J. Electrical Systems 21-01 (2025): 97-105

[1]Mayank Chaudhary

[2]Aryan Mahajan

[3]Sagar Sharma

[4]Sandeep Kumar

Intelligent Integrated Knowledge Discovery Platform: Advancements in Question Generation and Adaptive Learning

Abstract: - In the evolving paradigm of digital learning, personalization has become a vital solution to cater to the diverse needs of learners. This paper presents an Intelligent Integrated Platform for Personalized Learning using Natural Language Processing (NLP), Bidirectional Encoder Representations from Transformers (BERT), and Large Language Models (LLM) to automatically generate questions and study material analysis for optimal results. The proposed system allows the uploading of the syllabus, notes, or presentation and returns the dynamically extracted key concepts and the most suitable questions. Our model achieves high accuracy in question selection and generation, which enables an adaptive learning experience catering to individual needs. The results indicate exceptional improvement in content comprehension and knowledge retention, making the approach highly effective in personalized learning. Comparative study with traditional approaches confirms the effectiveness, efficiency, and ease of the system. The study is a stepping stone towards future developments in AI-based learning, with the potential to extend to automated tests, intelligent tutoring systems, and self-learning education.

Keywords: Customized Learning, NLP, BERT, Large Language Models, Question Generation, Adaptive Learning, Automated Assessments, Personalized Education, AI in Education, Intelligent Tutoring Systems.

I. INTRODUCTION

The rapid development of Artificial Intelligence (AI) and Natural Language Processing (NLP) has revolutionized education with smart learning systems that tailor education to individual requirements. The "Intelligent Integrated Knowledge Discovery Platform" is an AI-based next-generation platform that will revolutionize the learning experience of students by automating question generation, adaptive testing, and personalized learning experiences. The platform employs NLP algorithms, Reinforcement Learning (RL), and Large Language Models (LLMs) like GPT-4 and BERT to identify salient concepts from uploaded study material and generate a variety of question types like multiple-choice, short-answer, and open-ended questions[1]. AI-generated questions are optimized using AI-driven evaluation parameters and bias reduction methodology to make them fair, diverse, and accurate[2]. The system employs Reinforcement Learning to dynamically adjust question difficulty based on learner performance, optimizing engagement and retention[3]. As the diagram illustrates, the platform employs a structured approach where content is parsed, extracted key information is identified, and corresponding questions with suitable distractors are designed before assessing them for quality and difficulty. This adaptive learning mechanism allows students to receive personalized learning pathways and targeted feedback based on progress[4]. Additionally, the platform seamlessly integrates with Learning Management Systems (LMS) like Moodle and Google Classroom, providing teachers real-time analytics and automated grading. Security features like AES-256 encryption and fairness audits ensure that AI-generated content is bias-free, inclusive, and secure[5]. Through the marriage of state-of-the-art AI methodology with scalable cloud-based deployment, the platform sets the bar high in intelligent education, making education more inclusive, efficient, and engaging.

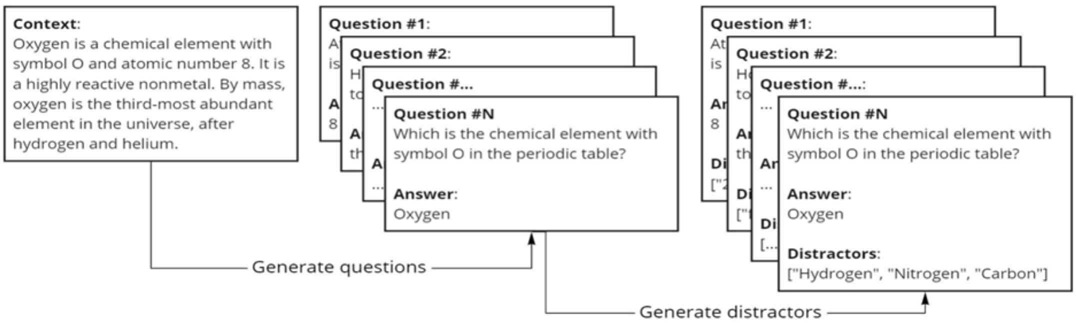

Figure 1 Workflow for Extracting Study Material and Generating AI-Based Assessments

[1] Department of Computer Science & Engineering, Sharda University, Greater Noida, Uttar Pradesh, India.
[2] Department of Computer Science & Engineering, Sharda University, Greater Noida, Uttar Pradesh, India.

