and password. After registration, the system sends a One-Time Password (OTP) to the learner's email for verification. The learner enters the OTP to validate their account. If the OTP is invalid, the system prompts the learner to resend the OTP and retry verification. Once the OTP is successfully validated, the learner proceeds to the placement assessment stage.

At this point, the system asks whether the learner will take the post-registration assessment. If the learner chooses to skip the assessment, they are directed to the dashboard with default learning modules. If the learner opts to take the assessment, they begin and complete the evaluation. Based on the assessment results, the system automatically suggests specific learning modules tailored to the learner's proficiency level. Finally, the learner gains access to the dashboard, where personalized lessons and activities are available for continued learning.

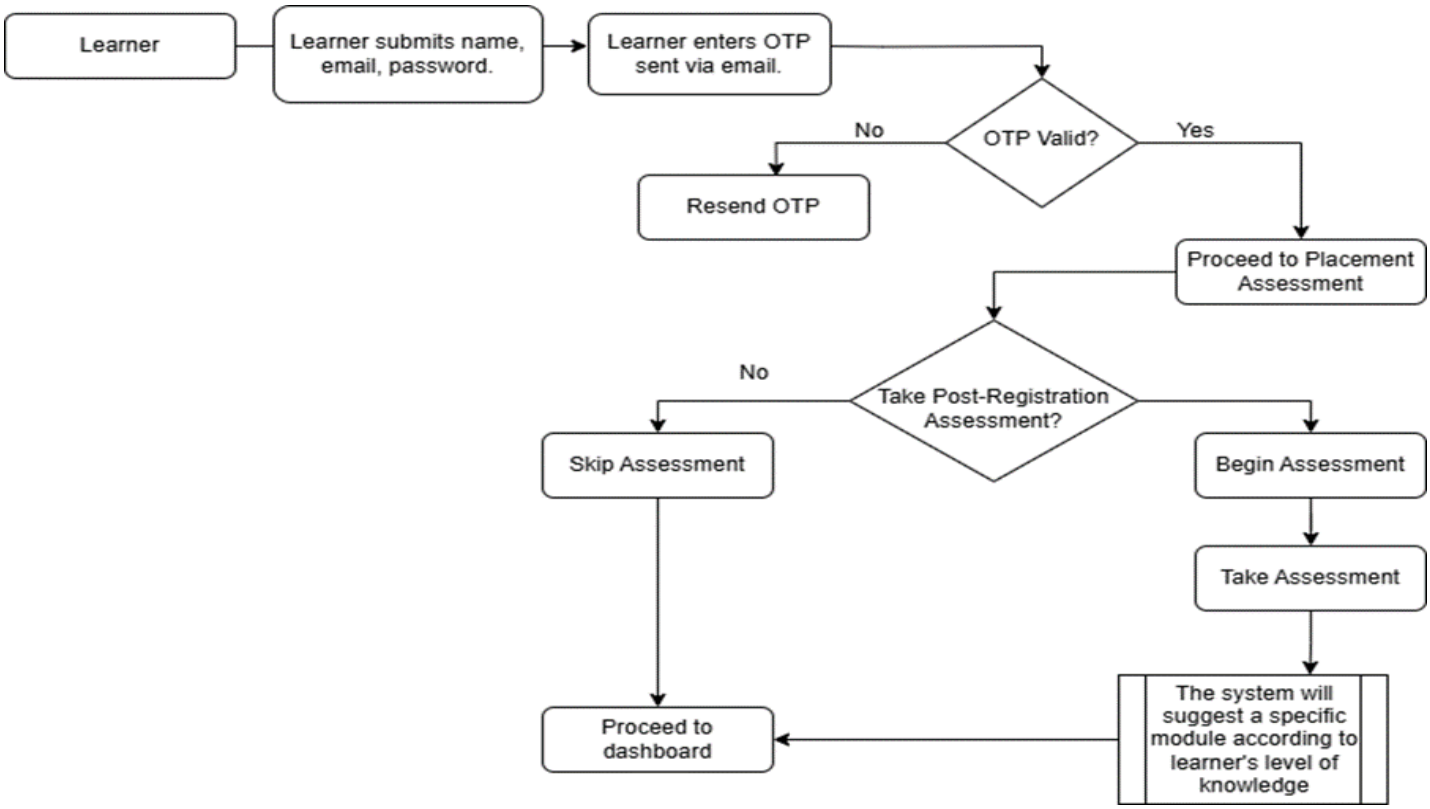


Figure 5. Learner Flow Diagram

Figure 5 illustrates the module learning flow within the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B. The process begins when the learner accesses the lesson content for the selected module. After reviewing the lesson, the learner proceeds to practice using voice recognition, which allows them to improve pronunciation and speaking skills interactively.

Once practice is completed, the system checks whether the learner has finished the module. If the module is not yet complete, the learner is directed back to the lesson content for further study and practice. If the module is completed, the learner moves forward to the module quiz, where they take an assessment to evaluate their understanding of the lesson.

After completing the quiz, the system displays quiz results and feedback, helping the learner identify strengths and areas for improvement. The learner's progress is then updated in the dashboard, ensuring accurate tracking of completed modules. Finally, the learner can either end the session or continue with the next module, maintaining a structured and personalized learning experience.

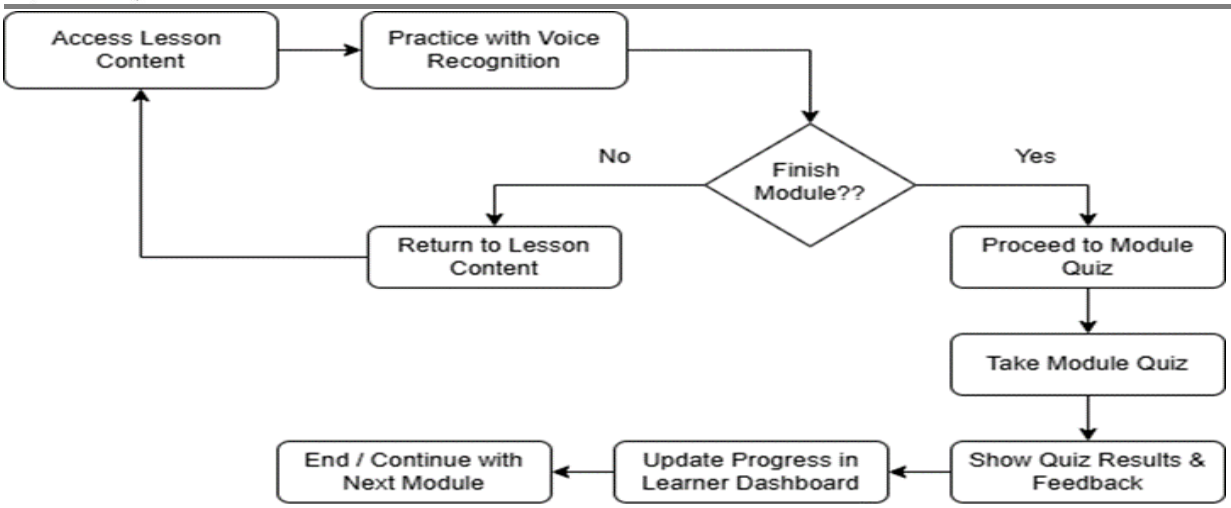


Figure 6. Learning Flow Diagram

Figure 6 illustrates the teacher account workflow within the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B. The process begins when the teacher logs into the system using credentials provided by the administrator. These default credentials include a username and password assigned during account creation. Upon first login, the system enforces a mandatory password change to ensure account security.

After updating the password, the teacher is redirected to the Teacher Dashboard, which serves as the central hub for managing instructional activities. From the dashboard, the teacher can create and manage lesson modules and assessments, tailoring content to meet JLPT N5 standards. Additionally, the teacher can view learner progress and assessment results, enabling them to monitor performance and provide targeted support. This workflow ensures secure access and efficient management of learning resources and learner data.

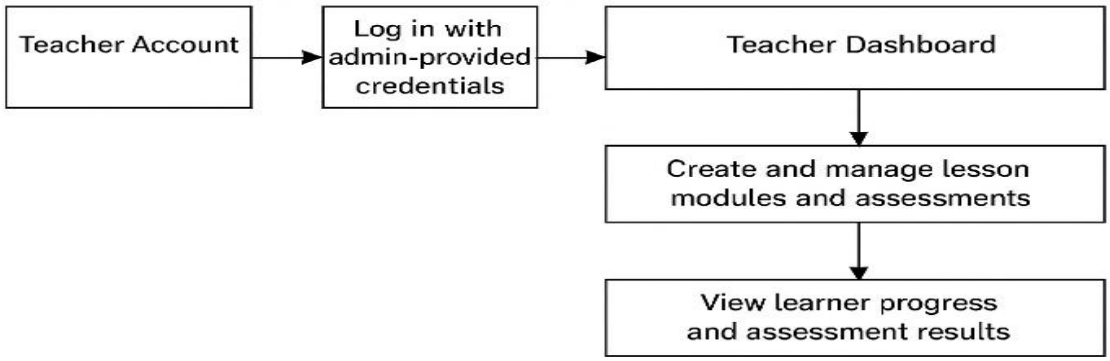


Figure 7. Teacher Account Workflow Diagram

Figure 7 illustrates the administrative functions within the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B. The process begins when the administrator accesses the Admin Dashboard, which serves as the central control panel for managing system operations. From this dashboard, the administrator performs several key tasks:

- User Management – Oversees teacher and learner accounts, ensuring proper access and security.
- Manage Course – Creates, updates, or removes courses offered within the LMS.
- Manage Contents – Organizes lesson materials, multimedia resources, and assessments to maintain instructional quality.

- Monitor Audit Logs – Reviews system activity logs to track changes, ensure compliance, and maintain security.

This structured workflow ensures that the administrator maintains full control over system configuration, content management, and user activities, supporting the smooth and secure operation of the LMS.

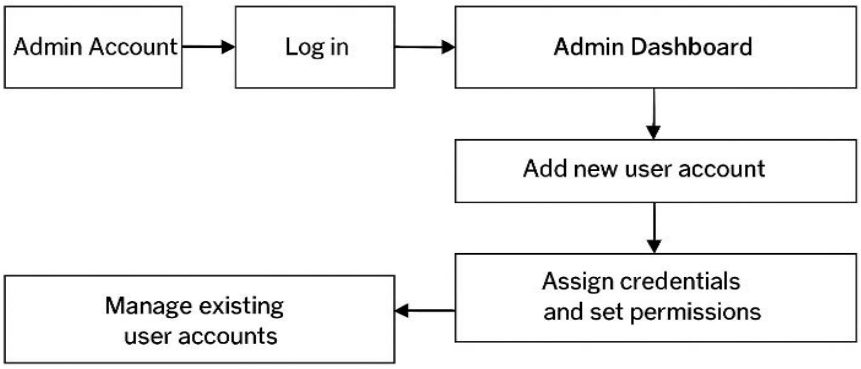


Figure 8. Admin Account Workflow Diagram

E. System Interface

The developed system provides an intuitive and user-friendly interface designed to support learners in navigating lessons and interacting with the AI tutor. Screenshots of key components are presented below to illustrate the system's functionality.

