International Journal For Research In Advanced Computer Science And Engineering

GREEN PUBLICATION

ISSN (Online): 2208-2107
Volume 12 Issue 01 2026
Available on-line at: https://ijfres.com/index.php/es/about

DOI: https://doi.org/10.53555/es.v11i1.xyz Received: 10-Feb-2026 Accepted: 20-Mar-2026

AUTOMATED EDUCATIONAL ASSESSMENT THROUGH QUESTION GENERATION FROM PDF RESOURCES

Jaykumar Patel*[1], Jay Panchal[2], Kaushal Panchal[3], Tirth Shah[4]

*[1]*Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India, ORCID ID- 0000-0001-7972-5597*
[2]*Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India, ORCID ID- 0009-0004-8638-8735*
[3]*Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India, ORCID ID- 0009-0002-5978-4033*
[4]*Department of Computer Science, Gujarat University, Ahmedabad, Gujarat, India, ORCID ID- 0009-0009-5392-6687*

Abstract

Assessment creation has been a very slow and time-consuming process for teachers who have had to spend a lot of time making sure questions reflect the objectives of the course. In such rapidly changing academic environments, this not only drastically reduces the amount of time that can be used for teaching but also the quality and coverage of assessments may vary. Our research focuses on an automated exam generation revolution to cope with exiting problems and simplify the whole exam process right from the course materials. The platform allows a teacher to upload a single or multiple PDF documents that can be used to generate the question according to the type of questions teacher want to generate like mcqs, blanks. Since there are limitations in data extraction from PDFs, the system has some features that will help the user input the specific topic he is interested in. The teacher can input the topics, and our system will generate questions only from these topics to meet the teacher requirements for the questions. Our system uses Gemini API to processes the content, which is an intelligent material analyser that produces relevant and well-structured questions based on the input it receives. Once questions for the exam are created, the exam can be sent to the students through the system. Students get the test and do it in the time limit. They also get their results immediately, so they know how they have performed. Teachers can use the time and effort saved by the automatic grading alongside the visual analytics provided, for instance, donut chart and performance summary, which help the teachers to easily find top performers, the students in need of support, and the trends happening in the class. The system will automatically generate questions, align content to educator defined topics, provide immediate results on assessments, allow for quicker, more accurate, and efficient evaluation of all types of students, thus helping to improve student learning outcomes by eliminating the necessity of manual administrative work on behalf of teachers; therefore, it will provide intelligence, efficiency and adaptability to today's educational systems.

Keywords: *Automated Question Generation, Assessment System, PDF Content Extraction, Intelligent Evaluation, Educational Analytics*

1. Introduction

Evaluation is very important within the process of teaching and learning. Through evaluation, an educator is able to determine if students have understood the material being taught, measure whether or not specific learning objectives were achieved by the students, and make changes to instruction to improve outcomes. Unfortunately, creating assessments in accordance with the traditional model takes a lot of time and is an event that will happen often for educators. Educators spend a significant amount of time developing questions that fit with their course objectives, determining the question's difficulty level, geography and/or balanced coverage of the curriculum. In the fast paced world of education, the manual process of creating assessments decreases valuable time a teacher has available for instruction and may also create disparities in the quality of assessments between the teacher's and other teachers' assessments. [1][5].As digital education and online platforms continue to expand rapidly, so too does the need for intelligent systems that can provide an automated and standardized means of assessing students [2] [5] [9]. This study aims to create a smart, AI-driven assessment platform to fulfil this demand [8]. This system provides a way for educators to upload one or more PDF documents containing course material from which examinations automatically generate questions based on topics, types of questions (e.g. multiple-choice questions (MCQ), fill-in-the-blank, and true/false), and categories. Using an Intelligent Content Analyzer (Gemini API) to analyze text contextually and create relevant/structured/objective aligned questions (3) (6) (7), the system will generate questions, deliver assessments, evaluate automatically, produce instant results and provide visual performance analysis of the student's assessment. This will allow for a reduction in manual effort, increased accuracy in assessments and increase the efficiency of teaching, thus improving student learning results. This study has five sections. Literature associated with assessment through question generation will be described in Section 2; Section 3 covers the methodology related to the study and implementation; Section 4 discusses the development of the system and results/output of the System; Section 5 will conclude the research study with future possibilities.

