IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

An Intelligent Automated Assessment System for Dynamic Question Generation and Student Performance Analysis

[1]Sudarshan Date, [2]Om Makwana, [3]Srishti Pandey, [4]Prathmesh Sable, [5]Dr. Atul Kachare

[1,2,3,4] Undergraduate Student, [5] Assistant Professor

[1,2,3,4,5] Computer Engineering,

[1,2,3,4,5] Shah & Anchor Kutchhi Engineering College, Mumbai, India

Abstract: Manual preparation of question papers is often repetitive, time consuming and difficult to scale especially when personalization and fairness is required. This paper proposes a dynamic question paper generation system using transformer based natural language processing along with semantic similarity analysis for the automation and improvement of the assessment process. The system is designed to generate multiple choice questions from existing content while taking into consideration topic coverage, difficulty level and cognitive categories based on Bloom's taxonomy. We apply semantic similarity methods to identify and remove duplicate or close questions, to ensure quality and diversity of the generated papers. The system also randomizes the question sets for each student, reducing the repetition and increasing the fairness of the assessments. It also has a performance analysis module which analyses the responses of the students and points out the weaker areas for improvement. The overall approach lessens the manual effort required from educators while supporting more adaptive and data driven assessment practices. The results indicate that the system can generate consistent and reliable question papers with relevance and diversity.

Index Terms - Dynamic Question Generation, Natural Language Processing, Transformer Models, Semantic Similarity, Bloom's Taxonomy, Automated Assessment, Adaptive Learning.

I. INTRODUCTION

Assessment is a vital section of the learning process because it assists in gauging the level of student comprehension of a subject and where their understanding is lacking. Nevertheless, manual preparation of question papers may be a tedious and time consuming process to teachers. This is even more of a problem when there is a requirement to design multiple sets of question papers, maintain fairness and make sure that topics and difficulty levels were adequately covered. Traditional approaches are also not flexible and able to adapt to various needs of students learning.

As digital education and artificial intelligence develop, the idea of automating academic work is gaining more and more attention. It is especially through Natural Language Processing (NLP) that it has become possible to effectively analyze and generate human language. The recent developments in transformer based models have further enhanced the capability of systems to produce meaningful and context aware text. This means that with these developments, it is now possible to automatically create questions based on the existing study material.

Nevertheless, the majority of the existing systems deal with only one aspect of the issue. There are tools which can create questions, however, they do not take into account such factors as the level of difficulty or duplication. Conversely, most internet-based assessment software provide test generation and assessment, but heavily depend on manually generated question banks. This indicates that there is an eminent

requirement to ensure a system that integrates intelligent question generation with structured assessment features.

This paper will present a proposed system of dynamic question paper generation that employs transformer based NLP methods as well as semantic similarity analysis. The system then creates multiple choice questions depending on the topics and level of difficulty selected and also ensures that the questions are aligned to the taxonomy of Bloom. Semantic similarity is employed to eliminate duplicated or similar questions in order to maintain quality. This system also gives a randomized set of questions to each student, which helps enhance fairness in assessments.

Moreover, the system has a performance analysis element which analyzes the student response and determines the weak points. It will enable more specific feedback and will facilitate adaptive learning. In general, the suggested strategy should have the following impacts: reduced manual workload on educators and enhanced quality, flexibility, and effectiveness of the assessment process.

The rest of this paper will be organized as follows. Section II presents the literature review and discusses existing approaches. Section III describes the proposed methodology and system design. Section IV provides the implementation and outcomes. Lastly, Section V of the paper wraps up the paper and provides recommendations on future research.

II. LITERATURE SURVEY

The recent innovations in Natural Language Processing (NLP) and the deep learning have played a significant role in the development of automated question generation systems. The earlier studies were mainly centered on rule-based systems, but more recent systems are based on neural networks and transformer-based architectures to produce meaningful and context-sensitive questions.

It was presented by Xinya Du, Junru Shao, and Claire Cardie (2017) in which a sequence-to-sequence learning framework was proposed to generate questions out of reading comprehension passages. Their model proved that encoder-decoder architectures are effective in generating the relevant questions, which can serve as a baseline in future studies. On the same note, Qingyu Zhou et al. (2017) investigated neural question generation on text with attention mechanisms and found it important to represent context in question generation to enhance quality of questions.

Moreover, Xingdi Yuan et al. (2017) introduced another improvement to the field in the form of text-to-text neural model of machines, which can understand and answer questions. This work helped to shift attention towards comprehension and generation tasks, in addition to making systems better prepared for understanding of the input contents before generating questions. Similarly, Luis Enrico Lopez et al. (2020) have utilized transformers to build language models to generate paragraph-level queries and showed that this model outperformed other previous models in terms of fluency and situational accuracy.

In addition, researchers have worked to integrate the concept of knowledge and reasoning into question generation. For instance, Lisa Bauer et al. (2018) focused on commonsense reasoning during multi-hop question answering processes to enable the development of more complex and meaningful questions. Similarly, Huang L. et al. (2018) proposed a method of using knowledge in order to generate question and answer, and used structured information to make the process relevant and accurate.

While research on neural models is still ongoing, some studies have been working on hybrid methods that integrate rule-based techniques along with machine learning methods. Thus, for example, Kulkarni A. et al. (2019) developed a framework for automatic MCQ generation using rules along with machine learning. The authors highlighted the importance of structured questions in education. Similarly, Zhou Y. et al. (2020) applied machine learning to create fill-in-the-blanks and validated the usefulness of the automatic assessment generation process.

There is also an increase in recent research focusing more on personalization and adaptive learning. One of these studies is the one conducted by Wang J. et al. (2020), who suggested a framework using deep learning techniques for question generation where questions are generated based on the learner's profile. Another study by Gkatzia D. et al. (2020) explored different dialogue generation models for online learning, emphasizing interactive and adaptive learning systems. Additionally, another study done by Yu X. et al. (2021) has provided effective strategies for creating concise answers for integration within question generation and answering systems.

Despite all these advancements, several limitations exist. First, most of the existing frameworks focus on generating questions rather than other challenges such as avoiding duplicates, balancing difficulty level, and integrating assessments into the assessment process. Secondly, some of the frameworks lack diversity in question creation, whereas others do not have any analysis or feedback of the learners' performance. In addition, the majority of the frameworks are standalone rather than incorporating them into a comprehensive platform that manages the entire assessment process.

These limitations will be addressed through a framework incorporating transformer-based question generation and semantic similarity analysis techniques to ensure unique and quality questions. It also integrates other features such as randomization of tests, difficulty balance through Bloom taxonomy, and performance analysis among others.

III. Research Gap

- The first major limitation that has been found in the course of this research study is that of the inability of systems to take inputs in the form of unstructured documents and subsequently generate appropriate questions from such unstructured inputs. Most of the approaches currently in use require processed or structured inputs, thereby limiting their practical application.
- Another very important limitation lies in the inability of the existing systems to generate questions based on different levels of cognition. These systems fail to make use of Bloom's Taxonomy, which is essential to vary the complexity of questions.
- Moreover, existing solutions are not very helpful in terms of deriving answers out of the source material and comparing them in an effective way. This decreases the general effectiveness of automated assessment systems and restricts their capability to offer valuable feedback.
- Additionally, it has not fully integrated platforms that integrate document processing, question generation, evaluation, and analytics into one platform. The majority of the tools available are standalone modules as opposed to full-blown web-based tools.
- To fill these gaps, the proposed system is modeled as a single web application of document input, generating questions of various levels of difficulty, evaluation of answers, and analytical results, which can enhance efficiency and usability.