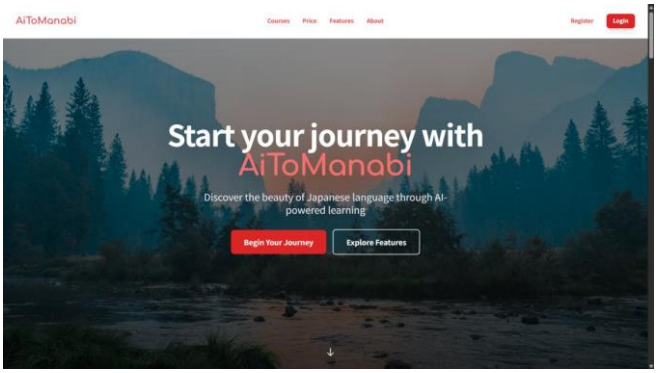


Figure 9. Landing Page

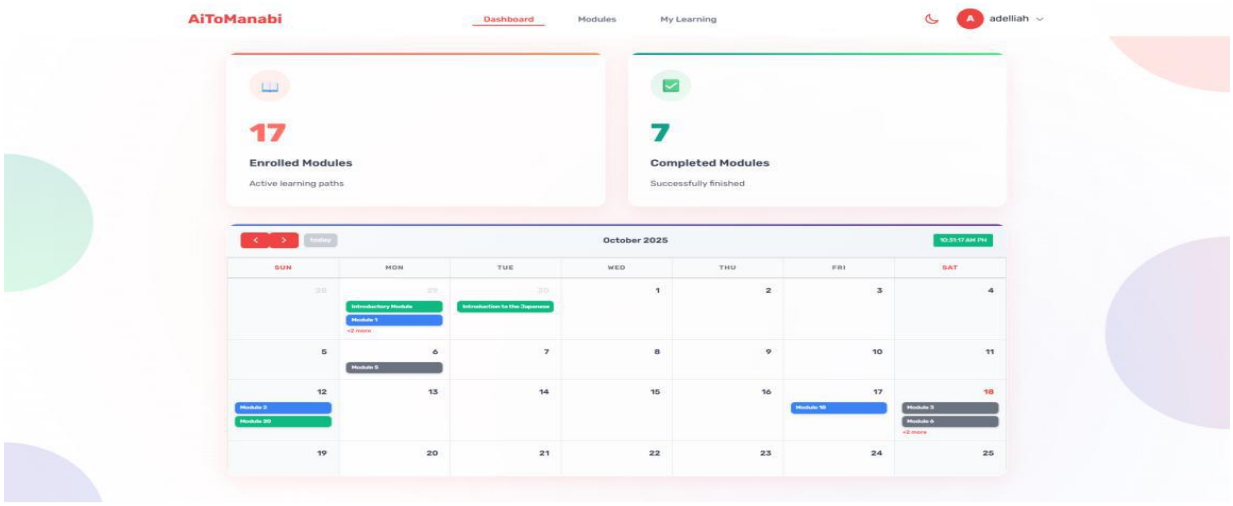


Figure 10. Learner/Student Dashboard

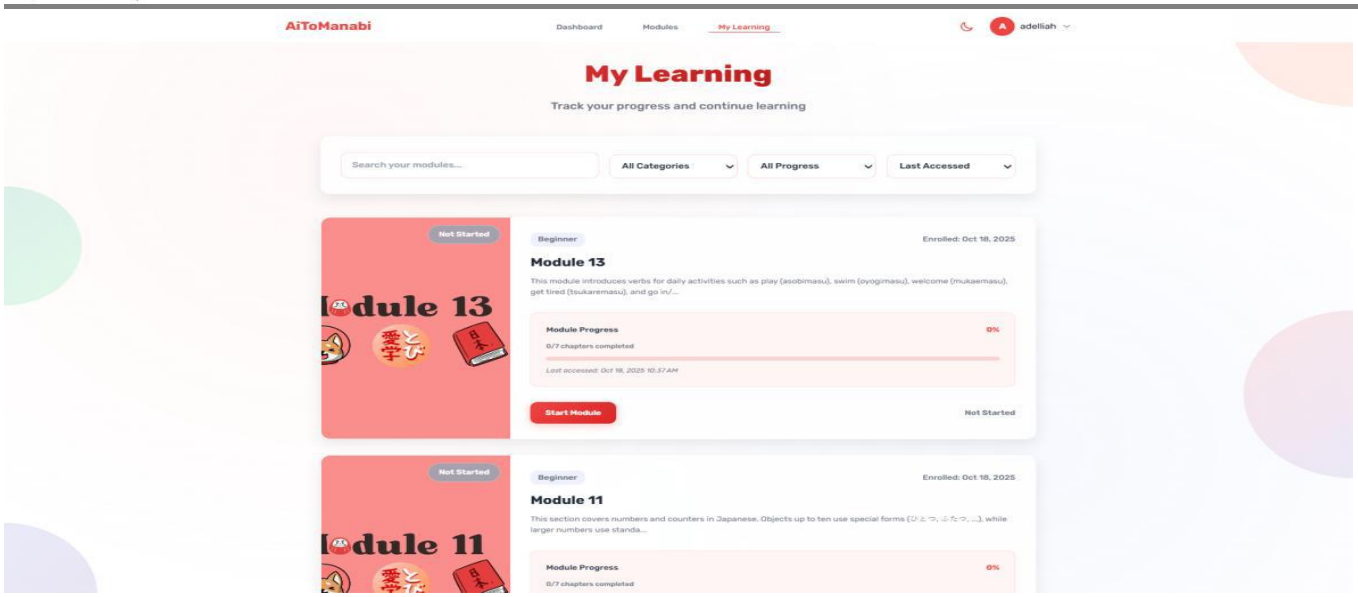


Figure 11. My Learning Page (enrolled modules)