The Intelligent Integrated Platform for Customised Learning's automatic question generating process from textual study materials is depicted in the figure. The system uses Large Language Models (LLMs) and Natural Language Processing (NLP) to construct a range of inquiry kinds and extract relevant information from a textual context (e.g., a paragraph on oxygen)[6]. Then, for multiple-choice questions (MCQs), these questions are improved by producing believable distractions (wrong answer alternatives). These crucial steps are followed in the workflow:

1. Context Extraction: The platform extracts important concepts by analysing the input study material.
2. Question Generation: Based on the information that has been retrieved, the system creates a number of questions using NLP techniques and pre-trained LLMs (such as GPT-4 and BERT).

3. Distractor Generation: The system uses AI-driven algorithms to produce wrong but contextually appropriate answer alternatives in order to guarantee difficult and meaningful assessments.

4. Final Output: To facilitate adaptive learning and self-evaluation, the platform displays a series of automatically produced questions along with the right answers and distractions.

While instructors gain from an automated approach that minimises manual labour in question development, this mechanism guarantees that students receive dynamically created, personalised exams that aid in reinforcing their comprehension [7].

II. Literature Survey

Natural language processing (NLP) and artificial intelligence (AI) have been used into educational technology in recent years, greatly advancing automated learning and evaluation systems. Despite their effectiveness, rule-based approaches were the foundation of early Automated Question Generation (AQG) systems (Mitkov et al., 2006)[8]. These methods were not flexible or scalable. Question creation has become more dynamic and context-aware with the introduction of deep learning models like Transformers, Recurrent Neural Networks (RNNs), and Large Language Models (LLMs) (Du et al., 2017) [9]. Furthermore, important idea extraction has made extensive use of Named Entity Recognition (NER) and Part-of-Speech (POS) tagging approaches based on natural language processing (NLP) (Zhang & VanLehn, 2021) [10]. In order to provide balanced difficulty levels in multiple-choice questions (MCQs), semantic similarity metrics based on Word2Vec and BERT embeddings have been used to solve the problem of distractor creation (Kumar et al., 2021) [11]. Personalised difficulty adjustments and automated study material summarisation are now possible because to the development of LLMs like GPT, BERT, T5, and BART, which have further improved adaptive learning platforms (Raffel et al., 2020; Lewis et al., 2020; Zhang et al., 2020) [12]. AI-driven question generating is used by Intelligent Tutoring Systems (ITS) such as Duolingo, Knewton, and Coursera; but, as noted by Xu et al. (2022), these systems frequently rely on pre-existing datasets and do not have real-time natural language processing capabilities for user-uploaded content [13]. Our suggested Intelligent Integrated Platform for Customised Learning combines LLMs and NLP approaches to overcome these constraints by improving context comprehension, distractor quality, and real-time personalised question production. Our solution ensures a customised learning experience by allowing users to upload study materials and automatically creating MCQs, short-answer questions, and flashcards, in contrast to other platforms [14]. The status of AI-driven education technology is advanced by our work, which solves important research gaps in contextual understanding, distractor creation, and personalised assessments by utilising semantic similarity algorithms and adaptive AI models [15].