2. Literature Review

The growing popularity of using digital formats for learning has driven the need for automated solutions to help with educational assessment. Educators are faced with many challenges, including the time and effort required to manually create a set of questions that are aligned with learning goals, cover the entire syllabus, and are appropriate for the difficulty level. Research shows that AQGs can help reduce the workload of the educator by providing timely, consistent, and high-quality assessment designs [1],[5].

Historically, early research on AQG systems focused on templated generating questions through either rule-based or template-driven approaches. These systems created a set of generated questions by applying a predefined set of linguistic patterns or syntactic rules to the source text. This method produced the initial level of automation for question generation but did not capture contextual meaning and/or semantic relationships of the education materials being used [4], [5]. Due to this, the generated questions lacked both diversity and contextual relevancy. Due to advances made recent years in natural language processing (NLP) and Machine Learning, researchers are increasingly employing data-driven methods to accomplish their goals. More specifically, neural networks and transformer architectures have been shown to provide enhanced performance when it comes to generating functional questions from text sources [6], [7]. Neural networks have made it possible for the system to examine factors of sentence structure and extract necessary elements for question production based on their relationship to other words as well as the overall context of the sentence. For example, Du et al. [8] provided empirical evidence that neural network question generation models trained using reading comprehension datasets produce quality, high-functioning questions, making them suitable for use in educational assessments. The latest advancements in large language models (LLM) have greatly improved the functionality of automated question generation systems. Models trained in advance such as BERT, as well as other transformer types of networks, are now able to recognize how contextual relationships exist in text documents and how to generate coherent question–answer pairs through the use of machine learning technology [11], [14]. Research into using LLMs for content creation as well as for automated assessment systems has also yielded results of greater grammatical correctness and semantic congruence in the resultant questions produced using the automated processes [3], [10], [15].One of the most significant avenues for furthering research regarding the integration of automated question generation and intelligent educational systems is through AI-powered tutoring and assessment systems which automatically generate questions, evaluate student responses, and provide performance analytics for instructors. The use of these types of systems reduces administrative burdens while also improving the personalization of learning experiences and developing data-based instructional strategies [9][12].Although improvements have been made in this area, there are still a number of limitations associated with existing systems. Many research studies being conducted assume that educational content will be available in a text-based format; however, teaching materials are typically provided in PDF format in academia. Due to the complexities of extracting meaningful textual content from PDF documents (the impact of formatting, embedded images, and complex structural forms), analysis of student responses to questions developed by these systems may be difficult or impossible to complete [13]. In addition, most of the contemporary automated question-generating methodologies are designed to work independently of exam delivery, automated evaluation, and performance analytics; rather, they create only the questions themselves.

As a result of these limitations, there is a growing need for a comprehensive system that integrates intelligent content extraction, automated question generation, online assessment delivery, and real-time performance analysis. The proposed system seeks to address these limitations by providing a fully integrated solution to generate a variety of

question types from educational materials in PDF format, including automated evaluation and analytical insights for educators. This type of system provides an opportunity to improve both the efficiency and effectiveness of modern educational assessment systems as well as the scalability of these systems.

3. Methodology

Using a modular and systematic approach, the proposed system will automate the entire assessments process from content ingestion to performance evaluation. This approach provides an overall workflow which allows for flexibility for educators while producing accurate and relevant questions.
The major stages of the system are:

3.1. PDF Upload and Content Extraction

- Educators upload a Course Time PDF containing course materials. The system obtains text from these PDFs, by extracting it through pdf processing algorithms, to enable them to generate questions for Tests[13].

3.2. Topic Specification

- Teachers provide the System with a specific topic or a specific keyword. The System will then only create Questions from the sections that match these words.

3.3. Intelligent Content Analysis

- The content retrieved will then be processed using the Gemini API to analyse its semantics and the Learning Purpose of the Content.

3.4. Question Generation

- The System has the ability to generate different Question types, based on the Teacher preference: e.g. MCQ, Fill in the Blanks, True/False.

3.5. Exam Delivery

- The Questions generated will be used to assemble a Test which will be assigned to the Student along with a Timer.

3.6. Automated Evaluation

- The Answers generated by the Student will be evaluated, in real-time (using pre-determined Answer Keys, and Logic).

3.7. Analytics and Reporting

- Results will be shown to Teachers in the form of Visual Dashboards & Performance Summaries.