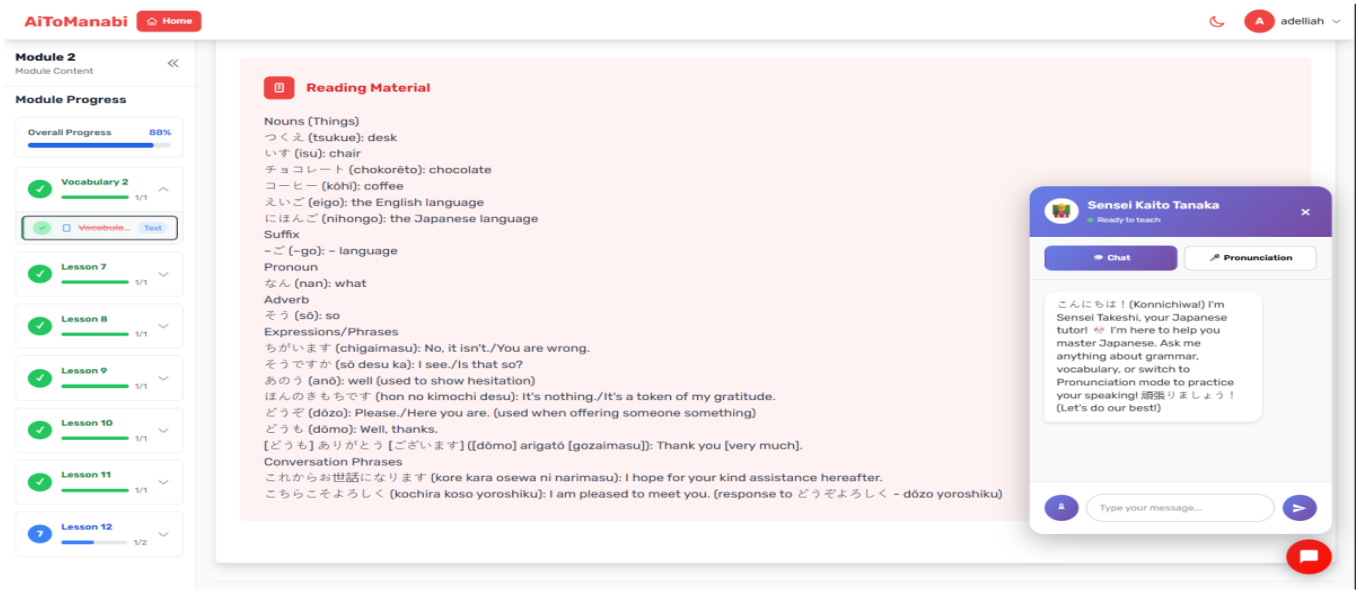


Figure 12. Lesson Content with AI Tutor

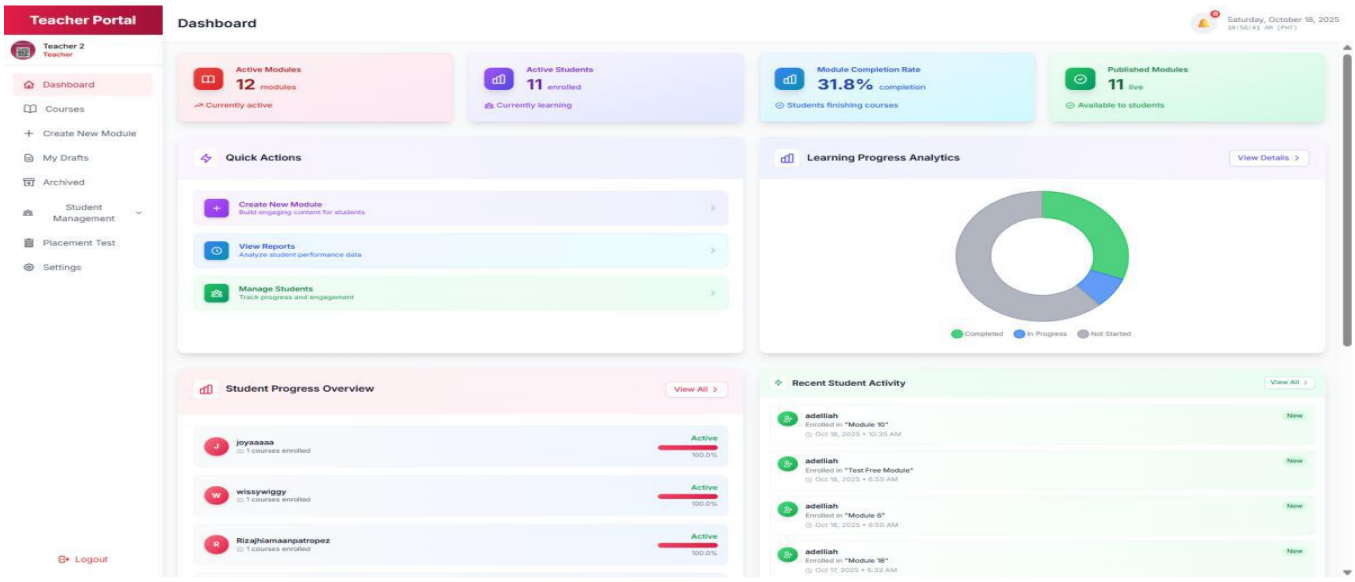


Figure 13. Teacher Dashboard

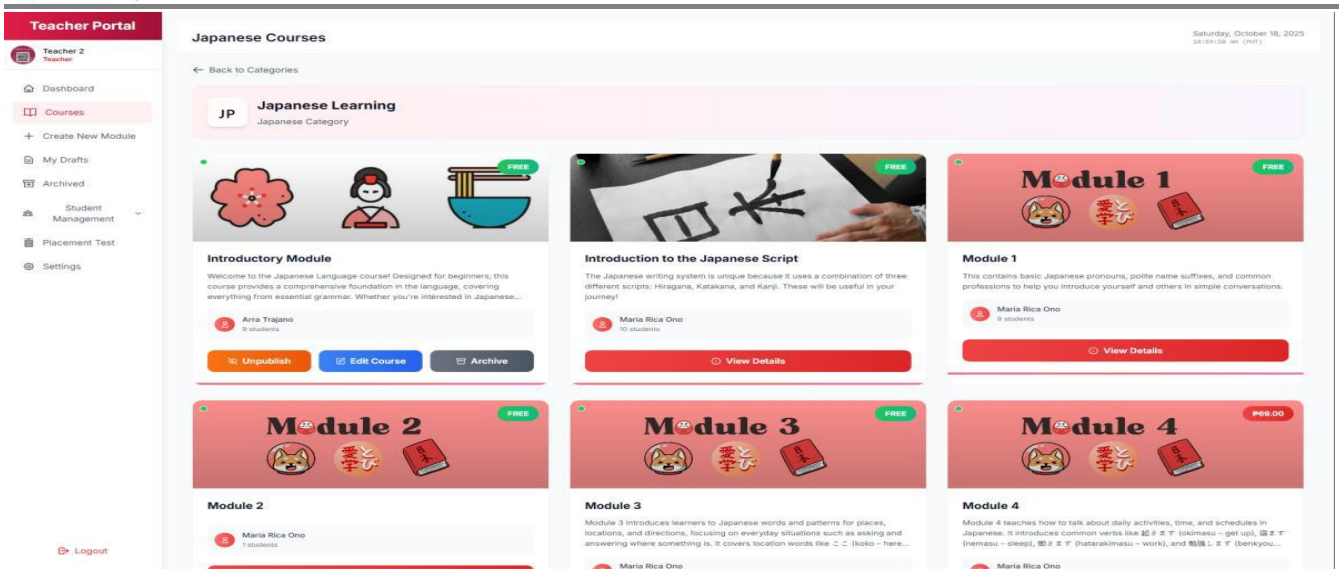


Figure 14. Modules Page

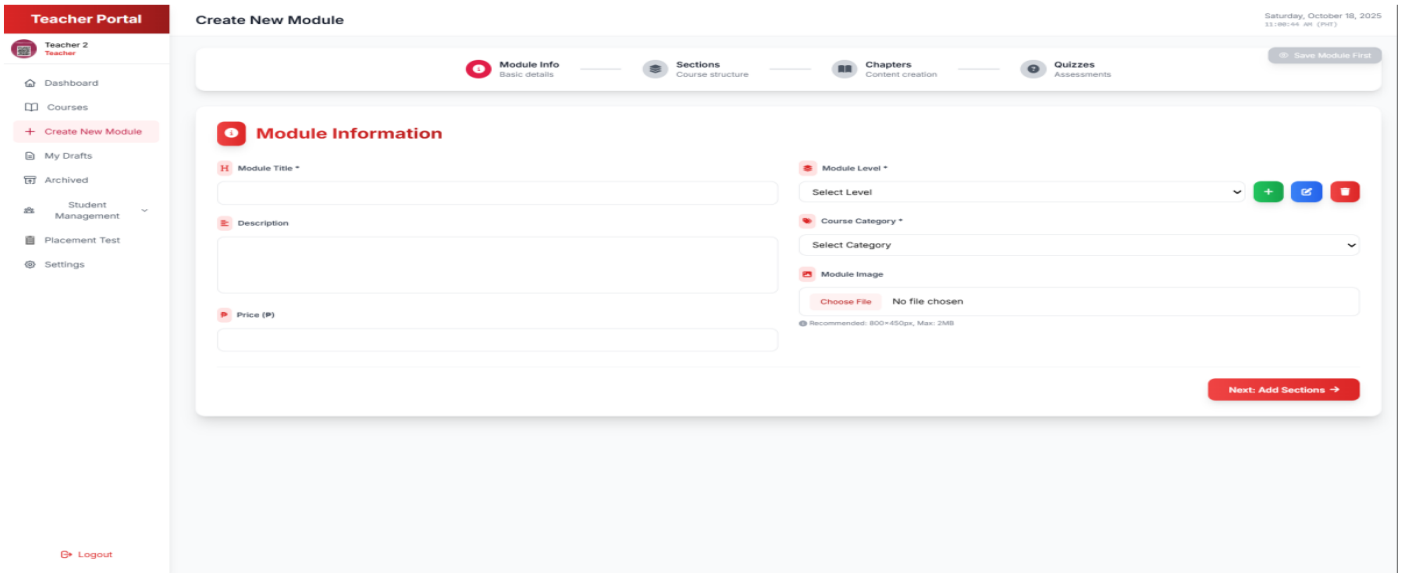


Figure 15. Module Creation

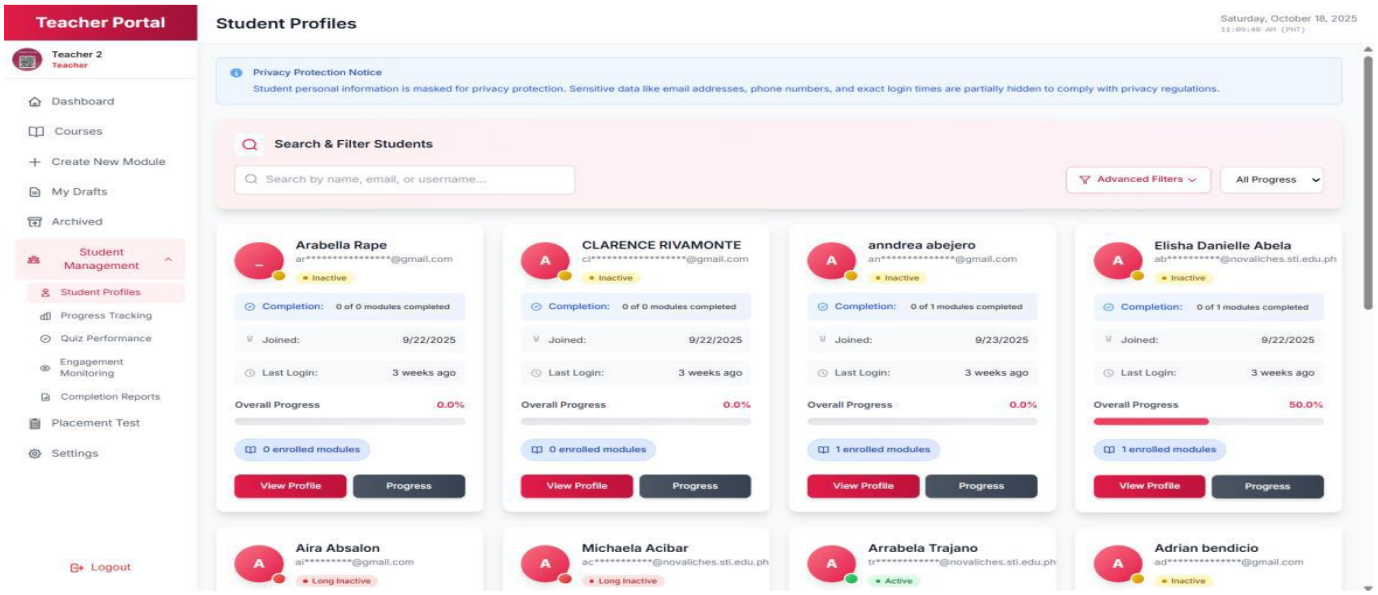


Figure 16. Students Profile

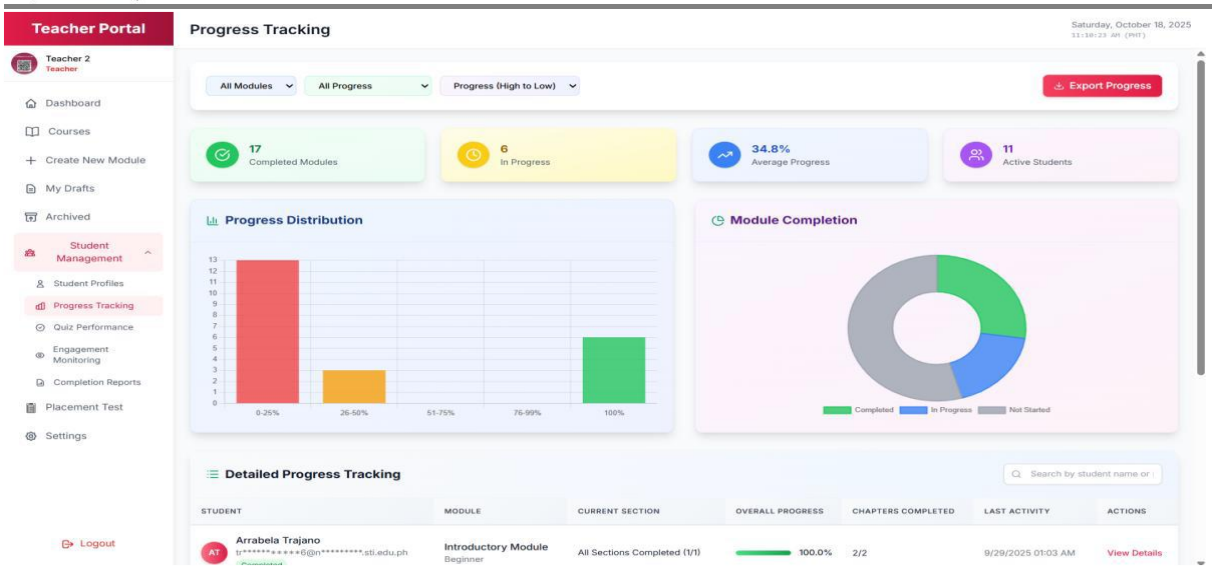


Figure 17. Progress Tracking

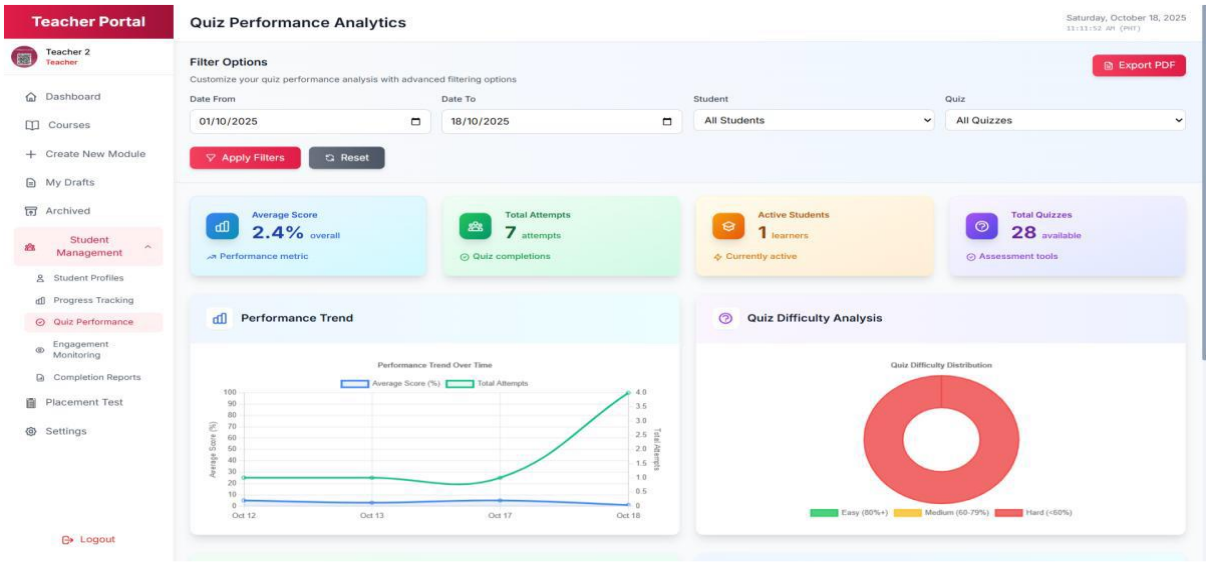


Figure 18. Quiz Performance Analytics

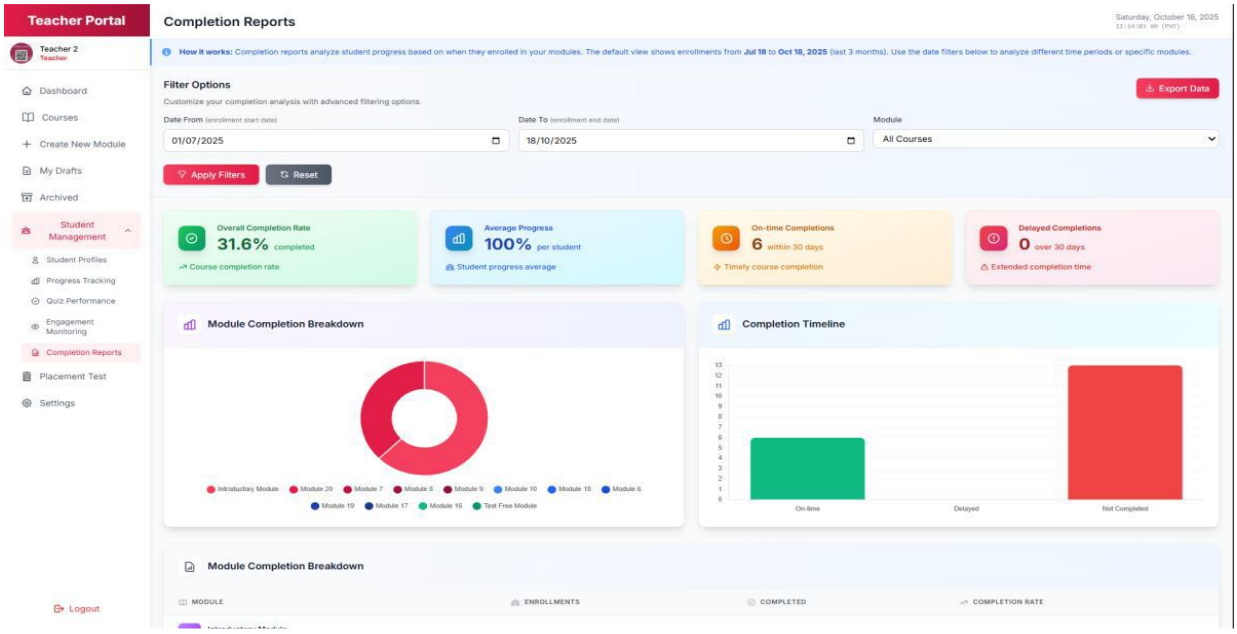


Figure 19. Module Completion Report

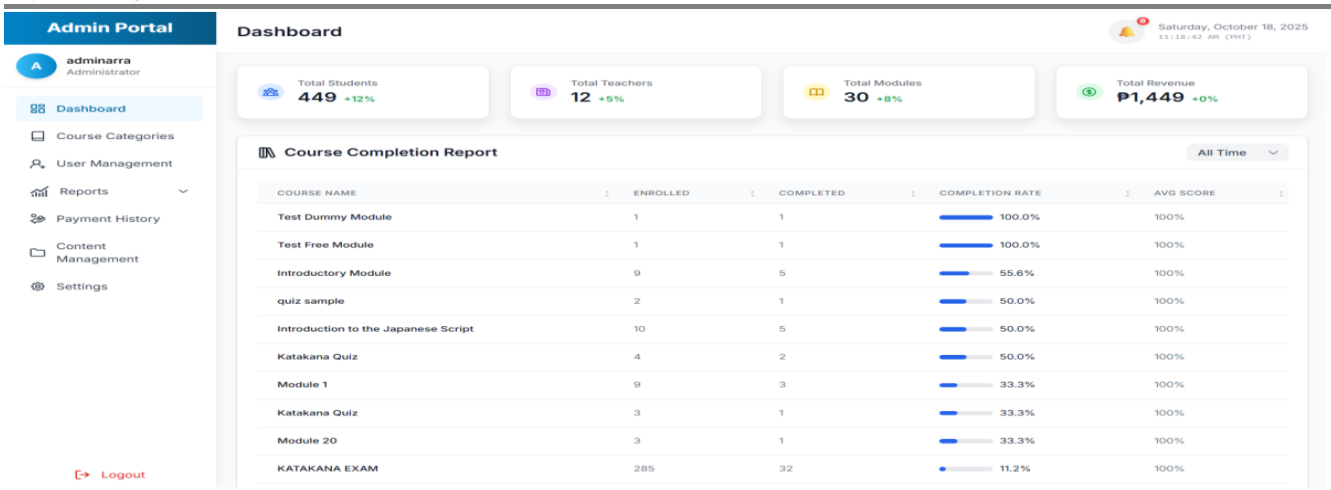


Figure 20. Admin Dashboard

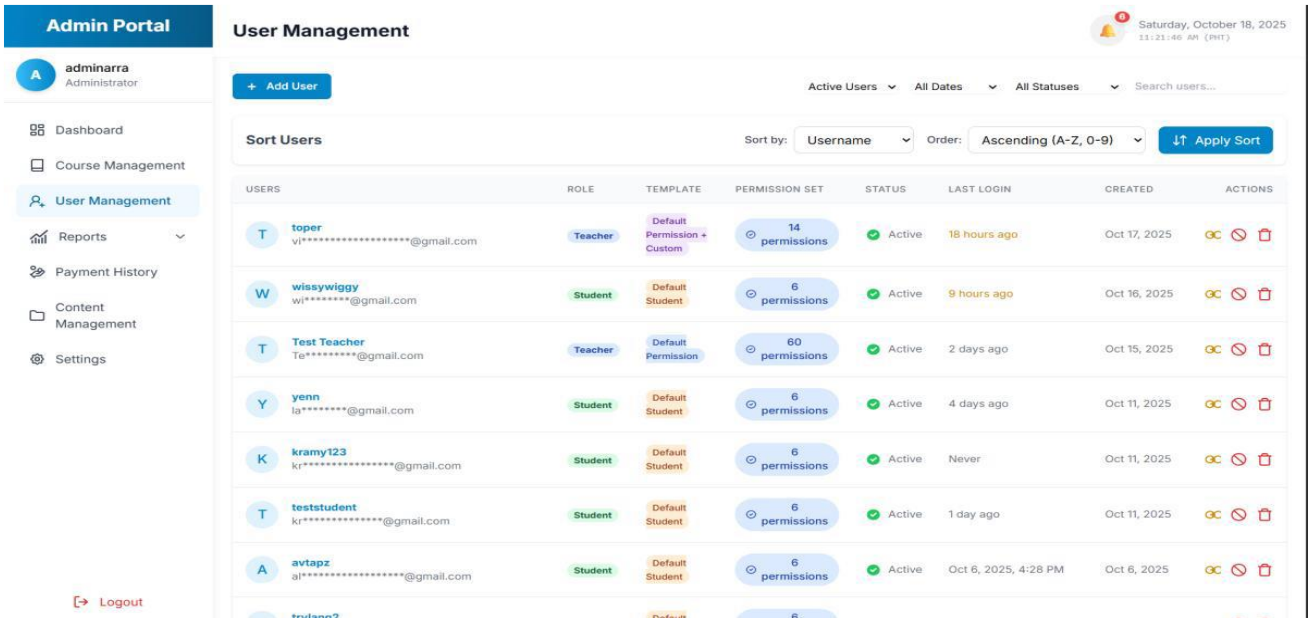


Figure 21. User Management