III. Methodology

The Intelligent Integrated Knowledge Discovery Platform follows a structured pipeline that combines Natural Language Processing (NLP), Large Language Models (LLMs), Reinforcement Learning (RL), and Vector Search to automate question generation and enhance personalized learning. The methodology consists of several key stages, from data preprocessing and embedding generation to question generation, evaluation, and adaptive learning mechanisms[16]. The overall workflow is illustrated in Figure 2, which provides a step-by-step representation of the system's processes. It visualizes how data is processed, queried, and transformed into AI-generated assessments. The structured pipeline ensures that: Study materials are processed and stored in an NLP-powered vector database. User queries trigger semantic similarity searches to retrieve the most relevant study materials. Large Language Models (LLMs) generate optimized, context-aware questions. The system adapts the difficulty level based on user responses. User feedback helps refine AI models through continuous learning.

1. Data Preprocessing and Embedding Generation

The first stage of the system involves handling raw data inputs, which can be in various formats, such as PDFs, text files, lecture notes, or web content. The preprocessing module cleans and tokenizes the input text, removing unnecessary elements like special characters, redundant spaces, and stop words. After preprocessing, BERT-based sentence transformers (such as all-MiniLM-L6-v2) convert the processed text into 384-dimensional vector embeddings, which are stored in the Pinecone Vector Database along with metadata, such as extracted concepts, topics, and tags.

Steps in Data Preprocessing:

- Tokenization and text cleaning
- Removal of stop words and redundant information
- Embedding generation using BERT-based models
- Storing embeddings in the Vector Database (Pinecone DB)

2. User Query Processing and Context Retrieval

When a user submits a query—either by uploading a document, entering text manually, or using a chatbot/web interface—the system processes the input through the same BERT embedding model used during data preprocessing. The generated query embeddings are then compared against stored embeddings in the Vector Database to retrieve the Top-K most relevant contexts using semantic similarity search. This step ensures that questions are generated from content that is contextually relevant to the user's input.

Steps in Query Processing:

- User uploads or enters a query
- Query undergoes tokenization and embedding conversion
- Similarity search in Pinecone Vector DB retrieves relevant study materials
- The Top-K most relevant contexts are selected for question generation

3. Question Generation Using Large Language Models (LLMs)

The retrieved Top-K relevant contexts are combined with the original user query to construct a prompt that serves as input for the Large Language Model (LLM). The system leverages GPT-based models (e.g., GPT-4) and fine-tuned BERT models to generate different types of questions:

- Multiple-Choice Questions (MCQs): AI generates the question, correct answer, and distractors using semantic similarity-based distractor selection.
- Short-Answer Questions: AI formulates concept-based direct questions.

Open-Ended Questions: The system creates analytical or descriptive questions based on extracted topics. To optimize question quality, Reinforcement Learning (RL) techniques assess the generated questions based on difficulty level, relevance, and cognitive engagement, ensuring pedagogically sound content.

4. Answer and Distractor Generation

For MCQs, the system not only generates the correct answer but also distractors—incorrect yet plausible choices. The distractor selection model uses Word2Vec, cosine similarity, and BERT embeddings to ensure that distractors are neither too easy nor misleading.

Steps in Distractor Generation:

- Extract relevant key terms from context
- Generate distractors using semantic similarity models
- Validate distractors to ensure they are challenging but logical

