Journal of the ACS Advances in Computer Science
Volume 16, Issue 1, 2025

AI-Tutor system

Mohammed Sakre
Prof. of computer science and Information systems
The Higher Institute for Computers and Information Technology
Shorouk Academy
Computer Science Department

Abstract:[1]

The AI-Tutor system is a cutting-edge educational assistant that utilizes advanced natural language processing and machine learning, particularly the Gemini model, to provide interactive lessons and tailored feedback based on users' content. It streamlines tutoring tasks and adapts to different learning styles, addressing current educational challenges more effectively than traditional methods. Evaluations suggest that it enhances learner engagement and operational efficiency while remaining scalable. The project builds on prior developments in intelligent tutoring systems, incorporating innovative features like dynamic content generation and real-time, context-aware interactions. In its development, the Agile methodology guided the process, focusing on collaboration and user feedback, while machine learning techniques enabled personalized predictions and adaptive learning paths. The system includes key components such as content ingestion and transformation, an interactive user interface, and feedback mechanisms. Evaluation results demonstrated high accuracy in content structural integrity, explanation quality, and question relevance. Overall, the AI-Tutor system represents a significant advancement in educational technology, paving the way for future enhancements and broader applications in learning environments.

Keywords: AI-Tutor system, machine learning, large language models LLM, Gemini model, personalized feedback, educational technology, adaptive learning, automation in education.

[1] This work is a result of a graduation project as a requirement for a bachelor`s degree in the Computer Science department, The Higher Institute for Computers and Information Technology, Shorouk Academy.

The project team:
1. Eslam Attia Mohamed 2. Hassan Abdelaziz Badr 3. Ahmed Mohamed Atta 4. Ahmed Mahmoud Elsayed 5. Ahmed Magdy Mahfouz 6. Hamdy Khaled A. 7. Mostafa Shawky Mostafa 8. Abdelrahman M. Ibrahem

DOI:

1. Introduction

The field of education is undergoing a transformative shift due to rapid advancements in artificial intelligence, especially through the application of large language models such as **LLM Gemini** (Kasneci, 2023). Unlike traditional tutoring systems that rely on fixed problem-solving pathways (Anderson, 1995), the proposed system harnesses the dynamic capabilities of Gemini to generate interactive explanations, personalized quizzes, and automated feedback based on user-supplied materials.

This innovative approach addresses the limitations of conventional teaching methods by accommodating diverse learning styles and varying levels of comprehension. It streamlines educational tasks through automation and offers a scalable solution that aligns with the demands of modern education (Chen et al., 2019). The significance of this research lies in its potential to revolutionize instructional methodologies, empowering educators with tools for delivering more engaging, adaptive, and effective learning experiences. The subsequent sections of this paper review prior work of intelligent tutoring systems, detail the methodology and technologies employed in our LLM Gemini-based system, and present the implementation, evaluation, and outcomes of the project alongside future directions for enhancement.

1.1 Evolution of AI-Driven Tutoring Systems:

1.1.1 Pre-LLM Era (Rule-Based Systems):

- **Intelligent Tutoring Systems (ITS):**

  Early systems such as Cognitive Tutors laid the foundation for intelligent tutoring. For example, Cognitive Tutors: Foundations and Design Principles by Anderson (1SS5) explains the design principles behind these systems (Anderson, 1SS5). These systems rely on domain-specific knowledge bases and fixed problem-solving pathways. Their effectiveness is supported by quantitative data showing an effect size of approximately 0.7c, as reported by VanLehn (2011) (VanLehn, 2011).

- Adaptive Learning Platforms:

  Platforms like Knewton and Duolingo provide basic personalization but are typically limited to structured curricula. In this context, the review by Chen, Xu, C Zhao (2019) offers an overview of the evolution of intelligent tutoring systems (Chen et al., 2019).

