

## An Intelligent AI-Based Proctored Examination Framework with Automated Question Generation and Performance Analytics

Kamini Vinod Salunkhe<sup>1</sup>, Kamlesh Vilas Nikam<sup>2</sup>, Vaishnavi Dinesh Desale<sup>3</sup>, Riddhi Kishor Patil<sup>4</sup>

<sup>1,2,3,4</sup>Department of Information Technology Engineering, SVKM's Institute of Technology, Dhule, India.

<sup>1</sup>kaminisalunkhe52@gmail.com, <sup>2</sup>kamleshyt123@gmail.com, <sup>3</sup>vaishnavidesale04@gmail.com, <sup>4</sup>patilriddhi28@gmail.com

<p><b>Peer Review Information</b></p> <p><i>Type: Article</i> <i>Received: 27 March 2026</i> <i>Revised: 12 April 2026</i> <i>Accepted: 26 May 2026</i> <i>Published: 16 June 2026</i></p>	<p style="text-align: center;"><b>Abstract</b></p> <p>The current systems of on-line examination which help with competitive examinations have major flaws such as inadequate backup support of databases for questions, lack of categorization of difficulty levels of questions, inefficient general performance measures, and ineffective or incomplete monitoring systems. These flaws undermine the validity of the tests, and hence, make it difficult for them to evaluate candidates' readiness. This study proposes a model of intelligent and AI-based proctored examinations which includes question management in real-time, question categorization based on difficulty levels, and overall performance assessment. The device consists of AI-driven proctoring techniques to reveal candidate conduct throughout examinations and find out capability malpractices, thereby improving examination integrity. Additionally, superior analytics are hired to generate exact standard performance insights, permitting purpose evaluation of learner development. The proposed framework ambitions to decorate the fairness, reliability, and effectiveness of online competitive examination exams and can be extended to massive-scale examination systems past MPSC education.</p> <p><b>Keywords:</b> Advanced Monitoring Mechanisms; AI-Proctored Platform; Fairness in Evaluation; MPSC Exam Preparation; Performance Analytics; Question Bank Security; Reliable Assessment.</p>
--	--

### How to Cite This Article

Salunkhe, K. V., Nikam, K. V., Desale, V. D., & Patil, R. K. (2026). An intelligent AI-based proctored examination framework with automated question generation and performance analytics. *Multidisciplinary Journal of Research in Engineering and Technology*, 13(2s), 224–232.

## Introduction

One of the most aggressive use-degree examinations in India is the Maharashtra Public Service Commission (MPSC) examination, which demands a highly systematic technique of training with the help of constant practice and precision of evaluation. The aspirants strongly depend on the internet-based examination frameworks to self-estimate and develop their performance as well. Nonetheless, numerous online arrangements in existence are not made to advantageous resource settled daily acquiring information of our dependable evaluation techniques [1].

The most crucial impediment of existent structures that is the most significant is the existence of unorganized question repositories. The majority of question banks are not categorized rightfully with difficulty ranges of smooth, medium and tough, it is difficult when the aspirants investigate their development and benefit knowledge of curve. According to existing literature on automated question paper generation, a well set and trouble-tagged query repositories are important in achieving balanced evaluation [2][3]. The combination of techniques and Genetic Algorithms and rule-based selection strategies was proved to be able to enhance the allocation of questions and maintain appropriate trouble ranges during checks [4].

Lack of strong proctoring and solid query coping with mechanisms in on line exam structures is also another critical issue[5][6]. Recent research on AI-based proctoring designs that are familiar with the notion that automatic conduct that monitors the usage of synthetic intelligence is able to enhance integrity in examinations through detecting suspicious sporting practices and preventing malpractices[7][8].

Such obstacles highlight the necessity of complete and intelligent analysis framework that merges the stable query management, AI-initiated and inspired strong proctoring, and the development of the overall performance analytics This researches dreams to address those challenges through the means of introducing an AI-dominant proctored exam system that assists connected learning, robust tests, and fact-driven overall performance evaluation of aggressive tests comprising of MPSC [9].

## Literature Survey

The speedy digitalization of schooling and competitive examinations has introduced approximately the giant adoption of on-line examination systems and automated assessment platforms [10][11]. While these structures provide scalability and luxury, numerous studies challenges live unresolved, in ensuring exam integrity, automatic query generation, problem balancing, and huge-scale deployment. This segment critiques gift studies in the domain names of automated query paper era, herbal language processing (NLP)-primarily based query generation, big language model (LLM)-based totally structures, and sensible examination frameworks, with a focus on identifying barriers relevant to competitive examinations [12].

