

AN INTEGRATED PLATFORM FOR ADAPTIVE EXAM PREPARATION AND DYNAMIC QUESTION GENERATION IN HIGHER EDUCATION

Project Reference No.: 49S_BE_2489

College: *PES University, Bengaluru*
Branch: *Department of Computer Science and Engineering*
Guide: *Dr. Manju*
Student(s) : *Ms. Roshna K Thomas*
Mr. Vedant Kamlesh Marathe
Mr. Parikshit Sharma
Ms. Sahana K Katenahalli

Keywords:

Adaptive Learning, Dynamic Question Generation, Academic Integrity, Personalized Education, Educational Technology, Natural Language Processing

Introduction:

1. Traditional assessment systems in higher education frequently suffer from several critical limitations that impede effective learning and evaluation. Conventional examinations often reuse questions from static repositories and fail to account for individual learning differences, resulting in decreased student engagement and increased opportunities for academic misconduct. Furthermore, existing question
2. generation approaches often lack coverage of diverse cognitive levels, as outlined in QG taxonomies.
3. Faculty members face substantial time investment in creating, updating, and maintaining diverse question banks while ensuring appropriate cognitive level distribution and academic integrity. Manual question authoring for multiple assessment cycles demands hours of effort per course, diverting time from teaching and research activities. Additionally, the predictability of recycled questions undermines assessment validity and creates opportunities for academic dishonesty that reactive detection systems cannot fully prevent.
4. This paper presents a comprehensive platform that addresses these challenges through automated, context-aware question generation. Unlike traditional Learning Management Systems that depend on static question repositories, the proposed system dynamically generates unique assessment items from uploaded course materials using fine-tuned language models. The platform employs Natural Language Processing techniques including document parsing, concept extraction via TF-IDF vectorization, and Low-Rank Adaptation for parameter-efficient model fine-tuning.
5. Each generated question is automatically classified according to Bloom's taxonomy using BERT-based classifiers, ensuring balanced cognitive coverage.

6. The system proactively enhances academic integrity by generating novel questions for each assessment cycle, eliminating predictability without relying solely on
7. post-hoc plagiarism detection. Faculty maintain pedagogical control through an intuitive review dashboard that enables rapid approval, editing, or rejection of generated items with full metadata visibility. By automating the labor-intensive aspects of assessment creation while preserving educator judgment, the platform substantially reduces faculty workload and supports scalable, secure academic evaluation.

Objectives:

1. To Mechanize the creation of questions enabling the swift production of high-caliber questions, with little human involvement
2. To prevent repetition and recycling of questions by using concept extraction, difficulty tuning, and template fallback mechanisms
3. To integrate gamification elements (XP, levels, streaks, achievements) to create interactive and engaging learning experiences for students
4. To support faculty by reducing manual workload and ensuring validated, accurate question sets ready for academic use

Methodology:

The methodology centers on creating an intelligent, modular system that automates question generation while maintaining faculty oversight and pedagogical rigor. The approach leverages Natural Language Processing techniques including document parsing, text summarization, concept extraction, and fine-tuned language model inference to transform course materials into diverse assessment items. Unlike prior work focusing solely on generation algorithms, this system provides an end-to-end platform integrating content ingestion, automated generation, cognitive

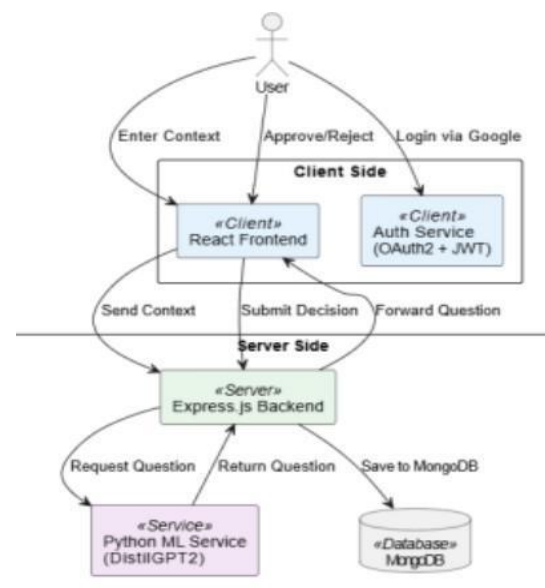


Figure 1: Project Architecture

classification, faculty validation, and secure export Workflows. The platform employs parameter-efficient fine-tuning through Low-Rank Adaptation (LoRA), enabling course-specific model adaptation without the computational overhead of full retraining. This approach allows the system to generate contextually relevant questions that reflect the specific terminology, concepts, and instructional focus of each course. Generated questions span multiple formats including multiple-choice, true/false, fill-in-the-blank, descriptive, numerical, and case-based items, each automatically tagged with cognitive level metadata according to Bloom's taxonomy.