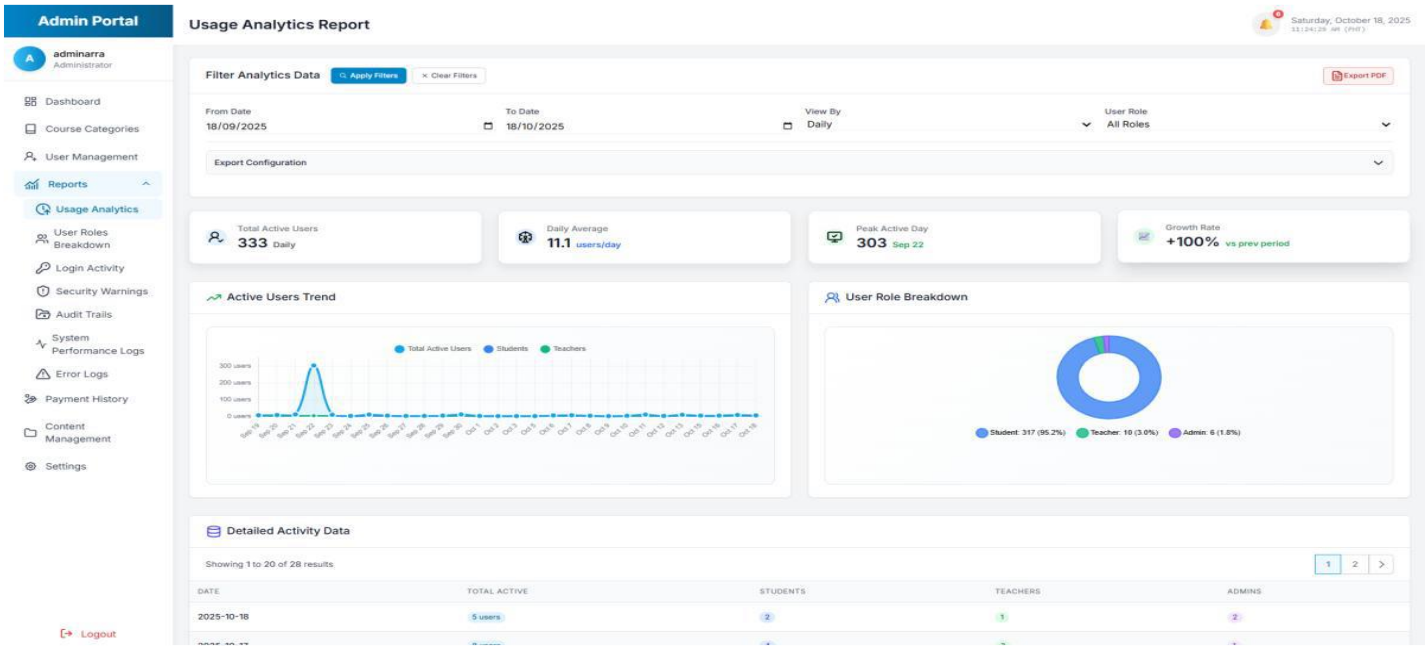


Figure 22. Usage Analytics

RESULT AND DISCUSSIONS

Evaluation Results of the Developed System

This presents the findings from the evaluation of the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B, analyzed according to the study's objectives. The evaluation involved three distinct respondent groups, each using a customized survey instrument aligned with their roles and expertise:

End Users: A total of 219 Tourism and Hospitality Management students enrolled in the Japanese language course at STI College Novaliches. Their evaluation focused on Learning Effectiveness, User Interface, Security, and System Quality, assessing how the system supports comprehension, engagement, accessibility, and overall learning experience.