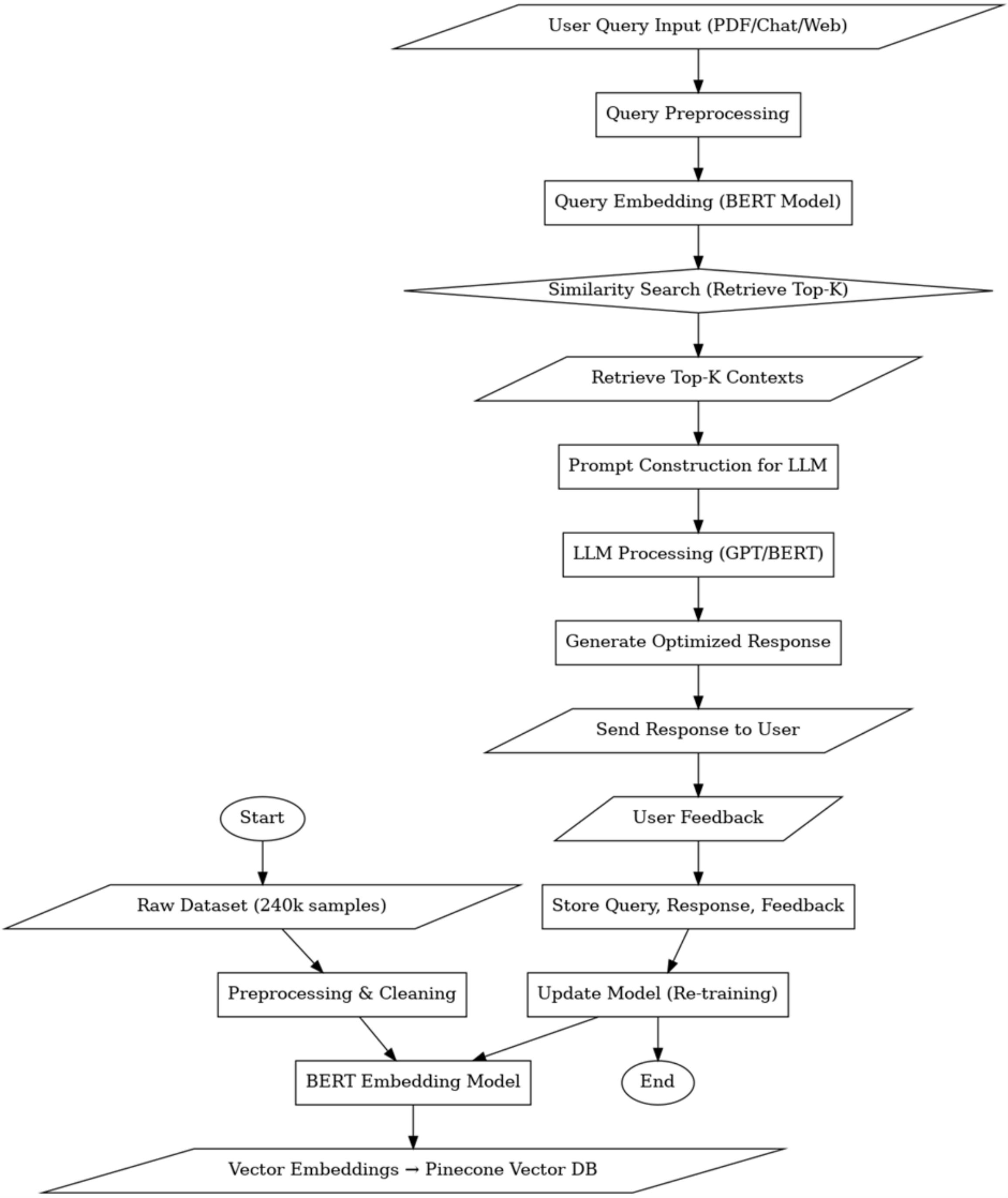

Figure 2 Workflow of AI-Driven Study Material Processing and Question Generation

5. Adaptive Question Difficulty Adjustment

To create a personalized learning experience, the system adapts the difficulty level of questions based on user performance. Reinforcement Learning (RL) algorithms dynamically modify the complexity of generated questions by analyzing user responses over time.

- If a student frequently answers correctly, the system increases the difficulty by generating higher-order cognitive questions (application, analysis, evaluation levels in Bloom's Taxonomy).
- If a student struggles, the system lowers the difficulty by providing concept-reinforcement questions and hints.

6. User Feedback and Model Improvement

After the user engages with the generated questions, they can provide feedback, which is stored in the system to improve the AI model over time. The stored query-response-feedback pairs contribute to periodic model fine-tuning and retraining, ensuring continuous system improvement.