4. Implementation and Result

4.1. GUI Screenshot

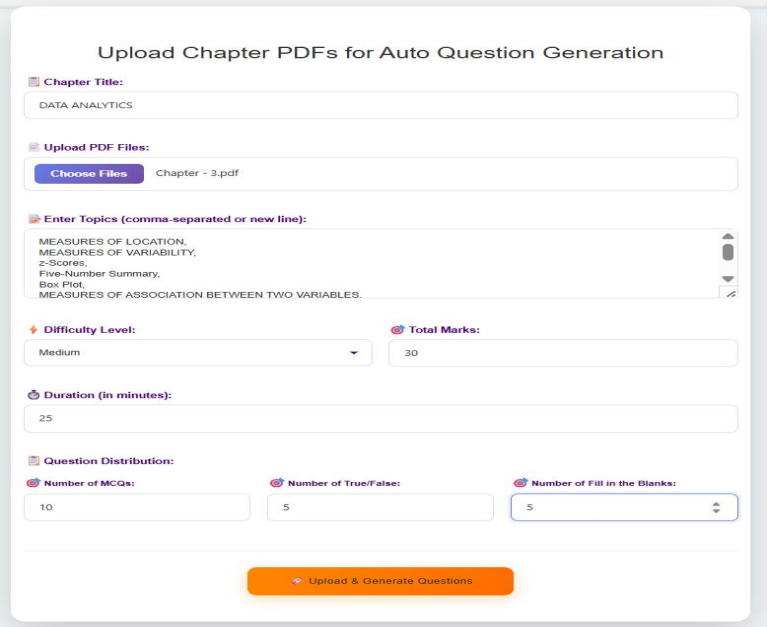
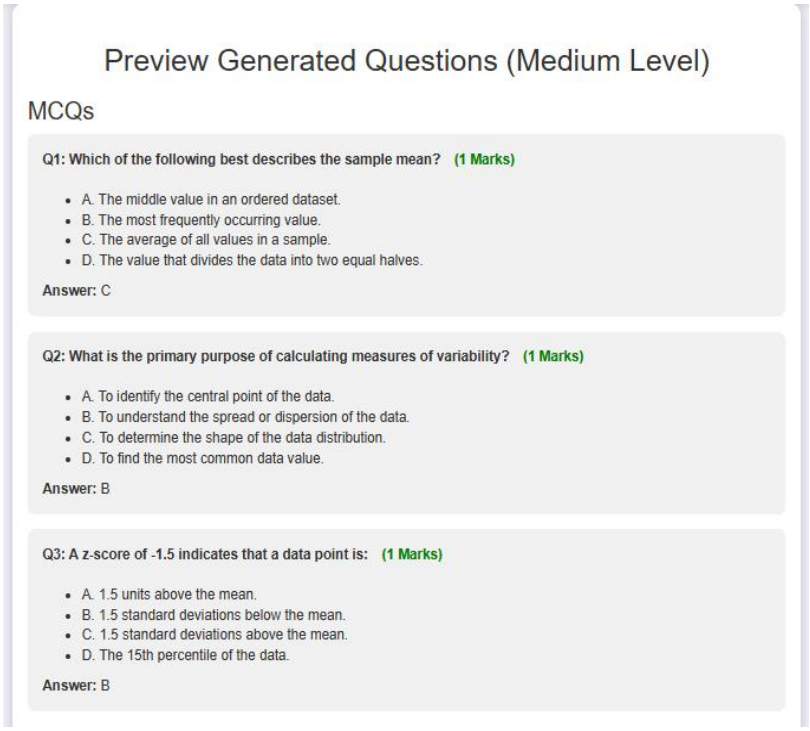


Figure 1. Auto Question Generation module

Figure 1 shows the Auto Question Generation Module. Here, the Teacher uploads the PDF Chapters, specifies the Topic, Level of Difficulty, Total Marks, Time Limit and Quantity of Questions (MCQ, True/False, or Fill-in-the-Blank). The System will then automatically generate the Assessment Question Paper, based on these parameters.



The Auto Question Generation module's Preview Generated Questions are shown on the screen in Figures 2 and 3. It shows questions that were automatically generated and sorted into MCQs, True/False, and Fill in the Blanks, along with the scores and the right answers. Teachers can check the questions before they save the quiz to make sure, they are correct and fit with the chosen level of difficulty and syllabus.