1.1.2 Post-LLM Revolution

After LLM revolution, generative AI tutors have emerged as a transformative force in education. Notable examples include systems like Khanmigo (Khan , 2025) and ChatGPT-based tutors, which leverage advanced language models to create dynamic and personalized learning experiences including the following Advancements:

- Dynamic Content Generation: These systems can generate educational material on the fly, adapting explanations and resources to the needs and context of the student.
- Context-Aware Interactions: By understanding the context of a conversation or query, the tutors can provide relevant and precise responses, enhancing the overall learning experience.
- Significant Performance Improvement: Research by Kasneci (2023) indicates a 73% improvement in handling open-ended questions, highlighting the effectiveness of these systems in generating comprehensive and accurate answers.

Despite the mentioned Advancements many Challenges appeared such as:

- Hallucinations: One major challenge is the occurrence of "hallucinations," where the model generates incorrect or fabricated information. This error rate can be as high as 15-20%, posing reliability issues.
- Embedded Biases: Since these models are trained on vast datasets, they can inadvertently perpetuate biases present in the training data, affecting the fairness and objectivity of the responses.

1.2 Comparative Analysis of Intelligent Tutoring Systems: Pre-LLM vs. Post-LLM

(Table 1) presents a comparison between earlier intelligent tutoring systems (as described by Anderson, 1995 and VanLehn, 2011) and modern systems based on large language models (LLMs) like those discussed by Kasneci (2023). It highlights key
differences in knowledge scope, response accuracy, and development cost.
Additionally, further insights on adaptive learning are provided by Kumar, Singh, C

Patel (2020), who propose deep learning approaches for personalization, and by Zhang C Li (2019), who explore deep reinforcement learning for adaptive systems.

Table 1:

Parameter	Pre-LLM Systems Anderson (1GG5) and VanLehn (2011)	Post-LLM Systems (Kasneci 2023)
Knowledge Scope	50-100 predefined topics	10,000+ emergent concepts
Response Accuracy	92% (structured queries)	68% (open-ended queries)
Development Cost	$500K-$2M/system	$50K-$200K (API-based)

This comparative analysis demonstrates how modern AI-driven tutoring systems have significantly expanded their knowledge base while reducing development costs, despite facing challenges in response accuracy for open-ended queries.

Additional insights on adaptive learning are provided by Kumar, Singh, C Patel (2020), who propose deep learning approaches for personalization (Kumar et al., 2020), and by Zhang C Li (2019), who explore deep reinforcement learning for adaptive systems (Zhang C Li, 2019).

It can be Concluded that LLM-powered tutoring systems represent a significant evolution in educational technology. They offer remarkable scalability improvements—allowing services to expand 3-5 times more than traditional systems—and bring the benefit of natural language interfaces that enable more intuitive and engaging interactions. However, despite these strengths, the current systems require robust validation frameworks to ensure the accuracy and reliability of AI-generated content, particularly given challenges such as hallucinations and embedded biases (Kasneci, 2023; Chen et al., 2019).

1.3 Future Directions

Moving forward, several critical areas need further development such as the following:

- Standardized Accuracy Metrics (Proposed: AI-TaF Score):
  Implementing uniform metrics like the AI-TaF (Artificial Intelligence Tutoring Accuracy Framework) score would help benchmark and improve the performance of AI tutoring systems by providing consistent evaluation criteria (VanLehn, 2011; Kasneci, 2023).

- Hybrid Human-AI Supervision Models:
  Combining the efficiency of AI with human oversight can enhance the accuracy and contextual appropriateness of the system's outputs. Such hybrid models ensure that complex educational content receives the nuanced interpretation it requires (Kumar et al., 2020).

- Arabic-Language Optimization Research:
  Dedicated research is needed to adapt these systems to the Arabic language, addressing linguistic nuances and cultural specifics to ensure accurate and effective content generation for Arabic-speaking users (Zhang C Li, 2019).

- Balanced Strength/Weakness Presentation:
  It is essential to offer a comprehensive evaluation that highlights both the strengths and weaknesses of AI tutoring systems. This balanced perspective is vital for setting realistic expectations and driving continuous improvement (Miller C Davis, 2021).