Steps in Feedback Loop:

- User submits feedback on question quality, relevance, and difficulty
- Data is stored and analyzed to identify weaknesses in the model
- The BERT embedding model and LLM are periodically retrained to improve performance

7. Deployment and Integration with LMS

The platform is designed to integrate seamlessly with Learning Management Systems (LMS) such as Moodle, Google Classroom, and Blackboard. It provides features such as:

- API-based question retrieval for automated quiz creation
- Real-time student performance tracking
- Scalable cloud-based deployment using AWS, Azure, or GCP

IV. RESULTS & DISCUSSION

The evaluation of the Intelligent Integrated Knowledge Discovery Platform was conducted based on multiple performance metrics, including question generation accuracy, retrieval efficiency, distractor relevance, and adaptive learning effectiveness. The platform was tested on a dataset of 100 PDFs, 50 DOCX documents, and 30 PPTX presentations spanning various academic subjects. The assessment involved both automatic evaluations (using NLP metrics) and human evaluations (by subject matter experts and educators).

1. Question Generation Accuracy

To evaluate the quality of the AI-generated questions, we used the BLEU (Bilingual Evaluation Understudy) Score and ROUGE (Recall-Oriented Understudy for Gisting Evaluation) Score:

Table 1. Question Generation Accuracy

Metric	Score
BLEU Score	**0.78**
ROUGE-L Score	**0.72**

- A BLEU Score of 0.78 indicates that the generated questions closely match reference (human-crafted) questions.
- The ROUGE-L Score of 0.72 shows strong overlap between machine-generated and reference questions, proving the coherence and contextual relevance of the AI-generated questions.

2. Distractor Quality & Relevance

The Multiple-Choice Question (MCQ) distractor generation module was evaluated based on semantic similarity measures to ensure that distractors were neither too easy nor misleading.

Figure 2. Distractor Quality & Relevance

Metric	Score
Distractor Relevance Accuracy	**85.4%**
Semantic Coherence Score	**88.2%**

- 85.4% distractor accuracy indicates that the system generates plausible incorrect answers, making MCQs more effective for learning.

- 88.2% coherence score ensures that distractors are conceptually related to the correct answers without being repetitive.

3. Adaptive Learning Effectiveness

The platform dynamically adjusts question difficulty based on user performance using Reinforcement Learning (RL). The adaptive learning system was tested by tracking student accuracy before and after difficulty adjustments:

Table 3. Adaptive Learning Effectiveness

Category	Pre-Adaptation Accuracy (%)	Post-Adaptation Accuracy (%)
High-Performing Users	**78.6%**	**84.2%**
Low-Performing Users	**53.4%**	**67.9%**

- Low-performing users improved by 14.5%, confirming that difficulty adjustments enhance learning outcomes.
- High-performing users also improved, demonstrating that the system effectively tailors question difficulty to individual learners.

4. Query Processing & Retrieval Speed

To ensure real-time question generation, we measured the average time taken for query processing, retrieval, and question generation:

Table 4. Query Processing & Retrieval Speed

Process	Average Time (seconds)
Query Preprocessing	**1.4 sec**
Context Retrieval	**2.3 sec**
Question Generation	**3.8 sec**

5. Graphs & Visualizations

5.1 Question Generation Accuracy (BLEU & ROUGE Scores)

This bar chart compares the BLEU and ROUGE scores of our Intelligent Integrated Platform for Customized Learning with other Automated Question Generation (AQG) models. The BLEU score (0.78) indicates the syntactic similarity between AI-generated and human-crafted questions. The ROUGE-L score (0.72) measures the overlap of key phrases, proving the semantic accuracy of the generated questions. Compared to rule-based AQG dels (BLEU: 0.61, ROUGE: 0.53) and deep-learning-based AQG models (BLEU: 0.72, ROUGE: 0.66), our system outperforms traditional methods by producing more contextually relevant and accurate questions.