IV. System Design

The proposed Dynamic Question Paper Generation system is designed as a modular and scalable architecture that integrates Natural Language Processing (NLP), transformer-based models, and assessment management components. The system follows a layered approach, consisting of input processing, question generation, post-processing, and evaluation modules, all connected through a centralized database.

A. Overall Architecture

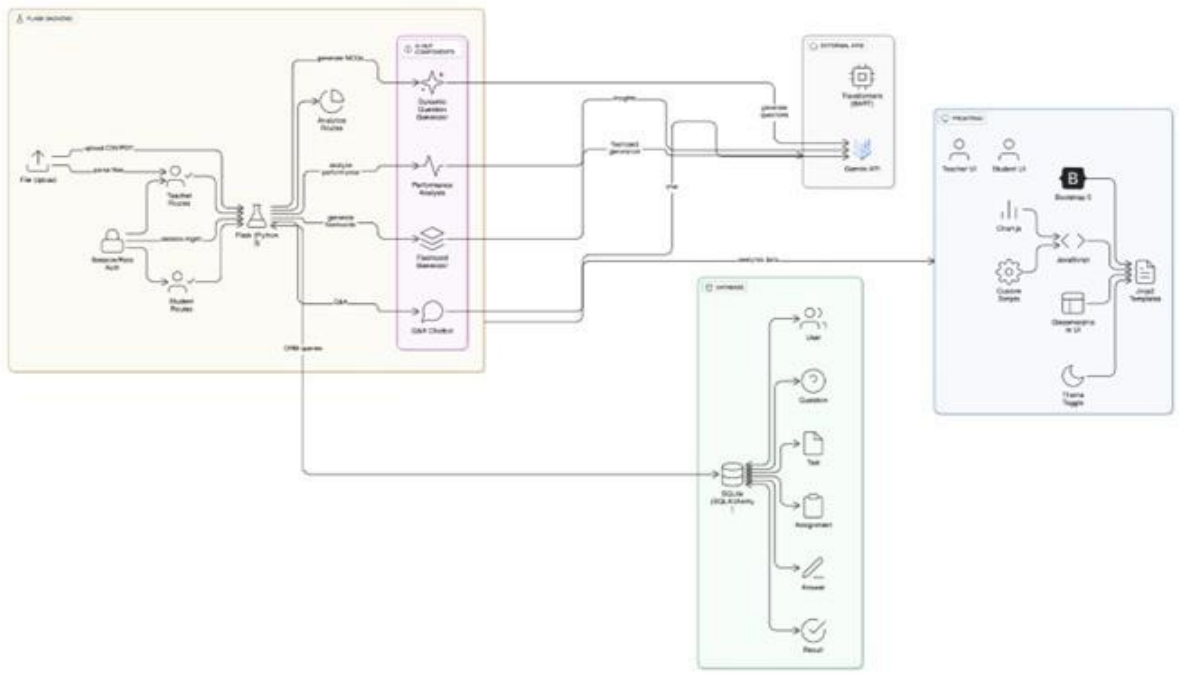

Fig. 4.1 System Architecture of the Proposed Dynamic Question Paper Generation System

This system is designed through four major layers:

Input Layer – This layer will take input from users in the form of documents (PDF, DOCX, TXT) and the selection of topics or themes.

- Processing Layer – This layer will make use of NLP processes to extract texts, clean them, and then identify important concepts in them.
- Question Generation Layer – This layer is used to create questions based on the extracted text and pre-defined parameters through transformer-based models.
- Output and Evaluation Layer – This layer creates the final questions, stores them in a database and also helps generate tests and evaluates their performances.

This sort of design will ensure that the system has the capacity for flexibility, scalability and can easily integrate with other AI models in the future.

B. Input/Text Processing Module

The first step involves getting input files uploaded by users. These input files are processed in the following manner through NLP-based methods:

- Text Extraction: Extract the raw text contained in input files.
- Text Cleaning: Removing stop words, punctuations and other unwanted characters.
- Sentence Segmentation: Sentence segmentation of extracted text.
- Concept Identification: Identification of important concepts using NLP.

This module plays an important role in ensuring only relevant and structured data is passed onto the question generation layer.

C. Question Generation Module

The main function of the system is question generation and is accomplished through transformer-based NLP models used to generate questions depending on the context.

- It makes use of pre-trained models like BART, T5, or any other model that can be used for the transformation of textual data into question-answer pairs.
- Also, the system can support various kinds of questions; however, multiple-choice questions (MCQs) are more of interest to the system.
- As far as questions' types go, they are generated based on interrogative patterns, including who, what, when, where, why, and how, as well as application and conceptual questions.

There are also control parameters that affect the process of question generation:

- Levels of Bloom taxonomy: knowledge, comprehension, application, analysis, evaluation, and creation.
- Difficulty level: easy, medium, and hard.
- Question count: number of questions per topic.

This approach will ensure a diversified type of questions and their alignment with academic standards.

D. Validation Module

After the question is generated, a series of actions aimed at increasing its quality will take place:

- Generation of options: generation of several options in case of a multiple-choice question.
- Correct answer selection: selection of an appropriate answer from the source material.
- Similarity analysis: elimination of duplicate and extremely similar questions using semantic similarity measures based on embeddings.
- Quality check: final validation for the sake of checking grammar relevance.

This component is crucial for ensuring the quality of the questions generated.

E. Question Bank, Database.

All the questions that have been tested and proven valid will be stored in one database that is structured to form a question bank. The metadata attached to each question includes but not limited to the following:

- Topic
- Level of difficulty
- Level according to Bloom's taxonomy
- Correct answers and alternatives.

Such systematic storing ensures that questions can be easily retrieved and reused when necessary for future assessments.

F. Test Generation and Usage Module

The system allows test generation using the existing question bank through the following processes:

- Random Question Selection: This will involve random selection of questions per individual student.
- Question Assignment: Tests will then be administered depending on the selected parameters.
- Analysis of performance: Performance analysis will be done to establish students' weakness and accuracy in answering questions.

Such a module ensures equity in evaluation and promotes adaptive learning.