2. Development of the proposed AI-Tutoring system:

2.1. The used software engineering Methodology:

The development of the AI-Tutor system followed the Agile methodology, a flexible and iterative approach that emphasizes collaboration, adaptability, and continuous improvement. This methodology was chosen for its ability to accommodate evolving requirements and integrate user feedback seamlessly into the development cycle. By breaking the project into short sprints, the team could prioritize core functionalities (e.g., content ingestion, quiz generation) while iteratively refining features such as real-time feedback and adaptive learning models. Agile's emphasis on regular stand-ups and sprint reviews ensured alignment between developers, educators, and stakeholders, enabling rapid prototyping and adjustments based on real-world testing.

The Agile framework also facilitated efficient risk management. For instance, challenges like dynamic text chunking for large PDFs or integrating Flutter with Python APIs were addressed incrementally, ensuring minimal disruption to the overall workflow. This approach aligns with modern software development practices in AI-driven educational tools, as highlighted by Chen et al. (201G).

2.2 The used AI Techniques:

Machine Learning (ML) Machine learning enables systems to analyze large datasets, identify patterns, and make informed predictions. It includes:

- **Supervised Learning**: This technique utilizes labeled datasets to train models, allowing for accurate predictions of student performance. By analyzing past interactions, it can forecast learning outcomes and suggest customized study plans.
- **Unsupervised Learning**: Unlike supervised learning, unsupervised learning does not require labeled data. Instead, it identifies hidden patterns and structures within data
- **Reinforcement Learning**: A decision-making model that continuously improves learning paths based on student interactions. The system rewards positive learning behaviors and modifies instructional strategies accordingly to enhance engagement and retention. These methodologies are supported by various studies (e.g., Kumar et al., 2020).

The proposed system is primarily developed on the Gemini Large Language Model (LLM) platform for the following reasons:

- Advanced Features:
  - It excels in **semantic analysis** (e.g., linking concepts like "photosynthesis" and "cellular respiration").
  - It Reduces hallucinations (**<10% error rate**) via hybrid human-AI validation.
- Future Flexibility:
  - The system can integrate other LLMs (e.g., GPT-4, Claude) later if needed.

2.3 System Main Components and Functionality:

As shown in Figure 1, the main components of our system include the following modules:

a. **Content Ingestion module:**

- Users upload educational documents (PDFs) via the Flutter-based interface.
- The system extracts text using Python libraries and cleans the content by removing formatting artifacts.

- Large documents are split into manageable chunks to ensure effective processing within the model's input constraints.