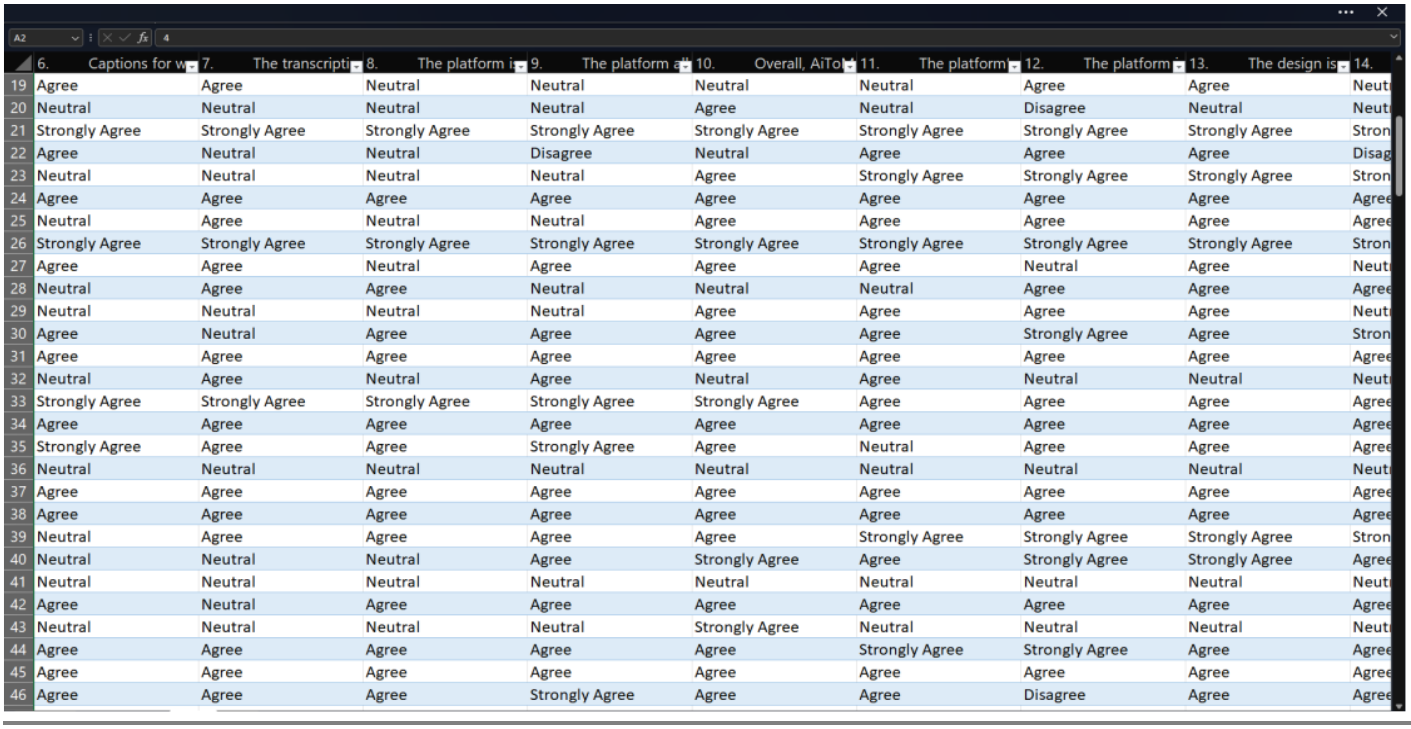
IT Professionals: Three specialists in web and software development assessed the system using a survey instrument aligned with the ISO/IEC 25010 Software Product Quality Model, covering Functionality, Performance, Reliability, and Maintainability.

Japanese Language Instructor: One professional educator evaluated the system's alignment with Learning Objectives, AI-based instruction, and its potential to improve pronunciation and comprehension, particularly for beginner-level learners.

Data were analyzed using descriptive statistical methods, including computation of mean, standard deviation, and weighted mean, interpreted using a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree).

5-point Likert Scale	
Mean Range	**Verbal Interpretation**
1-1.79	Strongly Disagree
1.8-2.59	Disagree
2.6-3.39	Neutral
3.4-4.19	Agree
4.2-5	Strongly Agree
Interval	
4/5=0.8	

Figure 23. 5-point Likert Scale



	6. Captions for w	7. The transcripti	8. The platform i	9. The platform a	10. Overall, AiTo	11. The platform'	12. The platform	13. The design is	14.
19	Agree	Agree	Neutral	Neutral	Neutral	Neutral	Agree	Agree	Neut
20	Neutral	Neutral	Neutral	Neutral	Agree	Neutral	Disagree	Neutral	Neut
21	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Stron
22	Agree	Neutral	Neutral	Disagree	Neutral	Agree	Agree	Agree	Disag
23	Neutral	Neutral	Neutral	Neutral	Agree	Strongly Agree	Strongly Agree	Strongly Agree	Stron
24	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agre
25	Neutral	Agree	Neutral	Neutral	Agree	Agree	Agree	Agree	Agre
26	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Stron
27	Agree	Agree	Neutral	Agree	Agree	Agree	Neutral	Agree	Neut
28	Neutral	Agree	Agree	Neutral	Neutral	Neutral	Agree	Agree	Agre
29	Neutral	Neutral	Neutral	Neutral	Agree	Agree	Agree	Agree	Neut
30	Agree	Neutral	Agree	Agree	Agree	Agree	Strongly Agree	Agree	Stron
31	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agre
32	Neutral	Agree	Neutral	Agree	Neutral	Agree	Neutral	Neutral	Neut
33	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Agree	Agree	Agree	Agre
34	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agre
35	Strongly Agree	Agree	Agree	Strongly Agree	Agree	Neutral	Agree	Agree	Agre
36	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neut
37	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agre
38	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agre
39	Neutral	Agree	Agree	Agree	Agree	Strongly Agree	Strongly Agree	Strongly Agree	Stron
40	Neutral	Neutral	Neutral	Agree	Strongly Agree	Agree	Strongly Agree	Strongly Agree	Agre
41	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neut
42	Agree	Neutral	Agree	Agree	Agree	Agree	Agree	Agree	Agre
43	Neutral	Neutral	Neutral	Neutral	Strongly Agree	Neutral	Neutral	Neutral	Neut
44	Agree	Agree	Agree	Agree	Agree	Strongly Agree	Strongly Agree	Agree	Agre
45	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agre
46	Agree	Agree	Agree	Strongly Agree	Agree	Agree	Disagree	Agree	Agre

	14. I can easily fi	15. I feel confide	16. The system c	17. The platform	18. The platform	19. The platform	20. The platform	21. The platform	22. W
19	Neutral	Neutral	Neutral	Agree	Neutral	Neutral	Neutral	Neutral	very
20	Neutral	Agree	Agree	Neutral	Neutral	Neutral	Neutral	Neutral	game
21	Strongly Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	none
22	Disagree	Neutral	Disagree	Neutral	Neutral	Neutral	Neutral	Neutral	The c
23	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	None
24	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	i gain
25	Agree	Neutral	Agree	Neutral	Agree	Agree	Neutral	Agree	I can
26	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Noth
27	Neutral	Agree	Agree	Agree	Agree	Neutral	Agree	Agree	Beggi
28	Agree	Agree	Neutral	Neutral	Agree	Agree	Agree	Agree	I imp
29	Neutral	Neutral	Neutral	Neutral	Agree	Agree	Agree	Agree	None
30	Strongly Agree	Neutral	Agree	Agree	Agree	Strongly Agree	Agree	Agree	Enab
31	Agree	Neutral	Neutral	Neutral	Agree	Agree	Agree	Agree	My sl
32	Neutral	Agree	Agree	Agree	Agree	Neutral	Neutral	Agree	the c
33	Agree	Agree	Agree	Strongly Agree	Strongly Agree	Neutral	Agree	Agree	Base
34	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	none
35	Agree	Agree	Agree	Agree	Neutral	Neutral	Neutral	Agree	I wou
36	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	n/a
37	Agree	Agree	Agree	Agree	Agree	Neutral	Agree	Agree	Noth
38	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	More
39	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	Strongly Agree	None
40	Agree	Neutral	Neutral	Agree	Agree	Strongly Agree	Strongly Agree	Strongly Agree	none
41	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	Neutral	None
42	Agree	Neutral	Agree	Agree	Neutral	Neutral	Neutral	Agree	None
43	Neutral	Neutral	Neutral	Neutral	Agree	Neutral	Neutral	Neutral	None
44	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Navig
45	Agree	Agree	Agree	Agree	Agree	Agree	Agree	Agree	none
46	Agree	Neutral	Neutral	Agree	Disagree	Agree	Neutral	Agree	Easy

Figure 24. Survey Raw Data

Figure 23 shows the 5-point Likert scale legend used for interpretation, while Figure 24 presents the raw data collected from the system evaluation surveys. The tables display individual responses recorded across multiple evaluation criteria, including captions, transcript accuracy, platform usability, design quality, security, and overall satisfaction. Each column represents a specific statement, and rows correspond to participant responses, showing varying levels of agreement.

The dataset served as the basis for computing descriptive statistics to assess the system's effectiveness in terms of learning experience, technical performance, and instructional alignment. These results provide insights into the strengths of the system and identify areas for improvement to enhance usability and learning outcomes.

System Testing

The evaluation of the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B was conducted under the supervision of a Japanese Language Instructor at STI College Novaliches. The testing process involved three phases:

Expert Review:

The Japanese Language Instructor assessed the system's alignment with JLPT N5 learning objectives, focusing on pronunciation, comprehension, and overall educational potential.

Three IT professionals specializing in web and software development evaluated the system based on functionality, interface design, performance, and security, using ISO/IEC 25010 standards.

End-User Testing:

A total of 219 Tourism and Hospitality Management students enrolled in Japanese language courses participated in the evaluation.

Students were introduced to the system, guided through account registration, navigation, and feature usage, including the AI-powered tutor.

Data Collection and Analysis:

After using the system for lessons and assessments, participants completed evaluation forms measuring Learning Experience, Usability and Interface, Security and Privacy, and System Quality and Performance.

Responses were analyzed using descriptive statistics, including mean, standard deviation, and weighted mean, interpreted through a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree).

Implementation Plan

The implementation followed a structured approach:

1. Preparation Stage: System presented to experts for approval and planning.
2. Orientation: Students briefed on system usage and testing procedures.
3. Testing Phase: Students used the system for lessons and quizzes, including AI features.
4. Data Gathering: Evaluation forms collected post-usage.
5. Analysis: Statistical interpretation applied to measure effectiveness and identify areas for improvement.

User Evaluation and Statistical Analysis

The system was designed to provide an accessible Japanese language learning experience for diverse learners, including travelers, overseas workers, and self-learners who may have limited access to formal classes. The user interface was developed to be intuitive and easy to navigate, allowing learners to progress through lessons and activities without technical difficulty. The system follows a modular structure, starting from basic lessons to vocabulary, grammar, and AI tutor guidance, supporting learners aiming for JLPT N5–N4 levels. Progress tracking and structured lessons help maintain learner motivation and focus.

End users evaluated the system across four areas: Learning Experience, Usability and Interface, Security and Privacy, and Quality and Performance. Using a 5-point Likert scale, the overall results showed a positive response to the system. Mean and standard deviations were computed to interpret the data. The following presents the findings for each category.

Learning Experience	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
1. The lessons in AiToManabi improved my understanding of the Japanese language.	64	119	30	2	4	4.08	3.63	Agree	1
2. The system's progress tracking tools (e.g., assessments, progress bars) are helpful in monitoring my learning.	48	126	41	1	3	3.98	3.52	Agree	4
3. The AI tutor provides interactive conversations that help me practice conversational Japanese.	49	125	42		3	3.99	3.53	Agree	3
4. Speaking practice with the AI tutor improves my confidence in using the Japanese language.	41	115	58	2	3	3.86	3.41	Agree	9
5. The AiToManabi AI tutor's feedback has helped me correct my mistakes in speaking the Japanese language.	44	117	51	2	5	3.88	3.44	Agree	7
6. Captions for words and phrases help me understand spoken Japanese better.	46	123	45	2	3	3.95	3.49	Agree	5
7. The transcriptions provided help me follow conversations more easily.	43	126	45	2	3	3.93	3.48	Agree	6
8. The platform is accessible even without formal language course enrollment.	42	121	44	8	4	3.86	3.43	Agree	9
9. The platform allows me to study flexibly, despite having a busy schedule.	45	113	54	3	4	3.88	3.44	Agree	8
10. Overall, AiToManabi has contributed to my progress in learning Japanese.	61	116	38		4	4.05	3.60	Agree	2
Overall Weighted Mean						3.94	3.50	Agree	

Figure 25. Learning Experience

The respondents agreed on the learning experience of the system with an overall weighted mean of 3.94 (Agree). This indicates that the system successfully provides a Japanese language learning environment that supports and guides learners effectively.