G. Workflow Overview

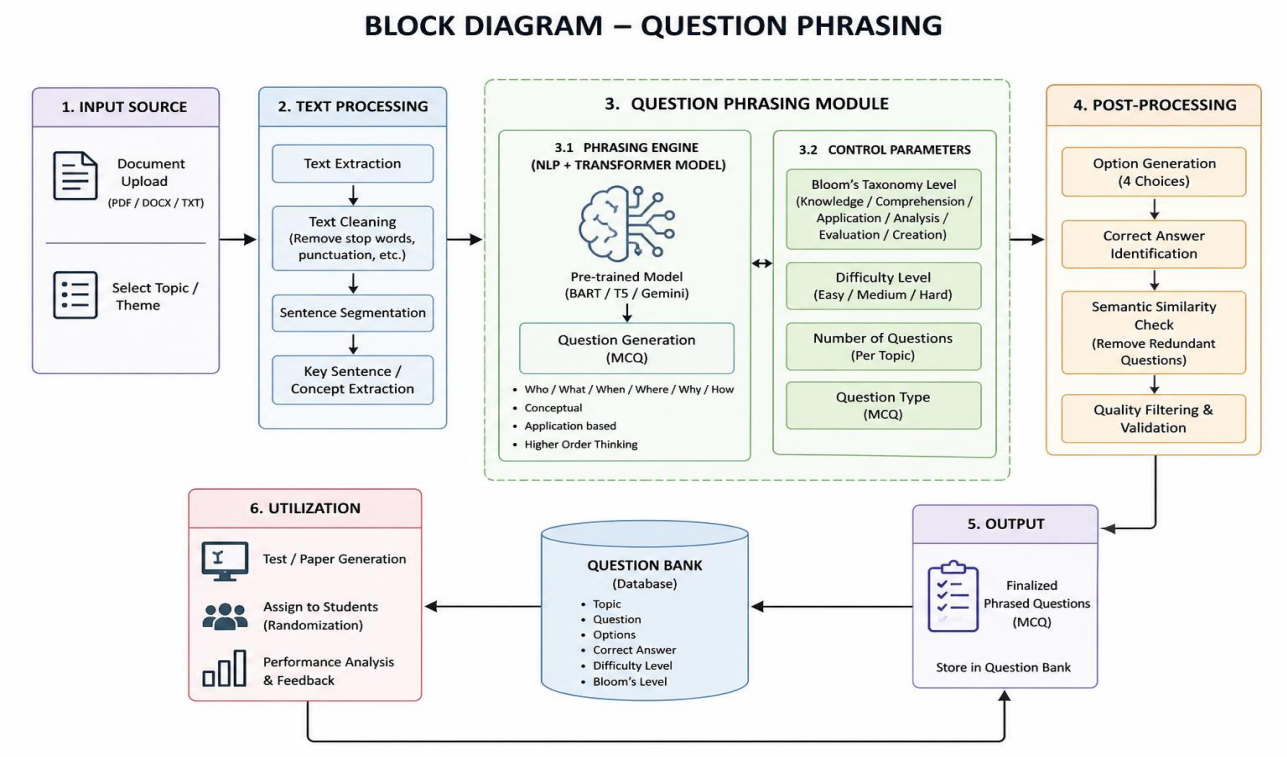

Fig. 4.2 Workflow of the Proposed Question Generation System

The overall system workflow can be outlined as follows:

1. A document or topic is uploaded by the user.
2. Extraction of text and preprocessing using NLP techniques.
3. Use of transformer-based models to generate questions.
4. Post-processing for ensuring quality and eliminating redundancy.

5. Storage of questions in the database.
6. Evaluation of students' responses.

V. IMPLEMENTATION AND RESULT

The proposed DQPG system has been implemented as a web application that incorporates the student assessment system. The system is a blend of NLP techniques, transformer model-based approaches, and web technology to provide a complete automation solution for generating questions, conducting tests, and analyzing student performance.

The data processing and AI-based generation performed on the backend side and the user interaction experience provided by the frontend are represented in the architecture of the proposed system. There is also a module aimed at students to take tests and analyze their performance.

A. Backend Implementation

The backend side of the proposed DQPG system has been implemented using Python programming language and Flask framework. RESTful API calls were developed using Flask to enable interaction between frontend and core components of the backend.

The backend uses NLP techniques provided by SpaCy library and transformer models to perform text preprocessing, concept identification, and question generation. Another task of the backend is to perform semantic similarity filtering to remove duplicated or redundant questions.

Additionally, the backend has business logic responsibilities such as:

- Question generation based on topic and difficulties.
- Blooms' taxonomy classification.
- Random assignment of tests to students.
- Analytic and performance evaluation.

B. Frontend Implementation

The frontend was implemented to ensure the user-friendly interface and responsiveness to user actions by using HTML, CSS, and JavaScript languages.

For the instructor, the interface provides:

- Documents uploading (PDF, DOCX, TXT).
- Selection of subjects, difficulties, and question types.
- Questions creation and editing.
- Tests creation and distribution among students.

For students, another test interface has been implemented, which consists of:

- Timed testing environment
- Single question navigation
- Automatic response recording
- Results presentation and submission.

C. Database Implementation

Relational database (MySQL/SQLite) is used for storing and managing the information. The database will consist of structured tables such as the following:

- users (teachers and students)

- questions (with additional meta information regarding topics, difficulty, bloom's level etc.)
- tests and assignments
- student feedback and outcomes.

Design will make possible the effective storing, retrieving and scalability of question bank and assessment information.

D. Question Generation and Download Module

System contains module which generates and exports questions in several formats. Once questions are created and validated, they can be downloaded in PDF, CSV, or Excel files.

This module interacts with the backend APIs in order to retrieve necessary data and transform it to the desired format, making question sets easy to share and reuse.

E. Student Evaluation and Analytics Module.

Specific student module is designed to take the test and analyze the performance of students. Such a module provides students the possibility to:

- take part in testing,
- each of the students will receive random sets of questions,
- real time response tracking,
- automatic scoring of responses.

Once the students complete their tests, system will be used to analyze students' performance by:

- calculating the accuracy and scores of each of them,
- identifying the weakest parts – topics or questions;
- analyzing students' answers;
- making conclusions based on responses.

Such a module contributes positively to the system as it makes possible adaptive learning and personal development of each learner.

F. Integration of Bloom's Taxonomy

Taxonomy of bloom is used by the system in order to classify and generate questions of the specified cognitive level. Cognitive domain consists of the following levels of thinking:

- remembering
- understanding
- applying
- analyzing
- evaluating
- creating

The system is capable of generating various levels of difficulties to evaluate students effectively. Moreover, it allows teachers to create appropriate tests.

G. Output

The system takes input information, processes it and creates unique subjective and objective type questions as an output. Generated questions become unique as a result of semantic filtering. They are then stored in the question bank to be used later.

VI. RESULTS AND DISCUSSION

In this section, the implementation findings of the designed system that includes both the web application of the teachers and the student's mobile application have been discussed. According to the results, the functionality of different modules has been proven, such as the question generating module, test making, and performance evaluation.

A. Web Application Interface (Teacher Side)

The web application would serve as the interface for instructors to control different features related to the process, such as question generating, test making, and grading. Teacher's dashboard has access to many modules, including assignments, tests, and analysis.

The dynamic question generator allows the instructor to generate different questions based on topics chosen and difficulty levels desired. Additionally, there are other functionalities that can be enhanced by AI, for example, the chat and revision functions to facilitate understanding the information.

Another feature included in the developed system is related to creating tests and controlling them while giving tests to students and monitoring their progress. This is made possible with the help of student's dashboard (web interface).