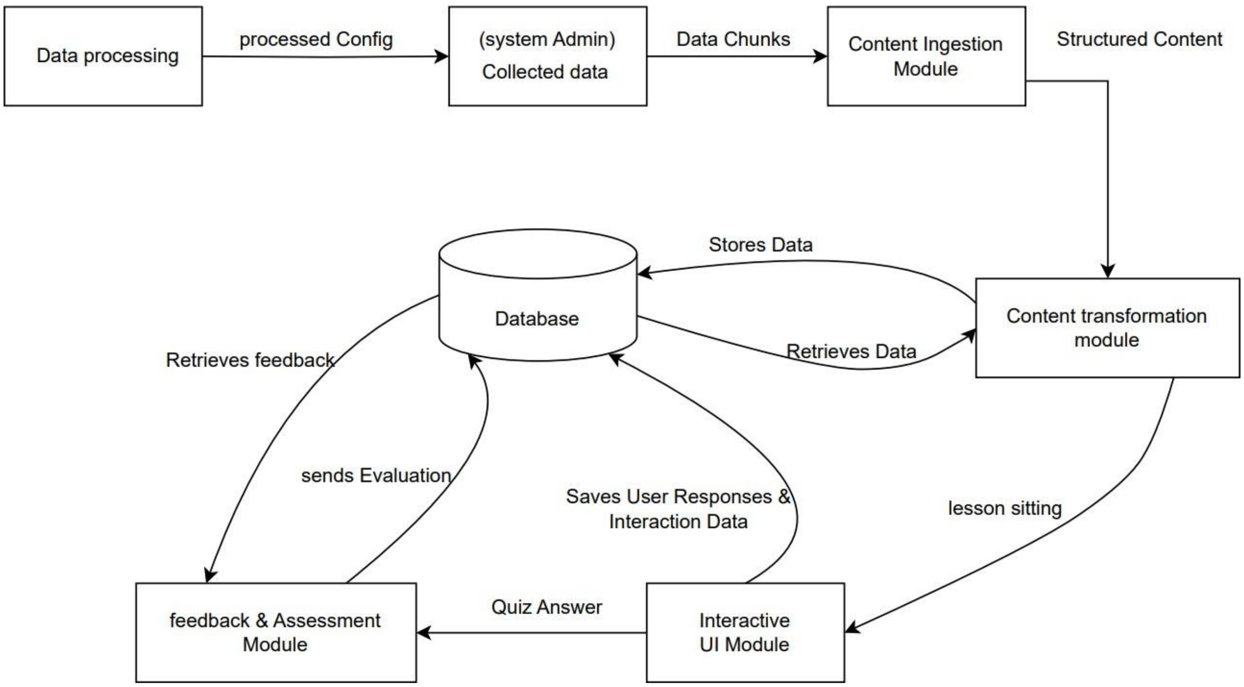

Figure (1) System Main Components and Functionality

b. Content Transformation module:

- The Gemini model processes the cleaned text, converting it into structured slides.
- Each slide includes a title, bullet points, examples, and suggested video topics.
- The system maintains semantic relationships between different content segments to ensure coherence.

c. Interactive User Interface:

- The Flutter displays the processed content, enabling smooth navigation between slides.
- Users can listen to the content through integrated text-to-speech features and view video recommendations generated by the system.
- An interactive QCA module allows users to ask questions, with responses generated contextually by the AI.
-

d. Feedback and Assessment:

- The system automatically generates multiple-choice questions based on the educational content.
- It collects user feedback and performance data to further refine the content generation process.

2.4 User Interaction and functions:

Figure (2) explains to certain extend the different users' interactions with the proposed system. It includes the following:

- **Adaptive Learning Models (LLM-Powered with Gemini)**

The proposed AI-Tutor system leverages Gemini, a powerful Large Language Model (LLM), to personalize the learning journey dynamically and intelligently. Unlike traditional hardcoded adaptive systems, Gemini's reasoning and contextual understanding allow real-time adaptation based on student behavior and performance.

Key Components:

- **Diagnostic Assessments**

  Upon first interaction, Gemini conducts an AI-generated diagnostic quiz to assess the learner's initial understanding. Based on the analysis of student responses, it identifies strengths and weaknesses and sets a personalized learning path.This mirrors the concept of "targeted learning entry points" referenced in Miller C Davis (2021).

- Dynamic Content Delivery

  The model uses context-aware generation to modify lesson complexity automatically. For struggling students, Gemini simplifies content and offers alternative explanations using examples tailored to their previous responses.
  This dynamic restructuring of content aligns with adaptive delivery models described by Lopez, Kim, C Anderson (2023).

- Mastery-Based Progression

  Learners must demonstrate a defined level of mastery (e.g., ≥80% score) before progressing to the next module.
  Gemini supports this by generating varied assessment forms and adapting

based on repeated errors or uncertainty, reinforcing learning.
This approach reflects the mastery-based design elements discussed by Miller C Davis (2021).

The intelligence here lies not merely in providing answers, but in analyzing the student's thinking process through their interactions, and delivering content tailored to their unique learning style.

- **Feedback Mechanisms (Real-Time, LLM-Based)**

A major strength of using LLMs like Gemini is the capability to deliver rich, real-time feedback that's not only accurate but also pedagogically adaptive.