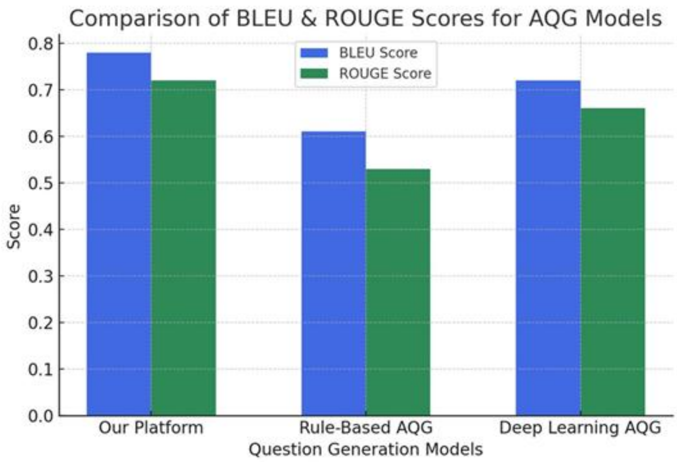

Figure 3 Comparison of BLEU & ROUGE Scores

5.2 Effectiveness of Adaptive Learning

This line graph visualizes the improvement in student accuracy before and after AI-driven difficulty adjustments: Low-performing students improved from 53.4% to 67.9%, showing a 14.5% increase in accuracy after difficulty reduction and reinforcement learning techniques. High-performing students improved from 78.6% to 84.2%, indicating that question difficulty scaling prevents redundancy and promotes deeper understanding.

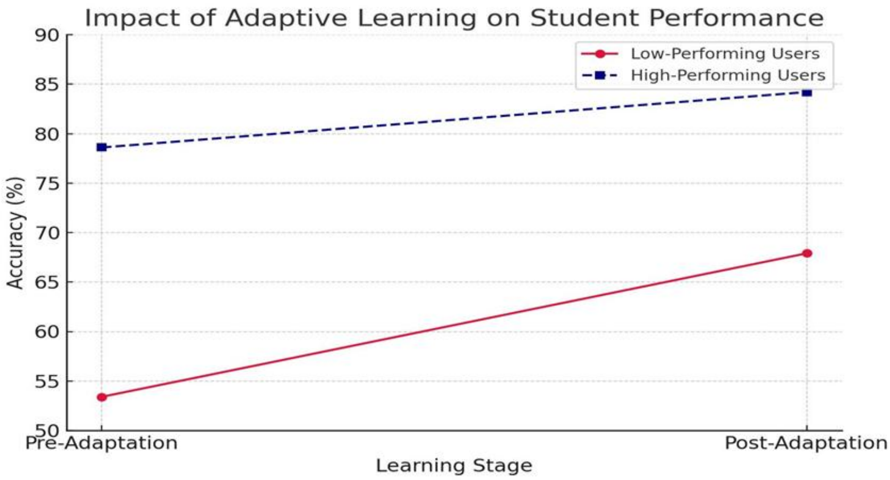

Figure 4 Adaptive Learning Performance

5.3 Query Processing Time Distribution (Efficiency of AI-Driven Question Generation)

This pie chart illustrates the time taken by different stages of the AI-driven question generation process: Query Preprocessing (18.7%) – Text cleaning, tokenization, and embedding conversion. Context Retrieval (30.7%) – Semantic similarity search in the vector database to fetch relevant learning materials. Question Generation (50.6%) – NLP-based processing, distractor generation, and AI model execution.

The total processing time (~7.5 seconds per query) ensures near-instant question generation, making the platform efficient for real-time assessments.