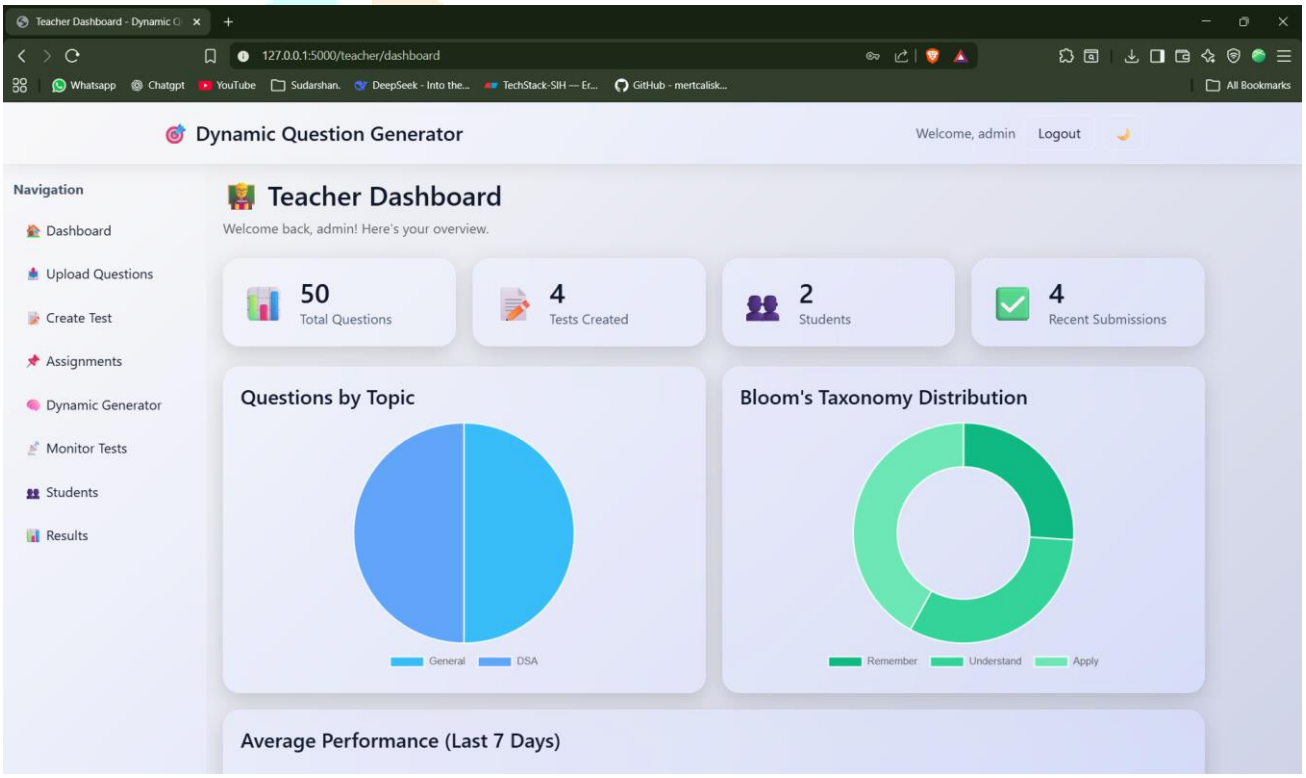

Fig. 6.1 Student's Dashboard

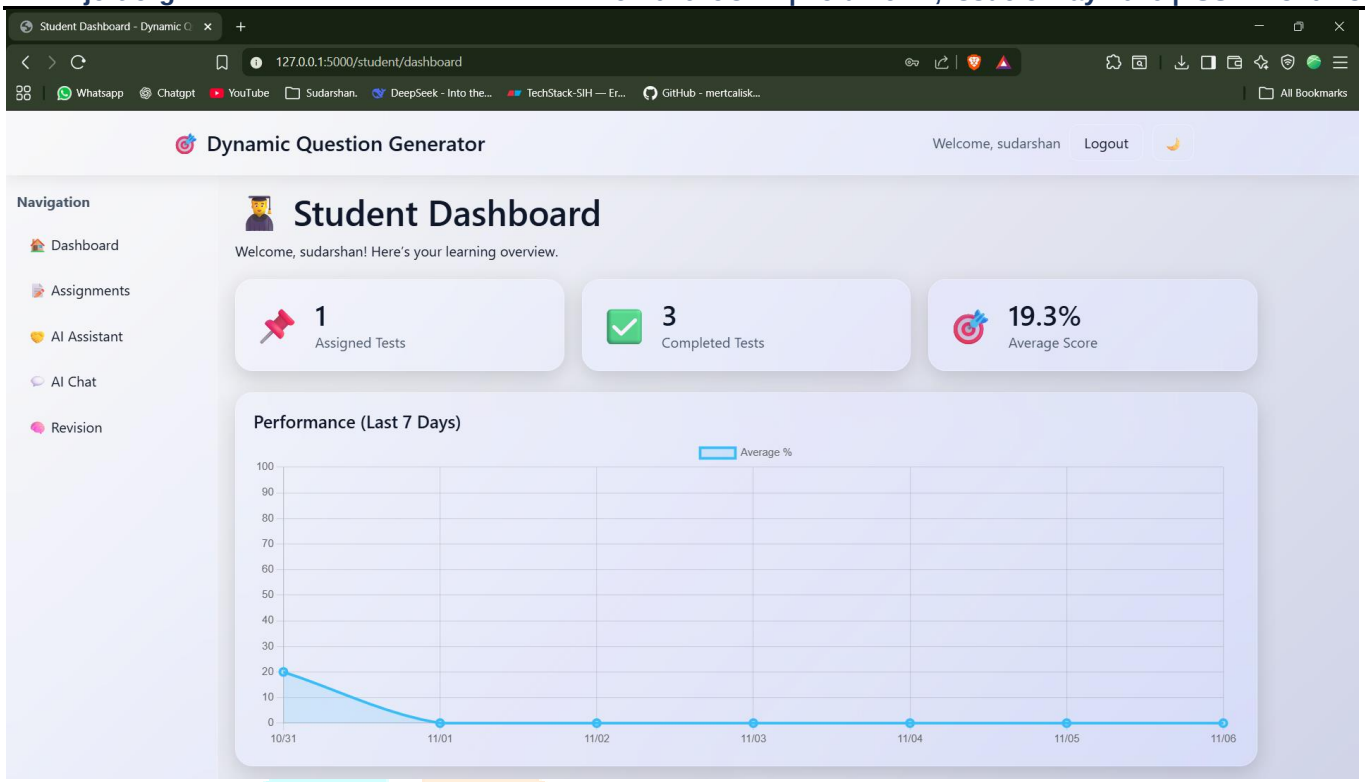

Fig. 6.2 Teacher's Dashboard

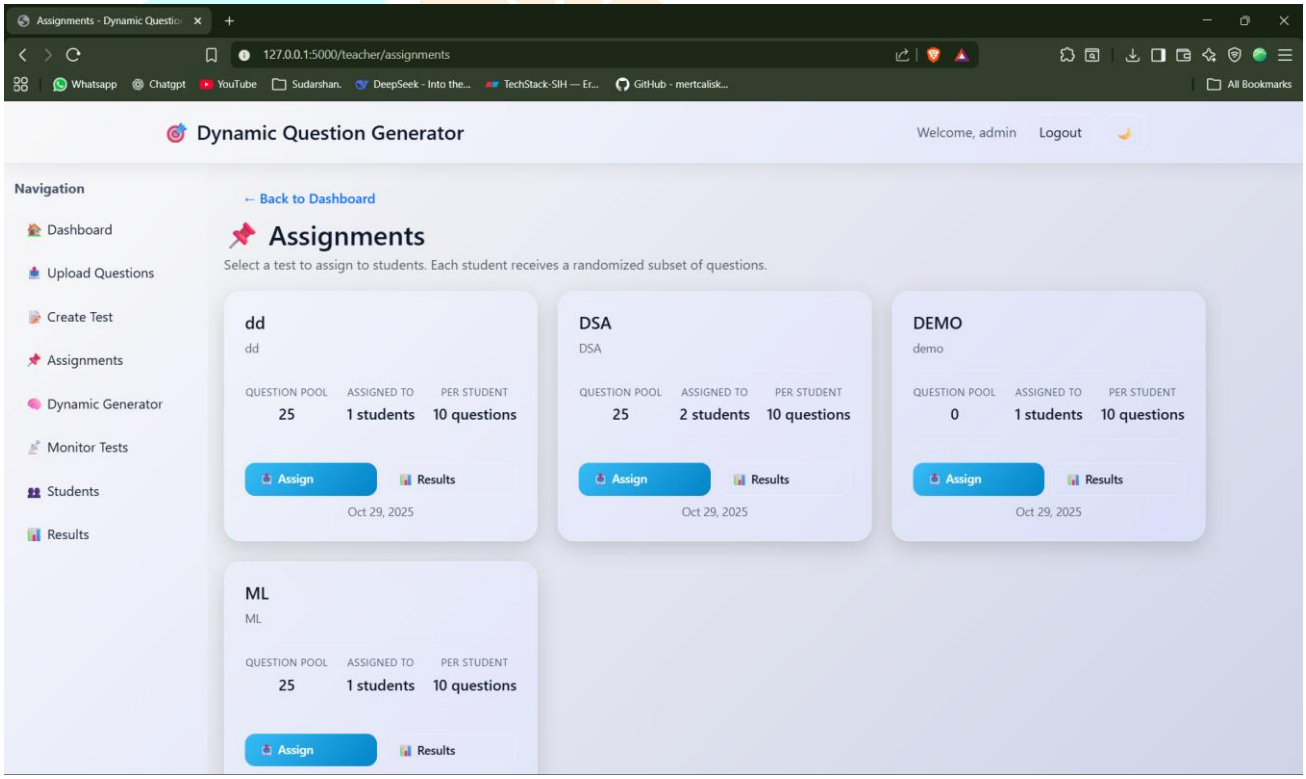

Fig. 6.3 Assignment Management Dashboard

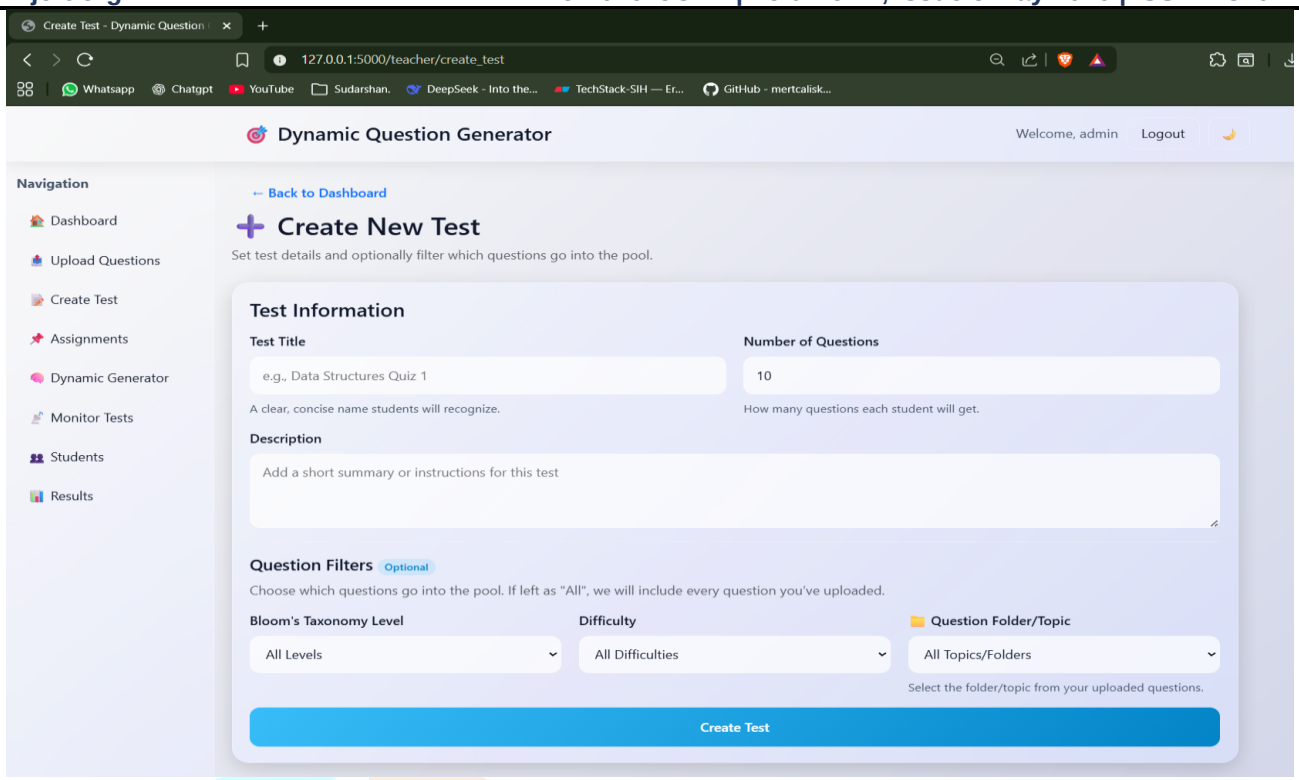

Fig. 6.4 Test Creation Interface

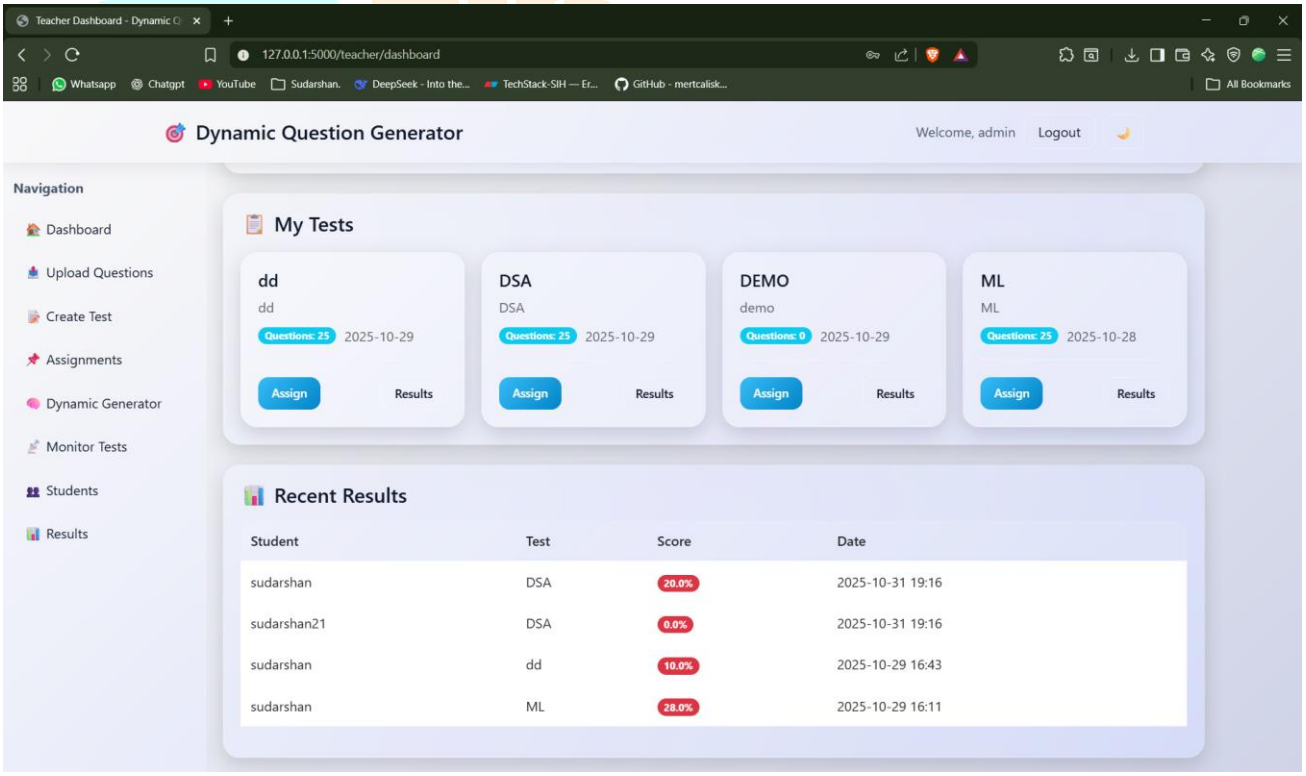

Fig. 6.5 Test List and Assignment Module