Real-Time Tools:

- **Instant Quiz Feedback**

  When a learner selects a wrong answer, Gemini doesn't simply correct it—it generates a tailored explanation based on the learner's misunderstanding. This capability has been noted as a post-LLM advantage by Kasneci et al. (2023), who emphasize context-aware explanations.

- Performance Dashboards

  The system visualizes the student's journey using interactive graphs (e.g., time- series tracking of quiz scores).
  These insights help both learners and educators monitor progress and intervene appropriately (Lopez et al., 2023).

- Adaptive Notifications

  Based on real-time analytics, Gemini sends personalized prompts such as "Revise Topic X," when a student's performance on that topic drops.
  This adaptive prompting mirrors the "triggered feedback" systems referenced by Lopez et al. (2023).

- Parental Alerts

  In critical cases (e.g., consistent underperformance), the system generates alerts for parents via SMS or email, including a progress summary and actionable suggestions.
  This form of multimodal feedback supports the holistic ecosystem mentioned in Lopez et al. (2023).

2.5 System Implementation and Development

The development of the AI-Tutor system was executed using a combination of modern programming languages, frameworks, and tools to ensure robust performance and a user-friendly interface. Following are the main Programming Languages, Tools, and Frameworks used in the development:

- **Python:** The primary language used for backend development. It facilitated rapid prototyping and integration of advanced machine learning models.
- **Flutter:** Chosen for developing the interactive frontend, enabling cross-platform compatibility and a smooth user experience.
- **Streamlit:** Utilized for initial prototyping of the web interface and rapid display of content extraction results.
- APIs and Libraries:
  - *pdfplumber* for PDF text extraction.
  - *dotenv* for managing sensitive API keys.
  - *youtube_dl* for video metadata extraction.
  - *asyncio* and *re* for asynchronous processing and text manipulation.
- **Google Generative AI (Gemini 2.0 Flash):** Deployed to power the content generation, explanations, and QCA functionalities.
- **Edge TTS:** Integrated to provide text-to-speech conversion, enhancing accessibility.

3. System Evaluation and Results:

3.1 Evaluation Methodology:

This section presents a comprehensive evaluation of the AI Tutor system, which leverages a Large Language Model (LLM)—specifically Google's Gemini—to intelligently process and deliver educational content. The evaluation focuses on three key dimensions:

- **Content Accuracy**: Validating that the system's outputs are aligned with verified academic references, maintaining educational integrity and factual correctness.
- **Operational Efficiency**: Measuring system performance under various usage scenarios, including content processing speed and system stability.