The system integrates Artificial Intelligence (AI) and voice recognition to provide real-time, interactive language practice. The AI acts as a virtual tutor, enabling learners to engage in short conversations, practice pronunciation, and receive instant feedback.

The AI system was developed through dataset preparation, model training, and evaluation. Japanese language data were gathered from open-source and custom materials, including words, grammar patterns, and everyday phrases. To enhance pronunciation assessment, phoneme analysis was implemented, allowing the system to compare learner speech with native Japanese phonemes and provide precise feedback.

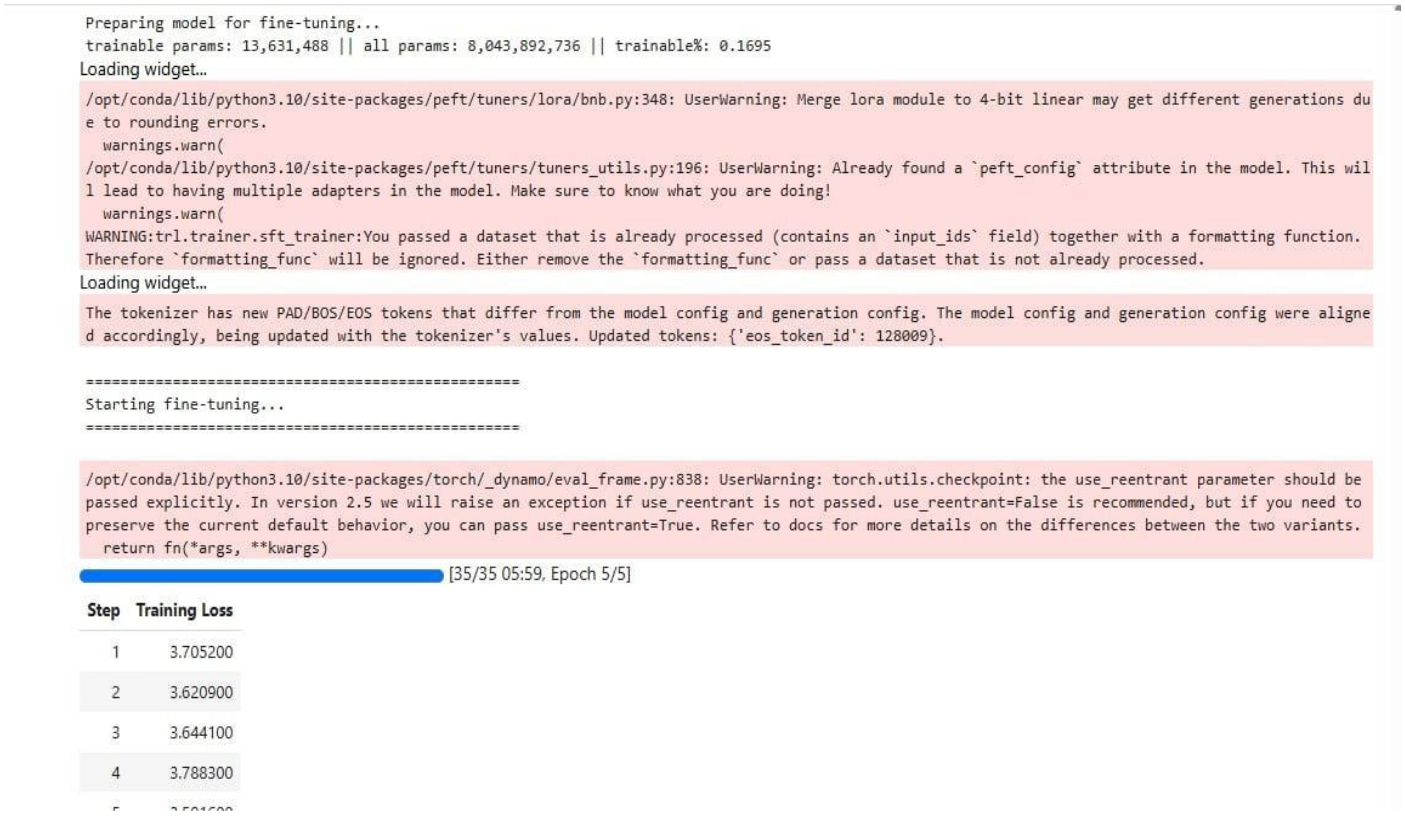


Figure 26. AI Training Process

Figure 27. Server Implementation

The machine learning model specializing in speech recognition and pronunciation analysis was trained using fine-tuned hyperparameters (epoch count, batch size, learning rate) and GPU acceleration for efficient deep learning computations. Evaluation metrics included accuracy, precision, recall, and F1 score.

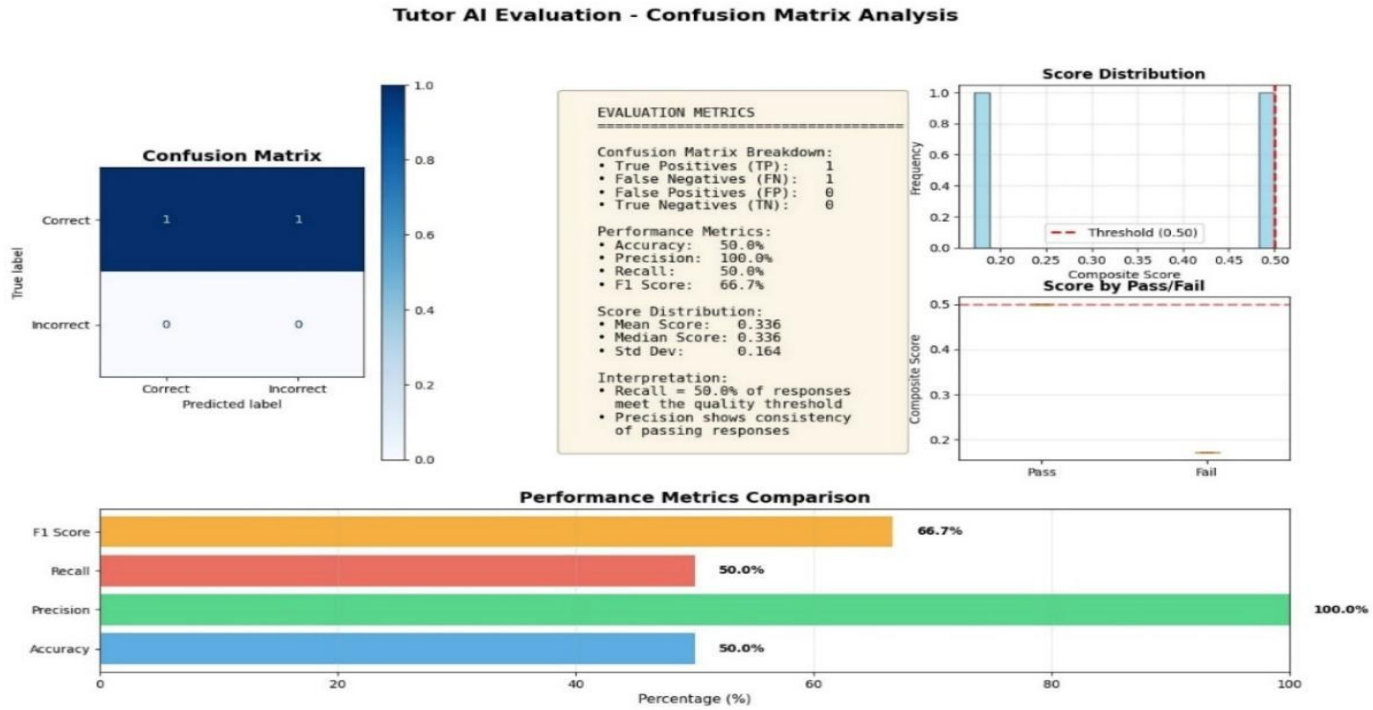


Figure 28. Confusion Matrix Analysis

Note: The confusion matrix analysis indicates an overall accuracy of 50%, precision rate of 100%, recall rate of 50%, and F1 score of 66.7%, showing moderate performance with room for improvement in detecting all correct responses.

Survey results under the Learning Experience category showed that speaking practice with the AI Tutor received a positive evaluation, with an overall weighted mean of 3.86 (Agree). Respondents confirmed that the AI feature enhanced their learning experience by providing immediate pronunciation feedback and opportunities for conversational practice.

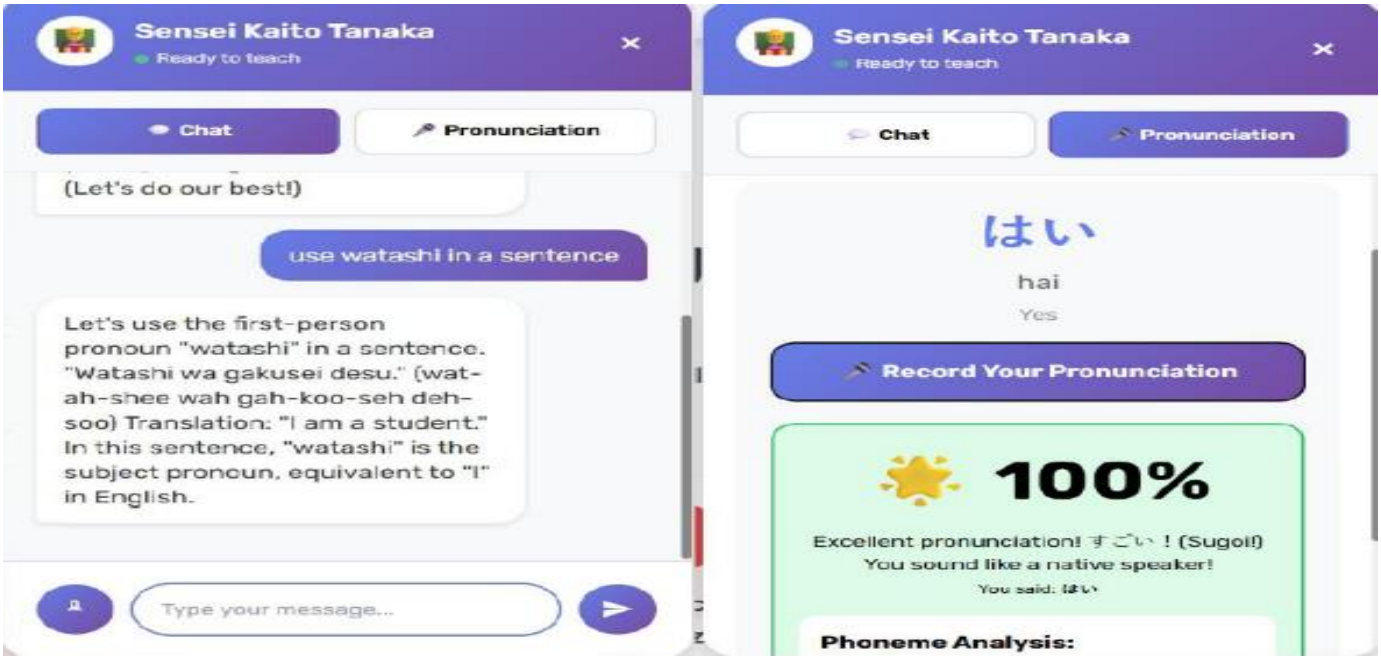


Figure 29. Sample AI Conversation/Pronunciation Practice

Figure 29 demonstrates real-time pronunciation detection and feedback during AI interaction, which learners agreed improved speaking confidence and fluency.