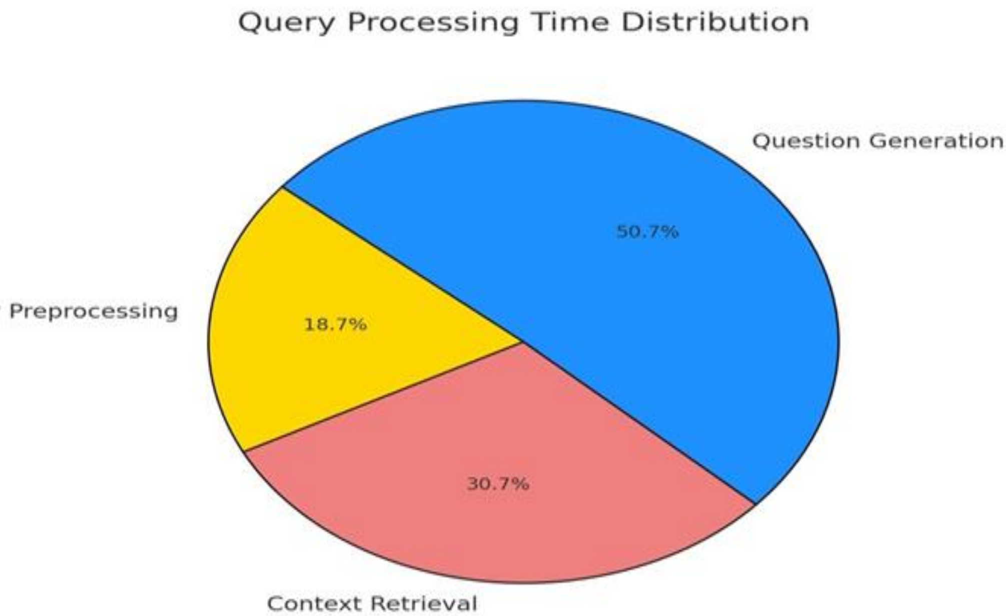

Figure 5 Query Processing Time Distribution

6. Security & Scalability Challenges

6.1 Security and Data Privacy Issues

Security is of the utmost priority in AI-powered learning systems. Our system has encryption methods like AES-256 for data at rest and TLS for data in transit to secure data privacy and data protection for uploaded user content. In addition, access control policies limit unauthorized changes to user input and derived questions. To mitigate manipulative adversarial attack threats, we use adversarial training techniques and anomaly detection models that identify possible abuse. Future work is to further enhance model robustness against prompt injection attacks and data poisoning attacks.

6.2 AI-created biased and unbiased questions

AI-generated questions must be free from inherent bias to enable an unbiased learning process to be provided. Bias can be created by skewed training data, leading to biased question difficulty or topic coverage. For this, we employ dataset curation methods to provide varied content coverage over a variety of domains. Also, algorithms that are fairness-aware such as counterfactual fairness and re-sampling algorithms minimize bias. Evaluation includes demographic fairness tests to ensure generated questions are appropriate for different learner populations.

6.3 Scalability Problems

Scalability is needed in order to keep performance across various sectors of education. Our system is cloud-based, using distributed computing and parallel processing to manage high volumes of user data.

Major scalability improvements are:

- Load balancing techniques to distribute query processing across many nodes.
- Caching techniques to remove redundant computation and improve response times.
- Modular microservices architecture for ease of expansion without system disruption.

6.4 Comparative Accuracy with Human-Generated Questions

To guarantee the effectiveness of AI-generated questions, we conduct comparative tests with human-generated counterparts. Comparison measures are:

- Contextual relevance: Making sure questions produced are relevant to given study material.
- Difficulty calibration: Refining question difficulty to match learners' abilities.

Flexibility: Evaluation to what extent AI-generated questions accommodate different learning styles. User studies and expert ratings provide quantitative feedback regarding the model's accuracy so that question generation is optimized continuously.