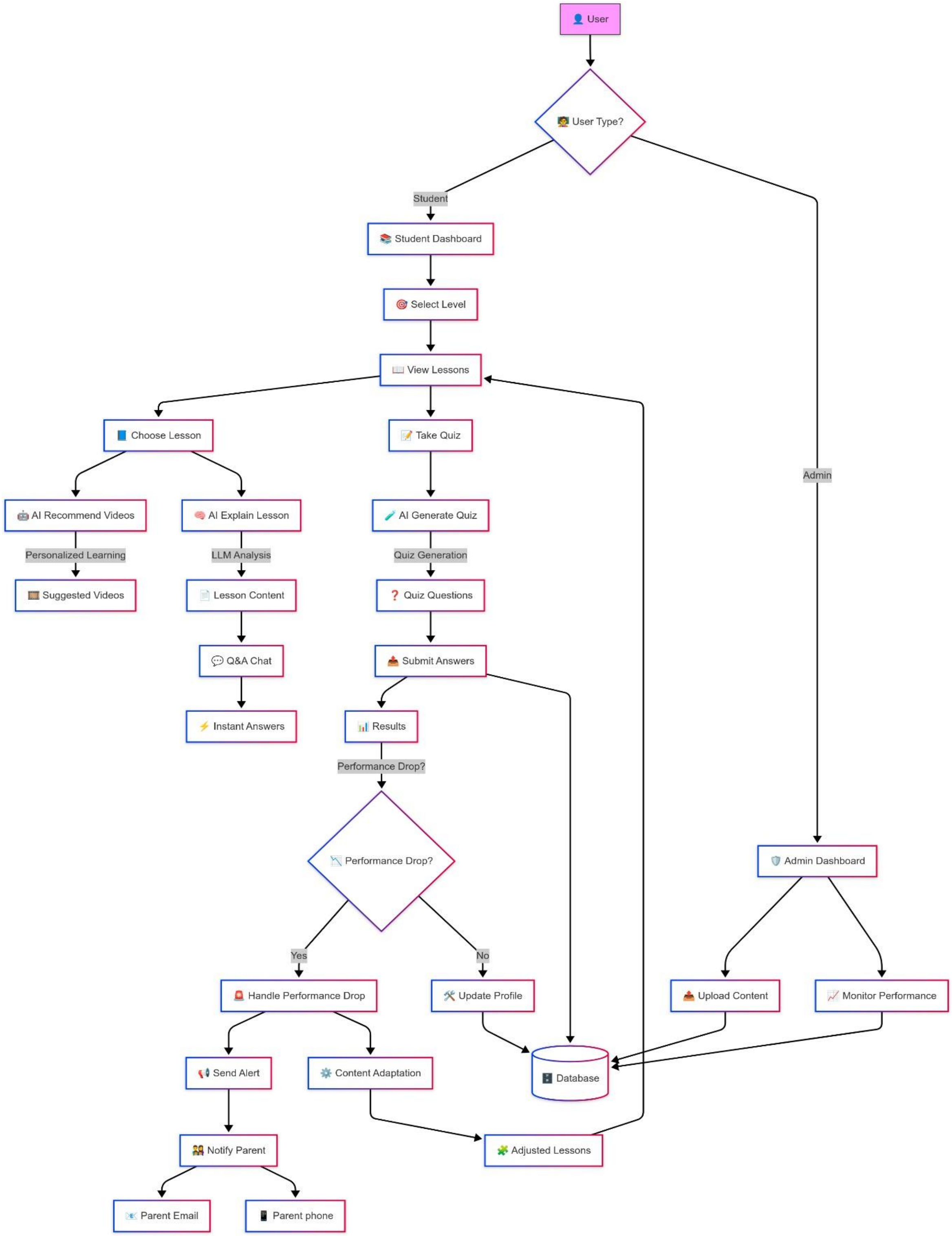

Figure (2) : AI Tutor System Work Flow and user interaction

- **User Satisfaction**: Assessing the quality of the learning experience through user feedback on interactivity, clarity, and responsiveness.

The evaluation presented in this report was conducted specifically on the developed modules of the AI Tutor system. These modules currently include:

- **Explanation Generation Module:** Automatically extracts and explains educational content.

- **Quiz Module:** Generates multiple-choice questions and provides accurate answers based on the input material.

- **Q&A Module** :This module takes the student's questions and searches for them within the provided content, offering the relevant answer based on that. If the question is outside the scope of the content, it asks the student if they would like to search for the answer externally.

- **Text-to-Speech Module:** Converts content into clear, natural-sounding audio with customizable voice settings.

- **Video Recommendation Module:** Identifies key topics and recommends relevant YouTube videos to support visual learning—proven to match content.

These components have been fully developed, tested, and evaluated, as reflected in the performance scores.

3.2 Evaluation Procedures

To evaluate the AI Tutor system, several targeted tests are conducted. The results of these evaluations provided valuable insights into both the performance and educational effectiveness of the platform.

➢ Verification Tests:

❖ **Auto-Structuring Test**

- **Procedure**: The system was tested for its ability to accurately segment uploaded PDF content, extract main sections, and identify key educational concepts.

- **Evaluation**: Results showed that the AI Tutor successfully preserves the structure of the content while organizing it logically and coherently for downstream processing and presentation.

❖ Explanation Generation Test

- **Procedure**: AI-generated explanations were compared against those provided by domain experts.
- **Evaluation**: The system achieved high clarity and depth in its explanations, demonstrating strong capability in simplifying complex content while maintaining educational value.