Figure 2. Preview Generated MCQ Questions

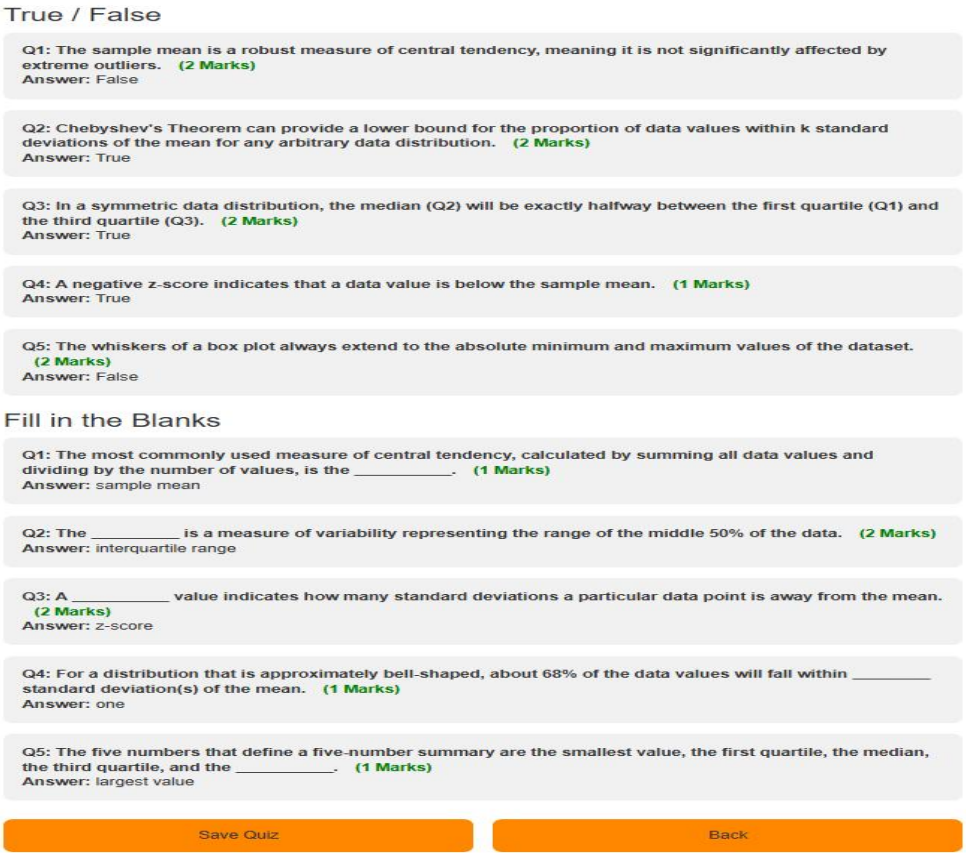


Figure 3. Preview Generated True/False and Fill in the blank Questions

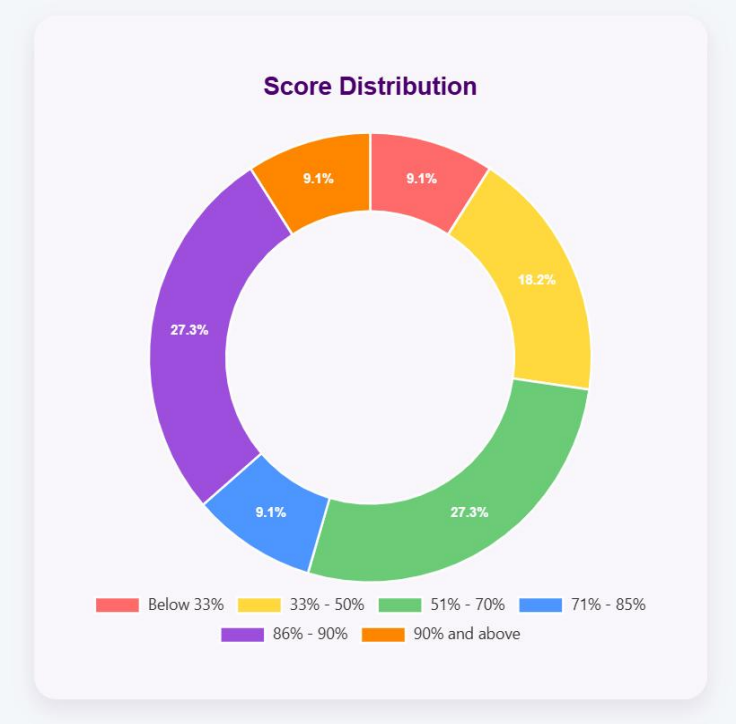


Figure 4. Score Distribution Analytics

Auto Question Generation module's Score Distribution Analytics can be seen in Figure 4 as a donut chart. It shows the different percentages of students within various ranges (e.g. Below 33%, 33-50%, 51-70%, 71-85%, 86-90%, or above 90%) allowing teachers to assess all students' achievement levels and evaluate areas where students need more assistance.