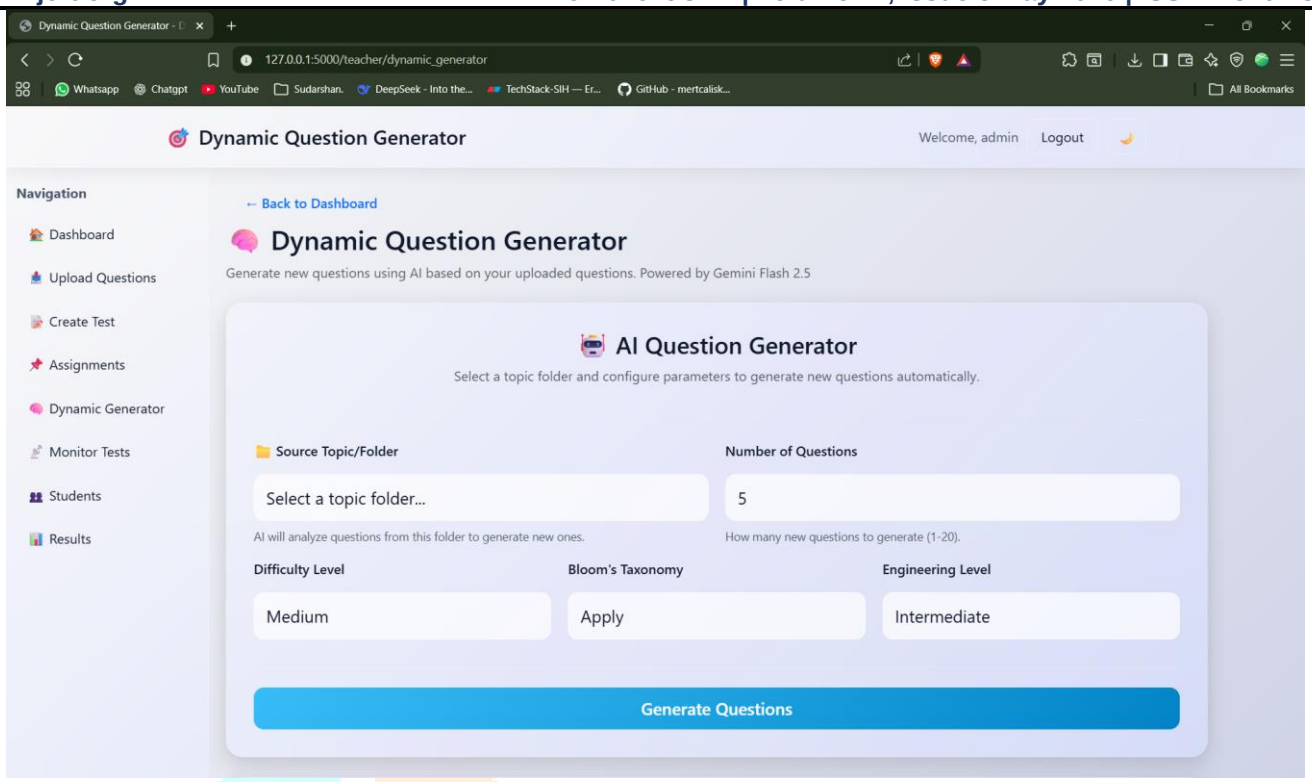

Fig. 6.6 Dynamic Question Generator Module

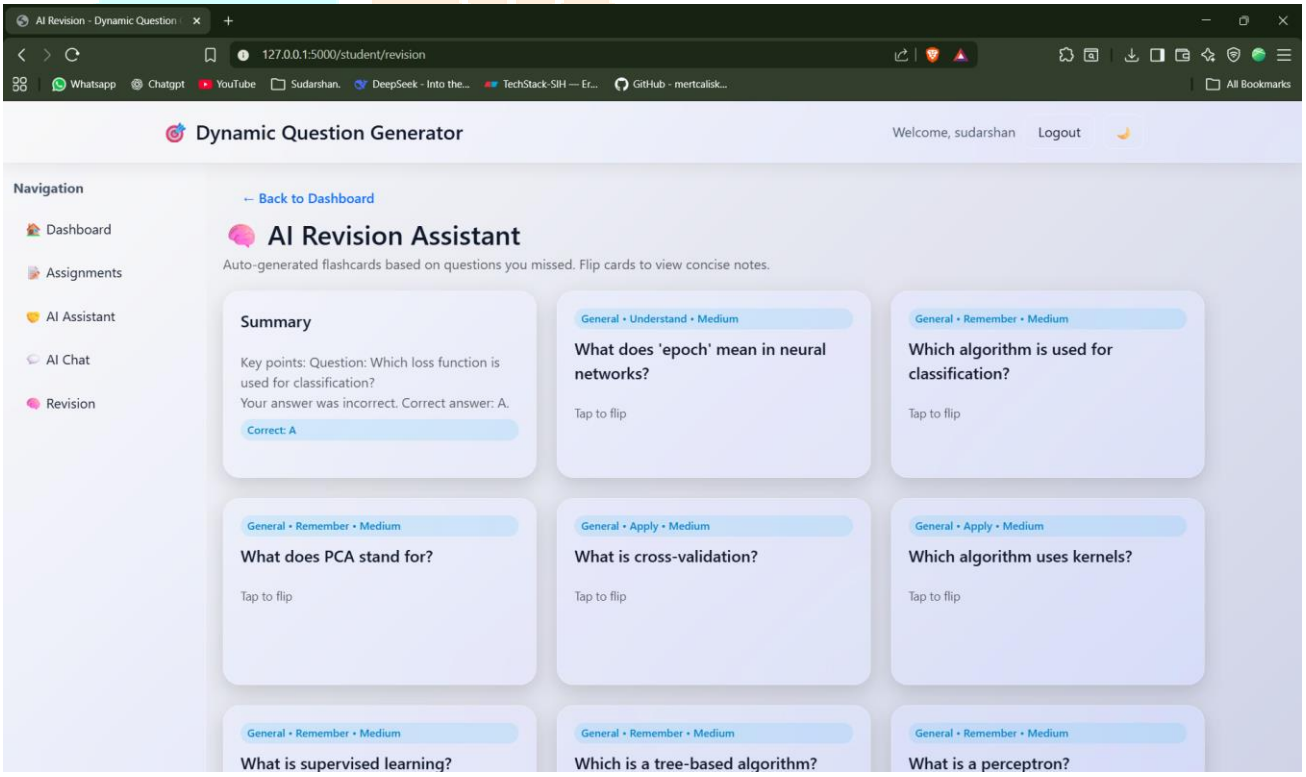

Fig. 6.7 AI-Based Revision Module

A. Mobile Application Interface (Student Side)

In order to improve access and usability, a mobile application was created to enable students to interact with the system. The mobile application will offer such features as secure log-in, profile management, and entrance to assigned tests and study materials.

Students are able to see their tasks, access learning materials and practice the generated questions using a mobile interface. Combination of study content and question formulation assists in maintaining the learning process that is not necessarily evaluated through formal means.