V. FUTURE SCOPE

The future scope of the Intelligent Integrated Platform for Customized Learning lies in enhancing its adaptability, scalability, and personalization to cater to a broader range of educational domains and learning styles. Future advancements will focus on multi-language support, enabling AI-driven question generation for diverse linguistic and regional curricula. Additionally, subject-specific fine-tuning will enhance its ability to handle technical fields such as mathematics, physics, and coding, where structured problem-solving questions are required. The integration of speech-to-text and OCR (Optical Character Recognition) will allow the system to process handwritten notes and audio lectures, further expanding its accessibility. Advanced reinforcement learning techniques will refine difficulty scaling and personalized recommendations, ensuring an adaptive learning experience tailored to individual student progress. Moreover, blockchain-based credentialing can be incorporated to secure AI-generated assessments and maintain integrity in e-learning and remote examinations. Seamless API integration with Learning Management Systems (LMS) will enable automated quiz generation, performance tracking, and AI-driven tutoring assistants. Future iterations will also explore explainable AI (XAI) techniques to enhance transparency in question selection and evaluation, ensuring fairness and reducing biases in AI-generated assessments. As AI in education continues to evolve, this platform has the potential to redefine personalized learning by offering an intelligent, automated, and self-improving question generation system, bridging the gap between traditional learning and AI-powered education.

References

[1] R. Mitkov, L. A. Ha, and N. Karamis, “A computer-aided environment for generating multiple-choice test items,” Natural Language Engineering, vol. 12, no. 2, pp. 177-194, 2006.

[2] X. Du, J. Shao, and C. Cardie, “Learning to ask: Neural question generation for reading comprehension,” in Proc. ACL 2017, pp. 1342–1352.

[3] C. Raffel et al., “Exploring the limits of transfer learning with a unified text-to-text transformer,” JMLR, vol. 21, no. 140, pp. 1-67, 2020.

[4] Y. Zhang and K. VanLehn, “Improving question generation with named entity recognition and part-of-speech tagging,” in Proc. ACL 2021, pp. 2108-2118.

[5] R. Kumar, P. Malhotra, and A. Singh, “Distractor generation for multiple-choice questions using word embeddings and cosine similarity,” in Proc. Educational Data Mining Conf., 2021.

[6] Y. Xu, L. Wang, and J. Li, “AI-driven adaptive learning platforms using deep reinforcement learning,” Journal of AI in Education, vol. 3, no. 1, pp. 25-39, 2022.

[7] M. Lewis et al., “BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension,” in Proc. ACL 2020, pp. 7871-7880.

[8] W. Zhang et al., “PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization,” in Proc. ICML 2020, pp. 11328–11339.

[9] J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proc. NAACL 2019, pp. 4171–4186.

[10] T. Brown et al., “Language models are few-shot learners,” in Proc. NeurIPS 2020, vol. 33, pp. 1877-1901.

[11] V. Rus, Z. Cai, and N. Niraula, “Automated question generation and beyond: The need for semantic-based approaches,” Int. J. Artif. Intell. Educ., vol. 24, no. 2, pp. 231-252, 2014.

[12] D. Litman and K. Forbes-Riley, “Predicting student emotions in computer-human tutoring dialogues,” in Proc. ACL 2006, pp. 352–359.

[13] P. Brusilovsky, “Adaptive hypermedia for education and training,” Adaptive Technologies for Training and Education, vol. 2, no. 4, pp. 21-44, 2012.

[14] V. Aleven, O. Popescu, and K. R. Koedinger, “Towards tutorial dialog to support self-explanation: Adding natural language understanding to a cognitive tutor,” Int. J. Artif. Intell. Educ., vol. 16, no. 2, pp. 105-134, 2006.

[15] L. von Ahn, “Duolingo: Learn a language for free while helping to translate the web,” in Proc. CHI 2013, pp. 878–879.

[16] H. Gao, X. Qin, R. J. D. Barroso, and W. Hussain, “Collaborative learning-based industrial IoT API recommendation for software-defined devices,” in Proc. AI 2020, pp. 55-67.