The system includes captioned text displaying Japanese words and phrases in romaji, hiragana, katakana, and kanji during interactions. This feature helps learners connect spoken sounds with written forms, improving comprehension. Respondents expressed satisfaction with the clarity of captions and layout but suggested improvements in navigation speed and screen responsiveness.

The highest-rated indicator was "The lessons improved my understanding of the Japanese language" (WM = 4.08), while the lowest-rated items were related to speaking confidence and platform accessibility without formal enrollment (WM = 3.86).

Usability & Interface	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
11. The platform's interface is visually appealing.	40	124	50	1	4	3.89	3.44	Agree	1
12. The platform is easy to navigate.	34	102	68	10	5	3.68	3.26	Agree	3
13. The design is consistent across different pages.	39	126	48	3	3	3.89	3.44	Agree	1
14. I can easily find the features I need without confusion.	33	100	72	11	3	3.68	3.25	Agree	4
Overall Weighted Mean						3.79	3.35	Agree	

Figure 30. Usability and Interface

The Usability and Interface category obtained an overall mean of 3.79 (Agree), with the highest rating for visual appeal and design consistency (WM = 3.89) and the lowest for ease of navigation (WM = 3.68).

Security & Privacy	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
15. I feel confident that my personal data is protected when using AiToManabi.	37	116	60	4	2	3.83	3.38	Agree	1
16. The system clearly informs me how my personal data is collected, stored, and used.	35	121	55	6	2	3.83	3.37	Agree	2
Overall Weighted Mean						3.83	3.37	Agree	

Figure 31. Security and Privacy

Security and Privacy evaluation resulted in an overall weighted mean of 3.83 (Agree), reflecting user confidence in data protection and clarity on data handling.

Quality & Performance	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
17. The platform meets my expectations for quality.	38	118	56	4	3	3.84	3.39	Agree	3
18. The platform meets my expectations for usability.	41	122	43	8	5	3.85	3.42	Agree	2
19. The platform responds quickly without unnecessary delays.	39	107	53	12	8	3.72	3.31	Agree	5
20. The platform maintains high quality standards in terms of performance.	41	114	55	5	4	3.84	3.40	Agree	4
21. The platform maintains high quality standards in terms of reliability.	42	123	46	4	4	3.89	3.45	Agree	1
Overall Weighted Mean						3.83	3.39	Agree	

Figure 32. Quality and Performance

Quality and Performance showed an overall mean of 3.83 (Agree), with reliability rated highest (WM = 3.89) and responsiveness lowest (WM = 3.72).

The IT Professionals evaluated the system using a survey instrument aligned with the ISO/IEC 25010 Software Product Quality Model, which measures eight core software quality attributes: Functional Suitability, Performance Efficiency, Compatibility, Usability, Reliability, Security, Maintainability, and Portability. A total of three (3) specialists in web and software development participated in this evaluation. The purpose of this assessment was to determine whether the developed system meets international software quality standards in its design, performance, and maintainability.

Adherence to International Quality Standards and Compliance

A. Functional Suitability	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
1. **Functional completeness** - Degree to which the set of functions covers all the specified tasks and user objectives.	0	3	0	0	0	4.00	0.00	Agree	2
2. **Functional correctness** - Degree to which a product or system provides the correct results with the needed degree of precision	1	2	0	0	0	4.33	0.47	Agree	1
3. **Functional appropriateness** - Degree to which the functions facilitate the accomplishment of specified tasks and objectives.	1	2	0	0	0	4.33	0.47	Agree	1
Overall Weighted Mean						4.22	0.38	Agree	

Figure 33. Functional Suitability

Functional Suitability obtained a weighted mean of 4.22 (Agree), indicating that the system features perform effectively and correctly, aligning with the system's functional requirements.

B. Performance Efficiency & Interface	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
4. **Time behaviors** - Degree to which the response and processing times and throughput rates of a product or system, when performing its functions, meet requirements.	2	1	0	0	0	4.67	0.47	Agree	1
5. **Resource utilization** - Degree to which the amounts and types of resources used by the system, when performing its functions, meet requirements.	1	2	0	0	0	4.33	0.47	Agree	1
6. **Capacity** - Degree to which the maximum limits of the system parameter meet requirements.	1	2	0	0	0	4.33	0.47	Agree	1
Overall Weighted Mean						4.44	0.47	Agree	

Figure 34. Performance Efficiency and Interface

Performance Efficiency achieved a weighted mean of 4.44 (Agree), reflecting optimized resource usage and fast response times attributed to Flask-Alpine.js integration and efficient database design.

C. Compatibility	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
7. **Co-existence** - Degree to which a system can perform its required functions efficiently while sharing a common environment and resources with other products, without detrimental impact on any other	1	2	0	0	0	4.33	0.47	Agree	1
8. **Interoperability** - Degree to which two or more systems, products or components can exchange information and use the information that has been exchanged.	2	1	0	0	0	4.67	0.47	Agree	1
Overall Weighted Mean						4.50	0.47	Agree	

Figure 35. Compatibility

Compatibility recorded a weighted mean of 4.50 (Agree), confirming that the system can coexist and integrate with other software without conflicts, supported by its modular structure and browser compatibility.

D. Usability	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
9. **Appropriateness Recognizability** - Degree to which users can recognize whether a system is appropriate for their needs.	2	1	0	0	0	4.67	0.47	Agree	2
10. **Learnability** - Degree to which a system can be used by specified users to achieve specified goals of learning to use the system with effectiveness, efficiency, freedom from risk and satisfaction in a specified context of use.	2	1	0	0	0	4.67	0.47	Agree	2
11. **Operability** - Degree to which a system has attributes that make it easy to operate and control.	2	0	0	0	0	5.00	0.00	Agree	1
12. **User error protection** - Degree to which a system protects users against making errors.	2	1	0	0	0	4.67	0.47	Agree	2
13. **User interface aesthetics** - Degree to which a user interface enables pleasing and satisfying interaction for the user.	2	1	0	0	0	4.67	0.47	Agree	1
14. **Accessibility** - Degree to which a system can be used by people with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use.	3	0	0	0	0	5.00	0.00	Agree	1
Overall Weighted Mean						4.78	0.38	Agree	

Figure 36. Usability

Usability obtained the highest rating among several attributes, with a weighted mean of 4.78 (Agree). IT Professionals highlighted the intuitive layout, clean interface, and clear labeling of menus, noting that the AI Tutor interface was especially user-friendly.

E. Reliability	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
15. **Maturity** - Degree to which a system or component meets needs for reliability under normal operation.	2	0	0	0	0	5.00	0.00	Agree	1
16. **Availability** - Degree to which a system or component is operational and accessible when required for use.	3	0	0	0	0	5.00	0.00	Agree	1
17. **Fault tolerance** - Degree to which a system or component operates as intended despite the presence of hardware or software faults.	1	1	0	0	0	4.50	0.50	Agree	2
18. **Recoverability** - Degree to which, in the event of an interruption or a failure or system can recover the data directly affected and re-establish	1	1	0	0	0	4.50	0.50	Agree	2
Overall Weighted Mean						4.75	0.35	Agree	

Figure 37. Reliability

Reliability achieved a weighted mean of 4.75 (Agree), demonstrating system stability and fault tolerance during prolonged sessions without major crashes or data loss.

F. Security	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
19. **Confidentiality -** Degree to which a system ensures that data are accessible only to those authorized to have access.	3	0	0	0	0	5.00	0.00	Agree	1
20. **Integrity -** Degree to which a system component prevents unauthorized access to, or modification of, computer programs or data.	3	0	0	0	0	5.00	0.00	Agree	1
21. **Non-repudiation -** Degree to which actions or events can be proven to have taken place so that the events or actions cannot be repudiated	3	0	0	0	0	5.00	0.00	Agree	1
22. **Accountability -** Degree to which the actions of an entity can be traced uniquely to the entity.	3	0	0	0	0	5.00	0.00	Agree	1
23. **Authenticity -** Degree to which the identity of a subject or resource can be proved to be the one claimed.	2	1	0	0	0	4.67	0.47	Agree	2
Overall Weighted Mean						4.93	0.21	Agree	

Figure 38. Security

Security received the highest overall rating, with a weighted mean of 4.93 (Agree). Evaluators confirmed that data confidentiality, integrity, and accountability were effectively maintained through Flask's authentication protocols and MySQL's secure data management.

G. Maintainability	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
24. **Modularity** - Degree to which a system or computer program is composed of discrete components such that a change to one	1	1	0	0	0	4.50	0.50	Agree	1
25. **Reusability** - Degree to which an asset can be used in more than one system, or in building other assets.	1	1	0	0	0	4.50	0.50	Agree	1
26. **Analyzability** - Degree of effectiveness and efficiency with which it is possible to assess the impact on a product or system of an intended change to one or more of its parts, or to diagnose a product for	1	1	0	0	0	4.50	0.50	Agree	1
27. **Modifiability** - Degree to which a product or system can be effectively and efficiently modified without introducing defects or degrading existing product quality.	1	1	0	0	0	4.50	0.50	Agree	1
28. **Testability** - Degree of effectiveness and efficiency with which test criteria can be established for a system, product or component and tests can be performed to determine whether those criteria have been met.	1	1	0	0	0	4.50	0.50	Agree	1
Overall Weighted Mean						4.50	0.50	Agree	

Figure 39. Maintainability

Maintainability earned a weighted mean of 4.50 (Agree), indicating that the systems modular programming structure allows independent updates of components without disrupting other functions, ensuring scalability and easier debugging.

H. Portability	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Mean	Standard Deviation	VI	Rank
29. **Adaptability** - Degree to which a product or system can effectively and efficiently be adapted for different or evolving hardware, software or other operational or usage environments.	2	1	0	0	0	4.67	0.47	Agree	2
30. **Installability** - Degree of effectiveness and efficiency with which a product or system can be successfully installed and/or uninstalled in a specified environment.	1	1	0	0	0	4.50	0.50	Agree	2
31. **Replaceability** - Degree to which a product can replace another specified software product for the same purpose in the same environment.	2	0	0	0	0	5.00	0.00	Agree	1
Overall Weighted Mean						4.72	0.40	Agree	

Figure 40. Portability

Portability received a weighted mean of 4.72 (Agree), confirming that the system runs consistently across laptops, tablets, and mobile devices, ensuring broad accessibility for learners.

The IT Professionals' assessment yielded a strong positive response, with an average weighted mean of 4.61 (Agree) across all eight ISO/IEC 25010 quality attributes. These results confirm that the system is functionally complete, efficient, secure, reliable, and maintainable, meeting key standards of software quality.

The Japanese Language Instructor evaluated the system based on its instructional effectiveness, AI-based support, and educational alignment with JLPT N5 standards. The instructor rated all indicators as "Strongly Agree", resulting in a mean score of 5.00 with a standard deviation of 0.00, indicating complete consistency in responses.

A. Learning Effectiveness	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
AiToManabi shows strong potential as a platform for teaching Japanese.	✓				
The system's feature of voice recognition and feedback could be valuable for language learners.	✓				
The integration of captions and text modules can help improve comprehension.	✓				
AiToManabi can effectively support learners at the beginner level.	✓				
The system design has potential to make Japanese learning more engaging compared to traditional methods.	✓				
AiToManabi can help learners achieve JLPT N5–N4 level proficiency.	✓				
The system can support learners who have no prior exposure to Japanese.	✓				
B. Learning Support	**Strongly Agree**	**Agree**	**Neutral**	**Disagree**	**Strongly Disagree**
The system can assist instructors in guiding students through structured lessons.	✓				
AiToManabi can be helpful for learners without access to formal classes.	✓				
AiToManabi could supplement traditional teaching methods effectively.	✓				
Learners can benefit from practicing conversation with an AI tutor.	✓				
The module-based approach could help learners progress step by step.	✓				
The platform's interface is visually appealing.	✓				
The platform is easy to navigate.	✓				

Figure 41. Portability

Figure 34 shows that the system meets the instructional and educational criteria expected of a Japanese language learning platform. The perfect rating confirms that the system effectively supports pronunciation practice, comprehension, and structured learning aligned with beginner-level Japanese language objectives.