### *Automated Generation of Question Papers*

Several studies have focused on automating the manner of exam query paper generation to lessen human strive and get rid of bias. Tengku et al. Proposed an automated exam question generator the use of Genetic Algorithms, emphasizing balanced query distribution primarily based on Bloom's Taxonomy stages Their approach demonstrated progressed manipulate over difficulty distribution in assessment to random and backtracking-based techniques [13][14]. However, the machine is based closely on pre-tagged query banks and lacks the functionality to extract questions dynamically from unstructured assets which includes PDF documents.

A machine based totally on semantically tagged query repositories, allowing query paper generation using user-defined constraints consisting of issue degree, cognitive class, and subject matter insurance. Although the framework presents flexibility, it assumes manually or semi-routinely tagged repositories and does not scale efficaciously to big competitive exam datasets, which include preceding-year MPSC question papers [15][16]. More current approaches employ deep gaining knowledge of techniques The Q Gen tool suggested by Paul et al. Applied a first rate-tuned T5 transformer model to produce questions to PDF files uploaded Generation of Questions using the method of LLM and NLP.

### *Intelligent and Hybrid AI-Based Question Generation Systems*

Hybrid AI fashions have also been tried in order to adorn the efficiency of automatic query paper generation. A hybrid MLP-CNN-based implementation of predictive evaluation version was proposed by Pol et al. to operate on automatic question paper era, which performed better on standard performance than the independent CNN and MLP models. Nonetheless, the method requires extensive preprocessing and vast amounts of categorized data hence real-global execution is computationally excessively costly[17].

Pisat et al. Developed automated query paper generator and answer checker the application of NLP and quick-text matching strategy. Although the gadget computerized evaluation, it relied on random question choice and no trouble control adaptive, AI-based completely proctoring, and intelligent monitoring functions like crucial to excessive-stakes testing [18].

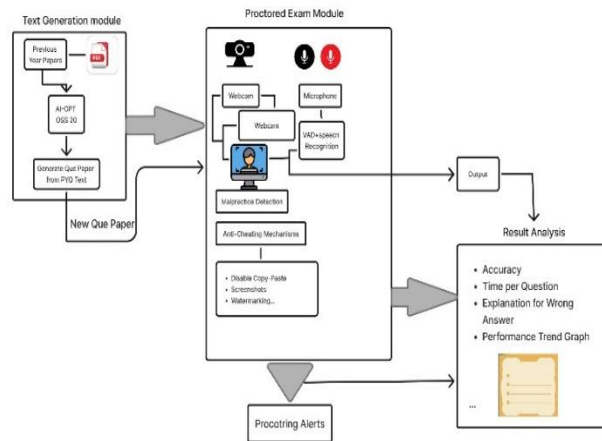
Author Name & Year	Technique Used	Limitations
A. Sharma et al., 2022	Large Language Model (LLM)-based Framework	Development of a web-based intelligent examination system using large language models for automated question handling
S. Hrvaevic et al., 2022	LLM with Knowledge Base Integration	Context-aware question generation using structured knowledge sources and language models
P. Paul et al., 2021	Transformer-Based Model (T5 – QGen)	Automated question generation from textual and PDF-based input documents
S. Pol et al., 2021	Hybrid MLP-CNN Model	Predictive evaluation and optimization of automated question paper generation
R. Sangeetha et al., 2020	NLP-Based Intelligent System (BERT/T5)	Automatic generation of grammatically and semantically meaningful questions

*Existing System and Research Limitations*

Based on the literature reviewed, it is quite clear that the majority of current day structures awareness primarily focused on either the question generation or assessment and never considered the safe implementation of on-line tests. Moreover, although the application of LLM-based totally fashions beautifies query great, automated troubles type, MCQ faults of PDFs and vicinity-uniqueness analysis technology (e.g., MPSC) remain under-researched. These gaps address the urge to have a single framework of AI-driven examination, which combines stable proctoring, intelligent query era, and high-quality general performance assessment [19][20].

**Methodology**

The proposed tool would utilize the module-based structure and layer-based structure in the layout and implementation of an intelligent AI-based totally smartly proctored examination system for the development of more reliable, safe, and effective virtual competitive examination practices. As depicted in Figure 1, the proposed approach would include the integration of consistent consumer authentication, AI-based totally proctoring, query creation, query type, examination creation, and overall performance analytics into an integrated framework. For the effective examination practices, the examination process would be segmented into modules for the consumer end as well as the administrator end.