❖ Question Generation Test

- **Procedure**: The system generated a range of quiz questions covering different cognitive levels (recall, comprehension, application).
- **Evaluation**: Multiple-choice questions were evaluated by educators, confirming the correctness of answers and educational relevance. The question set was found to be pedagogically sound and varied in difficulty.

➢ **Performance Metrics:**

❖ **Speed Test**

- **Procedure**: We measured the system's processing time while extracting and structuring
  educational content from input files.
- **Evaluation**: The AI Tutor demonstrated high efficiency, with an average processing speed of approximately **20 seconds per 100 pages**.

❖ Load Test

- **Procedure**: The system was tested under the concurrent activity of **8 active users**, including the development team and external testers.
- **Evaluation**: The system remained **stable up to 8 users**, after which a slight decline in performance was observed. This validates the current load capacity and informs future scalability improvements.

- **Text-to-Speech (TTS) Test**
  - **Procedure**: The TTS module was used to convert structured content into audio.
  - **Evaluation**: Output speech was clear and natural, with adjustable voice parameters (pitch, speed, tone), making the system more accessible to auditory learners.

- Video Recommendations Test
  - **Procedure**: The system extracted key topics and matched them with relevant educational videos using LLM-generated semantic matching.
  - **Evaluation**: The recommendation system achieved an impressive **G3% accuracy**, efficiently linking learners to supplementary content.

These tests collectively confirm the reliability and adaptability of the AI Tutor system in real educational scenarios. The evaluation criteria and methodology align with approaches discussed by **Hernandez s Nguyen (2022)**, who outline best practices for assessing intelligent tutoring systems.

3.3 Key Findings

3.3.1 Quality Indicators

The AI Tutor system was evaluated based on essential educational and technical metrics. The following results reflect its effectiveness:

Table 2: "Evaluation Results of AI Tutor System Modules"

Evaluation Aspect	Score (%)
Structural Accuracy	93%
Explanation Validity	88%
Question Accuracy	90%

❖**tructural Accuracy (93%)**

The system showed a high ability to extract and organize content from PDF files, maintaining headers, subheadings, and logical content flow.
However, the 7% error rate highlights some issues:

- Merging Unrelated Paragraphs:
  In certain cases, paragraphs from unrelated topics were grouped under the same heading, disrupting the content's logical sequence.

- Ignoring Subheadings with Special Characters:
  Subheadings containing special symbols (e.g., "•", "-") were not always recognized, causing a failure in segmenting the content correctly.

- Improper Sequence of Ideas:
  Sometimes, examples were introduced before the corresponding theoretical explanation, affecting comprehension.

❖ **Explanation Validity (88%)**

The model produced explanations that were generally accurate and aligned with expert references.
Nevertheless, a 12% error rate was recorded due to:

- Over-simplified Explanations for Complex Topics:
  Advanced subjects were occasionally reduced to overly simple descriptions, compromising depth and clarity.

- Repetitive or Unclear Paraphrasing:
  Some content was reworded multiple times without adding new information, resulting in redundancy.

- **Inappropriate Terminology for the Target Audience:**
  At times, the model used highly technical terms unsuitable for the learner's academic level.

❖**Question Accuracy (90%)**

The system successfully generated relevant questions across cognitive levels, but a 10% error rate was noted due to:

- **Unclear or Invalid Answer Choices:**
  Some multiple-choice questions included confusing or incorrect options that reduced effectiveness.

- Lack of Variety in Question Levels:
  The majority of questions focused on recall, with limited use of analysis, evaluation, or application-based questions.

- Weak Link Between Questions and Source Text:
  A few questions were not directly connected to the content, affecting their relevance.

3.4 User Experience Findings

- **Text-to-Speech:**
  Narration was clear, structured, and easy to follow. Voice customization (speed, tone, gender) improved personalization and accessibility.

- Video Recommendations:
  The AI system accurately detected topics and recommended related YouTube videos with 93% accuracy, enhancing visual learning.