The results from all three respondent groups End Users, IT Professionals, and the Japanese Language Instructor demonstrate that the system successfully fulfills its purpose as an accessible and flexible Japanese language learning platform. The seamless interaction between its components reflects a well-balanced integration of technology, education, and usability. The AI-driven feedback system significantly enhances learner engagement and fluency, while caption integration promotes comprehension and reading skills. Furthermore, compliance with software and ethical standards ensures the system's reliability and security.

End Users provided an overall rating of 3.85 (Agree), indicating that the system supports independent Japanese language learning in an engaging and easy-to-use manner. Respondents highlighted the AI Tutor and voice recognition as helpful tools for practicing pronunciation, and captions as effective for improving lesson comprehension.

IT Professionals gave an average score of 4.61 (Agree) based on the ISO/IEC 25010 quality model, confirming that the system meets international standards for functionality, reliability, security, and performance. They noted that the system is stable, efficient, and easy to maintain, with strong usability and portability features.

The Japanese Language Instructor rated the system 5.00 (Strongly Agree), affirming that the system aligns with JLPT N5–N4 beginner-level instruction. The instructor emphasized that the AI and caption features effectively support listening, reading, and speaking skills, making the platform pedagogically sound.

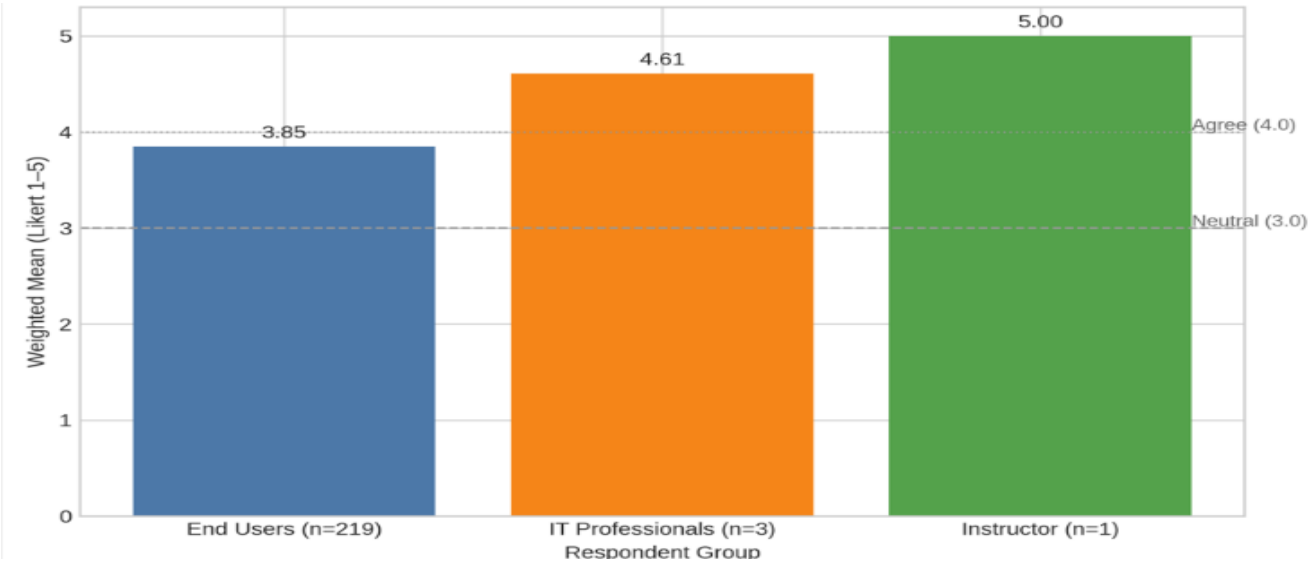


Figure 42. Comparative evaluation ratings across respondent groups (Likert 1-5). End Users (n=219: 3.85; IT Professionals (n=3) 4.61; Japanese Instructor (n=1): 5.00.

SUMMARY OF FINDINGS

The study successfully developed a web-based Japanese Language Learning Management System (LMS) integrated with an AI-powered tutor utilizing LLaMA-3.1-8B, designed to support JLPT N5–N4 learners. The system offers interactive lessons, real-time pronunciation feedback, and captions displaying Japanese scripts with romaji and English translations. It was implemented using Flask, Alpine.js, Tailwind CSS, MySQL, and AI components trained with Hugging Face Transformers and PyTorch.

Evaluation results from three respondent groups demonstrated high satisfaction:

- End Users (n=219): Rated the system positively across learning experience, usability, security, and performance, with an overall weighted mean of 3.85 (Agree). Learners highlighted the AI tutor and voice recognition as effective for pronunciation practice and comprehension.

- IT Professionals (n=3): Assessed the system using ISO/IEC 25010 standards, yielding an average rating of 4.61 (Agree). The system was found to be functionally complete, efficient, secure, reliable, and maintainable, meeting international software quality benchmarks.

- Japanese Language Instructor (n=1): Gave a perfect score of 5.00 (Strongly Agree), confirming the system's alignment with JLPT N5–N4 learning objectives and its effectiveness in improving pronunciation, comprehension, and learner engagement.

Overall, findings indicate that the AI-integrated LMS enhances speaking, listening, and reading skills, promotes learner engagement, and provides a user-friendly platform for independent Japanese language learning. Suggested improvements include expanding datasets, optimizing AI performance, refining usability, and adding higher JLPT levels and other languages for scalability.

CONCLUSIONS AND RECOMMENDATIONS

The study successfully developed and evaluated the Japanese Language Learning Management System with an AI-Powered Tutor Utilizing LLaMA 3.1 8B, designed to support JLPT N5–N4 learners through interactive

lessons, real-time pronunciation feedback, and captioned scripts. Evaluation results from three respondent groups demonstrated strong performance: End Users (n=219) rated the system

positively with an overall weighted mean of 3.85 (Agree), highlighting its effectiveness in improving pronunciation and comprehension through AI-driven features. IT Professionals (n=3) assessed the system using ISO/IEC 25010 standards, yielding an average score of 4.61 (Agree), confirming

compliance with international software quality benchmarks for functionality, reliability, security, and maintainability. The Japanese Language Instructor (n=1) gave a perfect score of 5.00 (Strongly Agree), affirming the system's alignment with JLPT N5–N4 objectives and its pedagogical soundness. Overall, the system enhances speaking, listening, and reading skills, promotes learner engagement, and provides a secure, user-friendly platform for independent Japanese language learning.

Based on the findings, and to further improve and scale the system, the following actions are recommended:

1. Optimize AI Performance - Enhance speech recognition and pronunciation feedback by expanding phoneme datasets and refining model accuracy.

2. Expand Learning Content - Include higher JLPT levels (N3–N1) and additional languages to broaden the system's reach.

3. Improve Usability - Refine navigation speed, screen responsiveness, and mobile compatibility for a seamless user experience.

4. Add Advanced Features - Integrate gamification, adaptive learning paths, and progress badges to boost learner motivation.

5. Strengthen Security - Implement multi-factor authentication and advanced encryption to ensure data privacy.

6. Conduct Long-Term Studies – Perform longitudinal evaluations to measure retention, engagement, and overall learning outcomes. By implementing these recommendations, the system can evolve into a scalable, high-performance platform that meets global standards for AI-assisted language education.

REFERENCES

1. JLPT Official Statistics (2024). JLPT in Charts: Changes in number of applicants (1984–2024). Available: https://www.jlpt.jp/e/statistics/
2. Hetz, G. (2025, May 6; updated Nov 27). Jargonic sets new standards for Japanese ASR. aiOla.ai.
3. Bajo, M., Fukukawa, H., Morita, R., & Ogasawara, Y. (2024, Dec 14). Efficient adaptation of multilingual models for Japanese ASR. arXiv:2412.10705v1.
4. Alotaibi, N. S. A. (2024). The impact of AI and LMS integration on the future of higher education: Opportunities, challenges, and strategies for transformation. Sustainability, 16(23), 10357. https://doi.org/10.3390/su162310357
5. Song, J. Y., Pycha, A., & Culleton, T. (2022). Interactions between voice-activated AI assistants and human speakers and their implications for second language acquisition. Frontiers in Communication, 7, 995475. https://doi.org/10.3389/fcomm.2022.995475
6. Meta (2024, July 22). Introducing LLaMA 3.1: Multilingual LLM with 128K context window. Meta AI Blog.
7. Hugging Face (2024, July 23). LLaMA 3.1 8B Instruct Model Card. Hugging Face Hub.
8. Bayly Castaneda, K., Ramirez Montoya, M.-S., & Morita Alexander, A. (2024, Aug 7). Crafting personalized learning paths with AI for lifelong learning: A systematic literature review. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1424386
9. Sharma, S., Mittal, P., Kumar, M., & Bhardwaj, V. (2025, Apr 4). The role of large language models in personalized learning: A systematic review of educational impact. Discover Sustainability, 6, 243. https://doi.org/10.1007/s43621-025-01094-z

10. Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2023). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. ReCALL, 36(1), 4–21. https://doi.org/10.1017/S0958344023000113
11. Alharthi, S. M. (2024). Siri as an interactive pronunciation coach: Its impact on EFL learners. Cogent Education, 11(1), 2304245. https://doi.org/10.1080/2331186X.2024.2304245
12. Purani, A. (2025, Jul 22). AI Integration in LMS Platforms: Personalized Learning Paths. SevenSquareTech.
13. Qian, C. (2025, Nov 19). The application of artificial intelligence technology in Japanese language education and its impact on learning effectiveness. Proceedings of SPIE, 13940, MLNNCS 2025, 139401O. https://doi.org/10.1117/12.3090376
14. Kaleci, D. (2025). Integration and application of artificial intelligence tools in the Moodle platform: A theoretical exploration. Journal of Educational Technology & Online Learning, 8(1), 100–111.
15. Department of Education (2025, Feb 20). DepEd launches AI center for education. Available: https://www.deped.gov.ph/2025/02/20/deped-launches-ai-center-for-education/
16. Department of Education (2025, Apr 29). DO 013, s. 2025 – Establishment of the Center for Artificial Intelligence Research for the Department of Education. Available: https://www.deped.gov.ph/2025/04/29/april-29-2025-do-013-s-2025-establishment-of-the-center-for-artificial-intelligence-research-for-the-department-of-education/
17. Department of Education (2025, Dec 19). AI research initiative ni Angara, nagbibigay ng data-driven, efficient reforms sa DepEd. Available: https://www.deped.gov.ph/2025/12/19/ai-research-initiative-ni-angara-nagbibigay-ng-data-driven-efficient-reforms-sa-deped/
18. DOST-ASTI (n.d.). iTANONG: A Natural Language Interface to Databases for Filipinos. Available: https://asti.dost.gov.ph/projects/itanong/
19. Rest of World (2024, Feb 19). The Philippines is creating a ChatGPT rival that speaks Filipino and Taglish. Available: https://restofworld.org/2024/filipino-ai-chatbot-launches-2025-local-languages/
20. University of Southeastern Philippines (n.d.). SePhi: USeP Chatbot AI Teaser. Available: https://www.youtube.com/watch?v=GcmnanHkmYU