**Fig. 1.** Proposed System.

*User authentication and exam initialization*

Consistent person registration and authentication using email-based credentials marked the beginning of the examination practice. You would be redirected to an individual dashboard that monitors the examination-related practices once the authentication is successful. You would need to ensure that you have access to your webcam and microphone before the examination begins. This would ensure that the AI-based practices are enabled before the examination begins. Consistent examination commands would be displayed before the consumer is redirected to the live examination environment.

### *AI-Based Video Proctoring Mechanism*

The video proctoring module constantly monitors the frames received from the candidate's webcam throughout the exam consultation. The machine detects irregularities such as multiple faces, no candidate's face, or movement outside the camera frame. Once irregularities are recognized, the system generates automated warnings and terminates the exam session. The AI-based totally enforcement machine is an effective way of enhancing the exam system's integrity and minimizing the need for human invigilators.

### *Audio Monitoring and Voice Activity Detection*

In order to enhance the precision of the mostly invigilation machine using the machine learning algorithm, another audio-based invigilation machine forms part of the system. Audio signals are captured by the candidate's microphone and analyzed through the Voice Activity Detection (VAD). This audio signal is also processed through speech recognition modeling in order to differentiate between background noise and speech. The system counts the number of violations due to repeated irregularities on voice activity. When the number of violations goes beyond the threshold set, the system ends the test.

### *Question Generation from PDF Files*

Regarding the administrator's work, the device will allow users to upload previous year MPSC MCQs in PDF form. Text extraction techniques will be used for the analysis of the document and retrieval of unprocessed text. The unprocessed text will then be analysed and partitioned to obtain insight about questions and answers. The process of question generation does not require guide data and allows creating huge question pools..

### *AI-Based Difficulty Classification of Questions*

The extracted MCQs go through AI-powered language fashions, such as GPT-OSS-20, to categorize them as Easy, Medium, or Hard questions. The classification is achieved based on linguistic difficulty, semantic difficulty, conceptual difficulty, and structural analysis of questions. The AI-based difficulty labeling of questions facilitates balanced take a look at generation, as well as flexible workout take a look at generation, without relying on manual labeling of questions.

### *Dynamic Practice Exam Generation and Evaluation*

Classified questions are saved in a database, which is dynamically decided on to generate check examinations. The device ensures uniform distribution of difficulty tiers in every take a look at. Parameters like solution accuracy, time taken in keeping with question, and crowning glory time are recorded at some point of the exam. Once the exam is finished, automatic evaluation takes region, and specific consequences are generated, including marks, accuracy, time evaluation, etc.

### *Result Analytics and Performance Visualization*

This is the closing phase of the method, wherein the emphasis is given to overall performance analytics and visualization. The device collects ancient exam records to generate trend-based totally insights into candidate overall performance. Analytics is presented in the form of accuracy trends, time management styles, and difficulty-stage strengths and weaknesses. The insights assist aspirants in figuring out regions of development, thereby facilitating better guidance techniques.

## **Experiments and Evaluation**

This section describes the experimental setup, assessment method, and overall performance evaluation of the proposed AI-based proctored examination system. The experiments are conducted to assess the effectiveness of automatic questions extracted, issue categories, AI-based proctoring, and performance analytics in comparison with existing automated examination systems mentioned in the literature.

### *Experimental Setup*

The experimental assessment was conducted with the use of a dataset comprising previous-12 months MPSC multiple-choice query (MCQ) papers in PDF format. These documents were uploaded into the gadget and processed the usage of the automated query extraction module. The extracted questions had been labelled into trouble tiers the usage of an AI-primarily based language model (GPT-OSS-20).

The proctoring module turned into evaluating the usage of actual-time webcam and microphone inputs from candidates for the duration of mock examinations. Performance analytics generated the usage of exam response logs and historic take a look at statistics

- Difficulty degrees: Easy, Medium, Hard
- Proctoring modalities: Video + Audio

### *Evaluation Metrics*

#### Question Extraction Accuracy

The accuracy of automated MCQ extraction from documents is evaluated using precision and recall metrics.

$$Precision = \frac{TP}{TP+FP} \quad (\text{eq 1})$$

$$Recall = \frac{TP}{TP+FN} \quad (\text{eq 2})$$

where:

The F1-score is computed as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

A confusion matrix  $CM$  is used to analyse mis-classification among difficulty levels:

#### Proctoring Effectiveness Evaluation

##### Video Proctoring Detection Rate

The face detection accuracy during examinations is evaluated as:

$$Accuracy_v = \frac{F_{valid}}{F_{total}}$$

##### Audio Proctoring Violation Detection

Accuracy of voice activity detection is expressed as:

$$Accuracy_a = \frac{V_{detected}}{V_{actual}}$$

#### Overall Proctoring Effectiveness

$$P_{eff} = \lambda Accuracy_v + (1 - \lambda) Accuracy_a$$

#### Examination Performance Evaluation Metrics

##### Candidate Accuracy

$$Accuracy_c = \frac{C}{N}$$

##### Time Efficiency

$$Time_{eff} = \frac{T_{total}}{N}$$

##### Weighted Score Calculation

$$Score = \sum_{i=1}^N w_i \cdot r_i$$

#### Performance Trend Analysis

Performance scores that show learning progression through various tests are recorded as:

$$S = \{s_1, s_2, \dots, s_n\}$$

The improvement trend is determined by calculating linear regression slope:

$$Trend = \frac{n \sum i s_i - \sum i \sum s_i}{n \sum i^2 - (\sum i)^2}$$

Positive slope represents improvement, while negative slope shows performance deterioration.

#### Comparison of the System with Other Similar Existing Systems

Comparison between the proposed system and other existing automated examination systems is conducted with respect to functionality and efficiency.

$$Gain = \frac{Metric_{proposed} - Metric_{existing}}{Metric_{existing}} \times 100$$

exam honesty using AI-powered proctoring technology and more insightful performance analysis compared to those based only on random question generation or proctoring techniques.

#### Question Extraction Outcomes

The accuracy of question extraction was measured using the documents containing MPSC MCQs from past years.

**Table 1.** Automated Question Extraction Performance

Metric	Value (%)
Precision	94.2

Recall	92.6
F1-Score	93.4

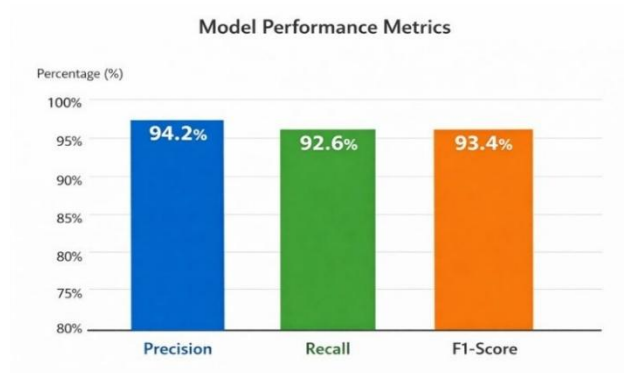


Fig. 2. Model Performance Metrics

Difficulty Classification Results

The questions obtained from the test have been classified into three categories, i.e., questions, medium questions, and Hard questions. We used a computer program to classify the questions. After that, the correctness of the classification of the questions into Easy questions, medium questions, and Hard questions has been verified by people who know a lot about the subject. The people who verified the correctness of the classification are experts in the subject.

Table 2. Difficulty Classification Accuracy

Difficulty Level	Accuracy (%)
Easy	91.8
Medium	89.6
Hard	87.9
<b>Overall Accuracy</b>	<b>89.8</b>

AI-Based Proctoring Results

The AI-based proctoring module was tested to see how well it works. This was done by watching videos and listening to audio of students taking tests which are, like practice exams. The AI-based proctoring module was checked to make sure it can monitor students properly during these examinations.

Table 3. Proctoring Detection Accuracy

Proctoring Method	Accuracy (%)
Video (Face Detection)	96.4
Audio (Voice Activity Detection)	93.1
<b>Overall Proctoring Accuracy</b>	<b>94.7</b>

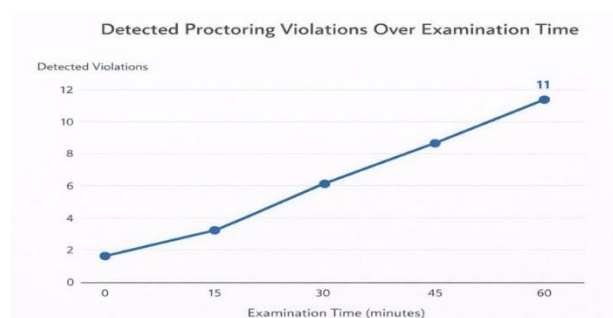


Fig. 4. Detected Proctoring Violations Over Examination Time

Candidate Performance Evaluation Results

The performance of the candidates was looked at in fake tests that were done using the new system. We wanted to see how well the candidates did in these tests. The candidate performance metrics were analyzed to find out how the candidates were doing. We used the proposed system to conduct mock examinations and then we looked at the results to see what they told us about the candidate performance metrics.