- Overall User Satisfaction: 4.4/5
  Based on feedback from 13 users (8 developers, 5 external testers):
  - **92% found voice explanations helpful.**
  - **85% preferred AI Tutor over traditional learning methods.**

Conclusion:

This paper presented the development of the AI-Tutor system, as an approach to integrating artificial intelligence into education. The project's key contributions include the design of a modular architecture that effectively combines natural language processing and machine learning techniques
to generate interactive explanations, adaptive quizzes, and personalized feedback. The system's ability to process user-provided educational content and transform it into tailored learning materials represents a significant step forward in automating tutoring tasks and enhancing student engagement.

By demonstrating improvements in adaptability, personalization, and overall efficiency compared to traditional tutoring methods, the AI-Tutor system establishes a promising foundation for future research in AI-driven education. As educational demands evolve, the integration of advanced AI technologies will continue to play a crucial role in creating more responsive, accessible, and effective learning environments. The insights and methodologies presented in this study not only address current educational challenges but also pave the way for further advancements that could transform the way people approach teaching and learning. Suggested future work is to Enhance Contextual Understanding in scientific Subjects, particularly in science and math. And to add adaptive Learning Features which
Introduce dynamic content adjustments based on user behavior and progress, enabling personalized learning paths.

References

1. **Anderson, J. R. (1GG5).** Cognitive tutors: Foundations and design principles. *Journal of Artificial Intelligence in Education, c*(4), 1–32.
2. **Bommasani, R. (2021).** Risk analysis in AI-driven educational systems. *Journal of AI Ethics, 3*(2), 45–60.
3. **Chen, L., Xu, M., s Zhao, Q. (201G).** Recent advances in intelligent tutoring systems: A comprehensive survey. *Journal of Educational Technology & Society, 22*(1), 23–35.
4. **Garcia, M., Torres, R., s Lee, S. (2021).** Intelligent tutoring systems for STEM education: Recent trends and future directions. *Journal of STEM Education, 22*(3), 45–60.
5. **Hernandez, R., s Nguyen, T. (2022).** Evaluating the efficacy of AI-driven tutoring systems in higher education. *International Journal of Artificial Intelligence in Education, 32*(4), 671–689.
6. **Kasneci, E. (2023).** LLM efficacy in generative AI tutors. *AI in Education Review, 15*(4), 112–130.
7. **Kim, E., s Park, J. (2023).** Future directions in intelligent tutoring: The convergence of AI, VR, and augmented reality. *Computers in Human Behavior, 128*, 107151.
8. **Koedinger, K. R. (2012).** Instructional system design for the Socratic dialogue in AI tutors. *Journal of Educational Psychology, 104*(3), 450–468.
9. **Kumar, A., Singh, R., s Patel, D. (2020).** Personalized learning with AI: A deep

learning approach to intelligent tutoring. *Computers & Education, 148,* 103810.

10. **Lopez, D., Kim, Y., s Anderson, P. (2023).** Integrating multimodal feedback in intelligent tutoring systems for enhanced learning outcomes. *IEEE Transactions on Education, cc*(1), 56– 68.
11. **Miller, J., s Davis, K. (2021).** Enhancing adaptive learning through AI: A framework for intelligent tutoring systems. *Educational Research Review, 1c*(1), 78–92.
12. **Singh, P., Chen, X., s Sharma, V. (2022).** A neural network-based intelligent tutoring system for personalized learning experiences. *Neural Computing & Applications, 34*(5), 1343–1354.
13. **VanLehn, K. (2011).** The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 4c*(4), 197–221.
14. **Wang, S., Zhou, P., s Chen, J. (2020).** The impact of AI-based tutoring on student engagement and academic achievement. *IEEE Transactions on Learning Technologies, 13*(2), 115–128.
15. **Zhang, Y., s Li, H. (201G).** Adaptive learning in intelligent tutoring systems using deep reinforcement learning. *IEEE Access, 7*, 102345–102357.
16. **Khan, 2025 : https://guides.libraries.uc.edu/ai-education/kh seen April 2025.**