5. Result and Discussion

This project validated intelligent question generation and automated assessment for use with pdf-based academic materials. In comparison to creating exams manually, the time to create exams was greatly reduced, while still aligning with course objectives. In particular, the topic-based filtering process successfully produced relevant and focused questions. Furthermore, with the Gemini API, the generation of questions that were grammatically correct and well-structured resulted in multiple formats to be used for each exam. As a result, students were able to receive immediate results which increased the transparency of the assessment process and helped students with self-evaluation. For instructors, the analytics dashboard provided valuable insight into class performance so they can identify high performing students, low performing students, and class trends in overall learning. In summary, while the automated intelligent assessment system performs well when creating assessments from text-based information, there are still limitations associated with creating assessments from complex diagrams, equations, and heavily formatted PDFs. Even with these limitations, the results support the assertion that intelligent automation can significantly improve the efficiency, consistency, and instructional decision-making of assessment items.

5. Conclusion

This article offers an efficient and effective solution to the traditional assessment challenges we face today. Combining AI analysed content with a topic-based question creation process, the platform maximizes both question relevance and educational alignment. Automatic evaluation and immediate feedback greatly reduce the workload of the educator, while providing the student with timely information about their performance. Additionally, through the use of visual analytics, data-based decisions can be made for teaching. Although there are some limitations to the handling of complex PDF elements, this system has brought intelligence, efficiency, and adaptability to current education assessment processes. Future improvements will enhance the understanding of content and expand the capability of assessments.

References

[1] A. Das, M. Majumder, and S. Phadikar, “Automatic question generation and answer assessment: A survey,” Research and Practice in Technology Enhanced Learning, vol. 16, no. 1, pp. 1–31, 2021, doi: 10.1186/s41039-021-00151-1.

[2] S. Bhowmick, A. Basu, and S. Das, “Automating question generation from educational text using deep learning approaches,” in Proceedings of the International Conference on Artificial Intelligence in Education (AIED), 2020, pp. 33–44.

[3] Y. Wang, X. Li, and W. Zhao, “Automatic question–answer pair generation using pretrained language models,” Machine Learning with Applications, vol. 15, p. 100527, 2024, doi: 10.1016/j.mlwa.2024.100527.

[4] D. G. Bhatia and N. B. Patel, “Automated question generation using natural language processing techniques,” International Journal of Advanced Computer Science and Applications, vol. 13, no. 2, 2022.

[5] A. Kurdi, J. Leo, and K. Parsia, "A systematic review of automatic question generation for educational purposes," International Journal of Artificial Intelligence in Education, vol. 31, pp. 121–204, 2021.
[6] Z. Zhang and M. Lapata, "Automatic question generation from text using neural networks," Proceedings of ACL, 2021.
[7] T. Chen, S. Li, and H. Jin, "Transformer-based automatic question generation for educational content," IEEE Access, vol. 10, pp. 35214–35225, 2022.
[8] M. Du, X. Shao, and M. Cardie, "Learning to ask: Neural question generation for reading comprehension," Proceedings of ACL, 2021.
[9] P. Kumar and R. Gupta, "AI-driven intelligent tutoring and assessment systems in education," Education and Information Technologies, vol. 28, 2023.
[10] A. Brown, J. Smith, and L. Wang, "Large language models for automated assessment generation," Computers & Education: Artificial Intelligence, vol. 4, 2023.
[11] H. Li, Y. Chen, and Q. Liu, "Educational question generation using BERT-based models," Applied Sciences, vol. 12, no. 7, 2022.
[12] M. Al-Shehab and T. Al-Azmi, "AI-based automated exam generation systems: A review," IEEE Access, vol. 11, pp. 56789–56803, 2023.
[13] K. Mishra and A. Singh, "PDF text extraction techniques for intelligent document processing," Journal of Information Processing Systems, vol. 18, no. 5, 2022.
[14] J. Devlin et al., "BERT: Pre-training of deep bidirectional transformers for language understanding," Proceedings of NAACL, 2019.
[15] OpenAI, "GPT models for natural language understanding and generation," OpenAI Technical Report, 2023.