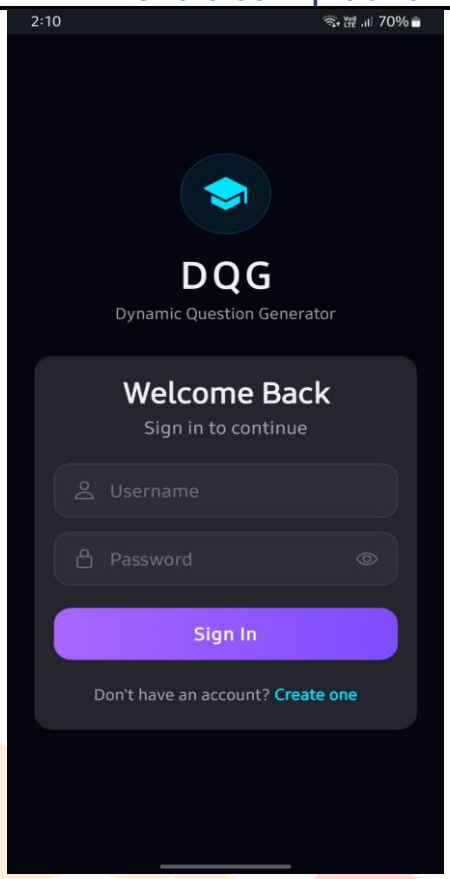

Fig. 6.8 Student Mobile Application Login Screen

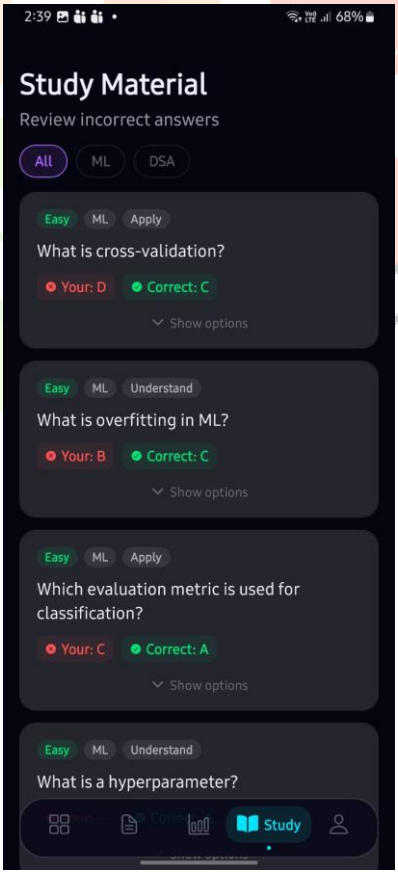

Fig. 6.9 Student Profile Interface

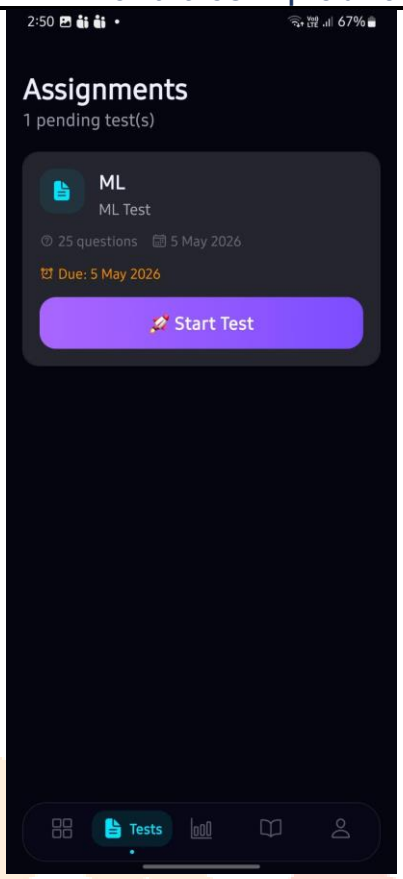

Fig. 6.10 Assignment List in Mobile Application

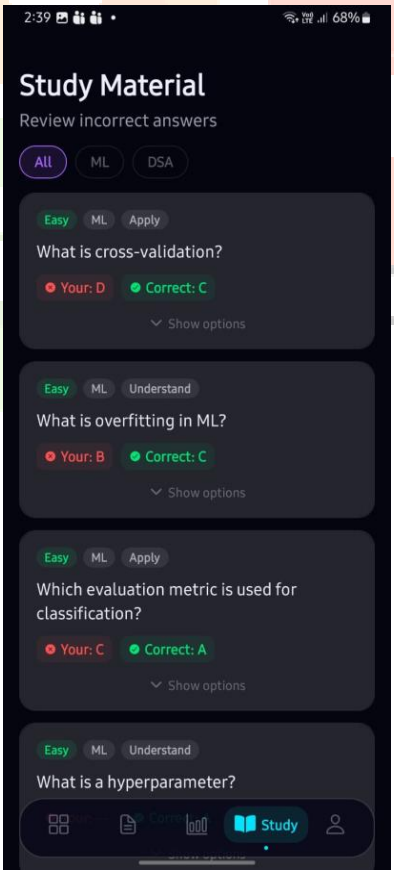

Fig. 6.11 Study Material Access Module

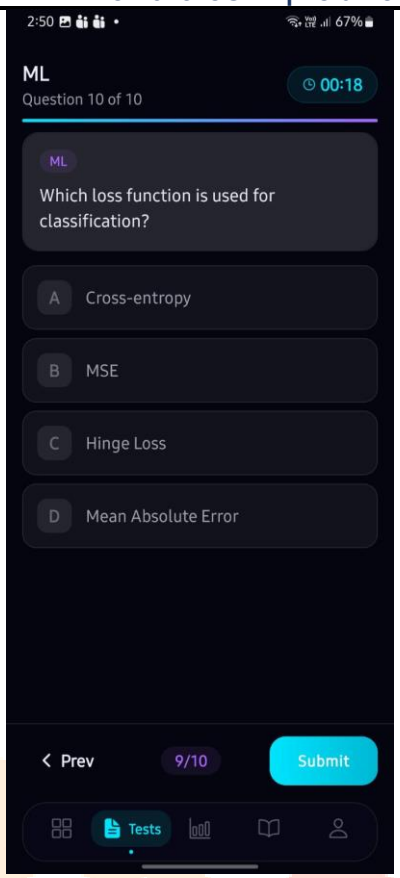

Fig. 6.12 Generated Questions in Mobile Application

B. Test Execution and Results

The system offers an interactive test environment whereby students are able to attempt assessments in an organized manner. A set of randomized questions is given to each student in order to achieve a fair and less duplicated result.

Once the tests have been submitted, the system will automatically assess the answers and produce performance scores. Detailed insights are given by the result module, such as accuracy and topic-wise performance to enable students to understand their strengths and weaknesses.

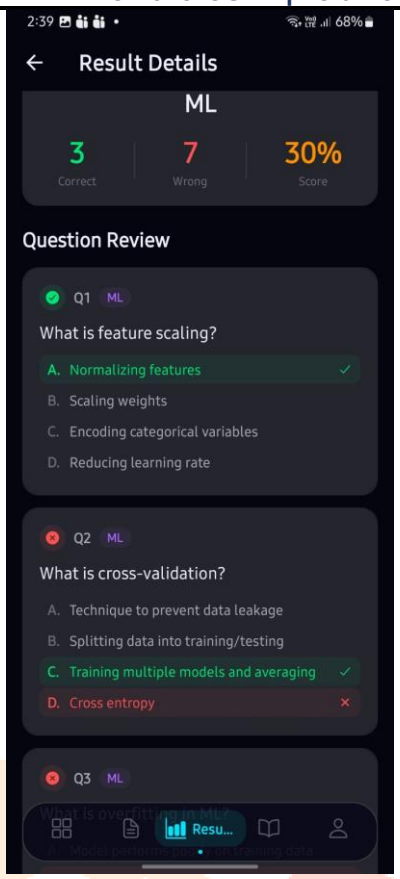

Fig. 16. Test Interface for Student Assessment

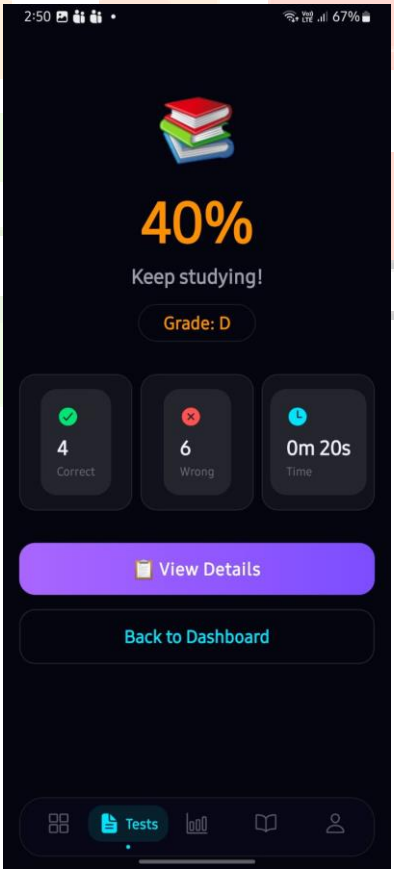

Fig. 17. Student Performance and Result Analysis

VII. CONCLUSION

This paper has suggested and applied a dynamic question paper generation system to overcome the weaknesses of the old manual assessment system. The system integrates the Natural Language Processing methods and transformer-based methods to automatically generate meaningful and diverse questions given documents. The system by incorporating the use of semantic similarity analysis will ensure that the questions that are generated by the system are unique and are not redundant.

The system is an all-encompassing assessment solution in addition to question generation, which is offered by its web and mobile applications. The web platform will enable instructors to create tests, manage assignments, and view the results of their students in a convenient way, whereas the mobile application will allow students to access the study materials, attempt tests, and review results in a convenient manner. The added features like the question classification based on the taxonomy of Bloom, the generation of randomized tests, and the performance analytics contribute to the quality and effectiveness of assessments.