Result and Conclusion:

1. The proposed platform was evaluated through a series of performance benchmarks, pilot deployments, and faculty feedback studies to assess its effectiveness in automated assessment generation.
2. The system demonstrates strong computational efficiency suitable for real-time educational environments. Question generation latency averages approximately 1.2 seconds per question on standard GPU hardware, enabling instructors to generate assessments quickly during preparation workflows. The BERT-based Bloom's taxonomy classifier achieved 87.3% accuracy, indicating reliable cognitive level classification of generated questions.
3. Faculty validation studies further highlight the system's practical usability. Approximately 91% of generated questions were accepted by instructors without major modifications, demonstrating the platform's ability to produce pedagogically meaningful assessment items. Time-motion analysis comparing manual question creation with AI-assisted generation showed a 65% reduction in assessment preparation time, reducing the average effort from 3.2 hours to 1.1 hours for a 10-question assessment.
4. The platform also ensures high answerability and structural quality of generated questions. The fine-tuned DistilGPT2 model achieved a training loss of 3.04 and evaluation loss of 2.91, indicating effective adaptation to educational content. Generated questions achieved a 95% answer correctness rate, while MCQ distractor quality reached approximately 70%, reflecting meaningful incorrect options based on common misconceptions.
5. In terms of assessment diversity, the system successfully generates multiple question types including multiple-choice, descriptive, true/false, and fill-in-the-blank questions. Cognitive distribution aligned with Bloom's taxonomy shows 50% Remember, 30% Understand, 10% Apply, 5% Analyze, 3% Evaluate, and 2% Create levels, ensuring coverage across different learning objectives. Difficulty distribution was maintained across Easy (50%), Moderate (30%), and Hard (20%) categories.
6. A pilot examination generated for course UE22CS352A (Machine Learning) demonstrated the system's ability to create balanced assessments containing a mix of MCQs and descriptive questions with automated difficulty calibration. Generated exams included structured metadata such as difficulty scores, cognitive levels, and estimated solving time, which assists instructors during review and validation.
7. The platform also integrates gamification features to improve student engagement. The system includes experience points (XP), streak tracking, leaderboards, and achievement badges. XP calculations occur in under 1 millisecond, and leaderboard synchronization operates asynchronously without affecting system responsiveness. These mechanisms encourage consistent participation and reinforce learning motivation.
8. Overall, the evaluation results demonstrate that the proposed system significantly improves assessment generation efficiency while maintaining

pedagogical validity. The combination of automated question generation, cognitive classification, and instructor validation workflows enables scalable and reliable academic assessment creation.

Future Scope:

The future scope of this project includes:

1. **Integration of Retrieval-Augmented Generation (RAG):** Incorporating RAG with dense vector databases can improve factual grounding by retrieving relevant course material during question generation, thereby reducing the chances of generating incorrect or ambiguous questions.
2. **Expansion of Training Dataset:** Increasing the size and diversity of the fine-tuning corpus with educational content from multiple disciplines will enhance the model's ability to generate high-quality questions for specialized subjects.
3. **Adaptive Difficulty Adjustment:** Future versions of the system can incorporate student performance analytics to dynamically adjust question difficulty and generate personalized assessments tailored to individual learning progress.
4. **Multimodal Content Processing:** Extending the system to process diagrams, charts, and graphical content using multimodal AI models would make the platform more suitable for STEM subjects that require visual assessment items.
5. **Enhanced Automated Grading:** Improving answer validation mechanisms and semantic similarity models can increase grading accuracy for descriptive questions and ensure reliable evaluation across different question types.
6. **Advanced Gamification Features:** Future improvements may include narrative-based quests, collaborative competitions, advanced achievement systems, and adaptive reward mechanisms to further increase student engagement.
7. **Cross-Domain Expansion:** The platform can be extended to support additional academic disciplines beyond computer science, enabling broader institutional adoption.
8. **Improved Model Training and Fine-Tuning:** Continuous model retraining using faculty feedback and newly generated assessment data can improve question quality and reduce rejection rates.