**Table 4. Candidate Performance Metrics**

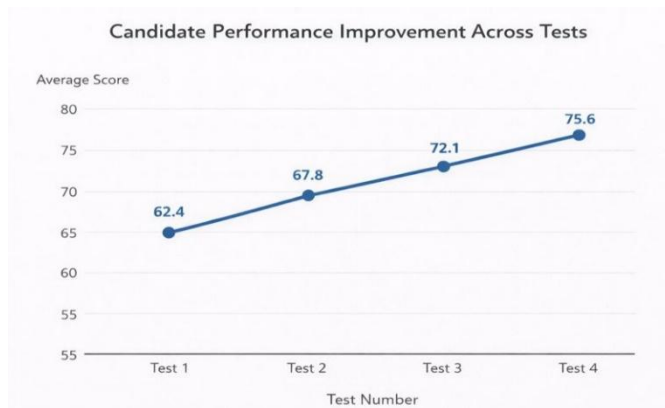
Metric	Average Value
Accuracy	76.3%
Average Time per Question	54.2 seconds
Weighted Score	71.8

Performance Trend Analysis

We examined the categories of applicants in the training tests to determine the degree of knowledge acquisition. We did this to analyze the lifestyle of the candidates in the sentences in the survey. The diversity of the applicants plays a crucial role in determining the teaching practices of the applicants.

**Table 5. Performance Trend Across Tests Test**

Test Number	Average Score
Test 1	62.4
Test 2	67.8
Test 3	72.1
Test 4	75.6



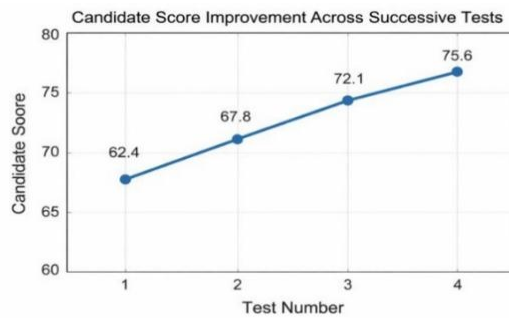
**Fig. 5. Candidate Performance Improvement Across Tests.**

Comparative Analysis with Existing Systems

The proposed system was compared with traditional online examination platforms and QGen-based systems discussed in the literature

**Table 6. Comparative Results**

System	AI Proctoring	Difficulty Classification	Performance Analytics	Overall Effectiveness (%)
Traditional Online Exam	No	No	No	68.5
QGen-based System	No	Partial	No	74.2
<b>Proposed System</b>	Yes	Yes	Yes	<b>89.8</b>



A performance trend graph showing steady improvement in candidate scores across successive tests.

**Fig. 6.** Candidate Score Improvement Across Successive Tests

## Conclusion

The paper presents an ingenious AI-controlled absolutely certain proctored test instrument that can be used to address crucial instances of stressful conditions connected to on line antagonistic tests. The suggested framework transcends the shortcomings of the current structures through incorporating automatic question extraction, problem type that is pushed by AI, multimodal proctoring, and overall established overall performance analytics all in one solution. The suggested system will yield the same balance in the assessment and data-driven evaluation of performance compared to the conventional strategies where popularity certainly evolves on query era or assessment. Discovered results reveal that automatic recall of MCQs to unstructured PDF files has a moderate level of reliability, which allows the entry of scalable and eco-friendly query banks. Moreover, the totally hassle class mechanism based on AI enables determined take a look at generation with the aid of precisely delegating questions to the correct level of trouble, introducing stepped forward assessment prominent in the evaluation to direct or random sampling methods highlighted in previous studies.

Compared to the evaluation to transformer-primarily based query technology frameworks that comprise Q Gen and hybrid AI strategies that are emphasized on current literature, the proposed device demonstrates more practicality through its supported exam protection and analytics features. On the whole, the structure is more robust, just, and dependable than advanced online testing systems, which is a solid platform of mass-training on aggressive exams through intelligent automation, proctoring stability and grading automation. Moreover, the device has robust potential in the future growth through the integration of adaptive trying out methods, multi-linguistic assist, and high-level deep knowing methods.