In general, the presented system minimizes the manual labor that educators need, and enhances the effectiveness of the assessment process. Simultaneously, it helps the students by making the learning process more interactive and adaptive. The findings indicate that the system can produce the appropriate questions, handle the assessments efficiently, and provide the useable information on the student performance.

The paper has emphasized the possibilities of introducing artificial intelligence in the education systems to develop scalable, intelligent and user friendly assessment platforms.

REFERENCES

[1] J. Doe, “The impact of machine learning on healthcare,” Journal of Medical Infor matics, vol. 15, pp. 234–245, July 2021.

[2] “Ieee guide for software requirements specifications,” IEEE Std 830-1984, pp. 1–26, 1984.

[3] N. Reimers and I. Gurevych, “Sentence-transformers: Multilingual sentence em-beddings using bert and transformers,” arXiv preprint arXiv:1908.10084, 2019.

[4] Luis Enrico Lopez, Diane Kathryn Cruz, Jan Christian Blaise Cruz, Charibeth Cheng(2020): Simplifying Paragraph-level Question Generation via Transformer Language Models.
https://paperswithcode.com/author/luis-enrico-lopez

[5] Xinya Du, Junru Shao, Claire Cardie(2017): Learning to Ask: Neural Question Generation for Reading Comprehension.
https://paperswithcode.com/search?q=author%3AClaire+Cardie

[6] Qingyu Zhou, Nan Yang, Furu Wei, Chuanqi Tan, Hangbo Bao, Ming Zhou (2017): Neural Question Generation from Text: A Preliminary Study
https://paperswithcode.com/author/furu-wei

[7] Liu, C., Sun, X., Hu, D., & Liu, X. (2020). Automatic Question generation Using Neural Networks.
https://doi.org/10.1007/s12652-019-01524-2

[8] Zhou, Y., Liu, X., & Gu, J. (2020). A Machine Learning-Based Approach for Generating Fill-in-the-Blank Type Questions
https://www.aclweb.org/anthology/C18-2020/

[9] Kulkarni, A., Singh, A., & Chakraborty, T. (2019). Automatic Generation of Multiple Choice Questions Using a Combination of Rule-Based and Machine Learning Techniques.
https://ieeexplore.ieee.org/document/8672953