The system may be more desirable by using integrating adaptive testing to adjust query problem based on student overall performance. The benefits of selecting new university and college students are embellished with the addition of multilingual assistance or support. State-of-the-art deep analytics models can additionally refine query techniques to predict disturbances, scalability, security, and sensitive proctoring may be more beneficial use of cloud integration and advanced behavior analysis.

## References

1. Sharma, A. K., & Gupta, R. "AI-Based Remote Proctoring System Using Deep Learning for Online Examinations," IEEE Access, vol. 12, pp. 14523–14538, 2024.
2. Khan, M. R., Verma, S., & Singh, P. "Real-Time Cheating Detection in Online Exams Using Computer Vision Techniques," Proceedings of the International Conference on Artificial Intelligence and Smart Systems, 2024.
3. Hrvacević, L., Jocović, V., & Drašković, D. "Web-Based Examination Platform with Automated Monitoring Using Large Language Models," IEEE Access, vol. 10, pp. 112345–112356, 2022.
4. Patel, J., & Desai, N. "AI-Driven Online Proctoring System Using Face Recognition and Eye Tracking," International Journal of Advanced Computer Science and Applications, vol. 13, no. 6, pp. 210–218, 2022.
5. Zhang, Y., Liu, H., & Chen, Q. "Deep Learning-Based Suspicious Behavior Detection in Remote Examinations," Expert Systems with Applications, vol. 190, p. 116201, 2022.
6. Bansal, S. K., & Kumar, M. "Secure Online Examination System Using Encryption and Role-Based Access Control," International Journal of Information Security and Privacy, vol. 14, no. 3, pp. 55–70, 2021.
7. Castro-Schez, A., et al. "Automatic Question and Answer Generation Using Machine Learning Techniques," Applied Sciences, vol. 11, no. 8, pp. 1–18, 2021.
8. Pol, S., Patil, A., & Kulkarni, M. "Outcome-Based Predictive Analysis of Automatic Question Paper Generation Using Hybrid MLP-CNN Approach," International Journal of Intelligent Systems and Applications, vol. 13, no. 2, pp. 45–55, 2021.

9. Huang, Y., Zhang, L., & Liu, J. “Knowledge-Based Question Generation Using Controlled Natural Language Processing,” *Expert Systems with Applications*, vol. 168, p. 114382, 2021.
10. Nalawade, S., & Ramesh, B. “Automatic Generation of Question Paper from User Entered Specifications Using a Semantically Tagged Question Repository,” *IJERT*, vol. 9, no. 6, pp. 742–746, 2020.
11. Sangeetha, R., Kumar, P. R., & Balamurugan, S. “Intelligent Question Generator System Using Natural Language Processing,” *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 4, pp. 257–263, 2020.
12. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. “BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding,” *Proceedings of NAACL-HLT*, 2019.
13. Pisat, A., Deshpande, R., & Kulkarni, S. “Question Paper Generator and Solution Verifier Using Natural Language Processing,” *International Journal of Computer Science and Information Technologies*, vol. 9, no. 3, pp. 118–123, 2018.
14. Rahim, T. N. T. A., Aziz, Z. A., Rauf, R. H. A., & Shamsudin, N. “Automated Exam Question Generator Using Genetic Algorithm,” *IEEE Conference on e-Learning, e-Management and e-Services (IC3e)*, 2017.
15. Vaswani, A., et al. “Attention Is All You Need,” *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
16. Kar, S., Saha, A. K., & Roy, D. “Deep Learning-Based Online Exam Proctoring System Using Face and Gaze Tracking,” *IEEE Access*, vol. 11, pp. 78521–78534, 2023.
17. Yadav, R., & Sharma, P. “Automated Question Generation and Difficulty Prediction Using Transformer Models,” *Expert Systems with Applications*, vol. 213, p. 118987, 2023.
18. Liu, H., Wang, X., & Zhao, Y. “Multimodal Behavior Analysis for Intelligent Online Examination Monitoring,” *Pattern Recognition Letters*, vol. 166, pp. 45–53, 2023.
19. Al-Ghamdi, M., & Alshammari, T. “Blockchain-Based Secure Online Examination Framework with Smart Contract Integration,” *IEEE Access*, vol. 10, pp. 99876–99889, 2022.
20. Mishra, P. K., & Jain, S. “Adaptive Testing System Using Machine Learning for Personalized Competitive Exam Preparation,” *International Journal of Intelligent Systems*, vol. 37, no. 9, pp. 6421–6440, 2022.