

AI-Driven Quiz Generation and Evaluation Platform for Government Recruitment Exams

Shriyash Khot¹, Karan Pujari², Gayatri Toraskar³, Ms. Poonam Kadam⁴

^{1,2,3} Student, Department of Computer Engineering, Metropolitan Institute of Technology and Management (MITM), Sindhudurg Maharashtra, India

⁴ Ms. Poonam Kadam, Assistant Professor and Project Guide, Department of Computer Engineering, Metropolitan Institute of Technology and Management (MITM), Sindhudurg Maharashtra, India

DOI: 10.5281/zenodo.20627607

ABSTRACT

Government recruitment examinations such as Police Bharti, Army Recruitment, Railway Recruitment Board exams and banking recruitment tests attract millions of aspirants across India every year. These examinations require structured preparation in subjects including mathematics, reasoning ability, general knowledge, and English. Traditional preparation systems rely on printed books, coaching institutes, and static online mock tests that provide limited evaluation and feedback mechanisms. This research proposes an AI-Driven Quiz Generation and Evaluation Platform designed to modernize competitive exam preparation through artificial intelligence, data analytics, and automation. The system integrates Optical Character Recognition (OCR) technology to convert scanned government recruitment exam papers into interactive quizzes. Using Natural Language Processing techniques such as TF-IDF vectorization and cosine similarity, the system evaluates answers intelligently rather than using simple keyword matching.

The proposed system also incorporates transformer-based generative AI models capable of generating contextual explanations for quiz answers. These explanations help students understand concepts in depth instead of memorizing answers. Additionally, performance analytics are implemented using clustering algorithms such as K-Means to categorize learners into performance groups including high performers, moderate performers, and learners who require improvement.

Keyword: Artificial Intelligence, OCR, Natural Language Processing, Competitive Exams, Quiz Generation, Adaptive Learning.

1. INTRODUCTION

Government recruitment examinations are among the most competitive examinations in India. Aspirants prepare for months or even years to succeed in examinations such as Police Bharti, Army Recruitment, Banking exams, and Railway Recruitment Board tests. However, many students face difficulties due to the lack of structured feedback during their preparation.

Traditional exam preparation platforms provide static question banks and simple score-based evaluations. These systems rarely analyze detailed learning patterns or provide personalized recommendations. With advancements in Artificial Intelligence and Machine Learning technologies, modern educational platforms can deliver adaptive learning experiences and intelligent tutoring systems.

The proposed research aims to develop an AI-driven quiz generation platform capable of automatically converting scanned question papers into digital quizzes. By integrating artificial intelligence techniques, the system provides automated evaluation, explanation generation, and performance analytics. This approach significantly improves learning outcomes and helps aspirants focus on weak subjects more effectively.

2. PROBLEM STATEMENT

Traditional mock test platforms provide limited analytics and often fail to identify the strengths and weaknesses of students. Students only receive a final score without understanding their performance patterns. Therefore, there is a need for an intelligent system capable of generating quizzes automatically, evaluating responses intelligently, and providing personalized feedback for exam preparation.

Objectives

1. Automate quiz generation from scanned exam papers using OCR.
2. Provide intelligent evaluation using Natural Language Processing.
3. Generate AI-based explanations for quiz answers.
4. Track subject-wise performance analytics for students.
5. Categorize learners using machine learning clustering techniques.

3. LITERATURE REVIEW

Several online platforms and educational technologies have been developed to support students preparing for competitive examinations. Most existing exam preparation platforms provide large question banks, mock tests, and score-based evaluation systems. However, these platforms typically focus on delivering static content rather than providing intelligent learning support or personalized feedback mechanisms. As a result, students often receive only their final scores without understanding their strengths, weaknesses, or subject-wise performance patterns.

Research in the field of Intelligent Tutoring Systems (ITS) has demonstrated that adaptive learning systems significantly improve knowledge retention and learning efficiency. Intelligent tutoring systems use artificial intelligence techniques to analyse user behaviour, track learning progress, and provide personalized recommendations. These systems enable students to receive targeted feedback based on their performance, which helps them focus on areas that require improvement. Studies in AI-based education platforms indicate that adaptive learning environments can enhance student engagement and improve overall learning outcomes.

Recent advancements in Natural Language Processing (NLP) have also contributed significantly to the development of intelligent educational platforms. NLP techniques enable computers to understand and process human language, allowing systems to analyse textual data such as quiz questions, answers, and explanations. Techniques such as Term Frequency–Inverse Document Frequency (TF-IDF) and semantic similarity measures are widely used to analyse text and evaluate descriptive responses. These techniques help educational platforms move beyond simple keyword matching and instead evaluate answers based on meaning and context.

4. PROPOSED SYSTEM

The proposed AI-driven quiz generation and evaluation platform is designed to modernize the preparation process for government recruitment examinations by integrating artificial intelligence, natural language processing, and data analytics. The system provides an intelligent learning environment that automates quiz generation, evaluates user responses, and provides detailed performance insights for students preparing for competitive exams.

The platform consists of several interconnected modules including the user interface module, backend processing services, OCR module, AI explanation engine, and performance analytics engine. Each module performs a specific function within the system architecture and collectively enables automated quiz generation, intelligent evaluation, and adaptive learning support.

The primary objective of the proposed system is to reduce manual effort in quiz preparation and provide students with a personalized learning experience. Unlike traditional exam preparation platforms that provide only static question banks, the proposed system dynamically generates quizzes from scanned exam papers and evaluates answers using artificial intelligence techniques.

System Architecture

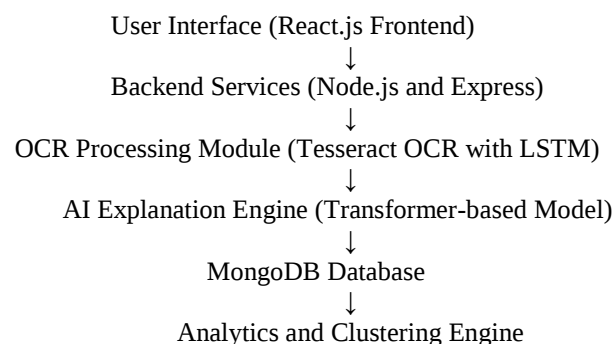


Table 1 - Technology Stack Used in the Proposed System.

Technology	Module	Purpose
React.js	Frontend	User Interface
Node.js	Backend	API and server logic
MongoDB	Database	Store quiz data
Tesseract OCR	OCR	Extract text from papers
Transformer AI	AI Engine	Generate explanations

5. ALGORITHMIC FRAMEWORK

The proposed AI-driven quiz generation and evaluation platform integrates several machine learning and natural language processing algorithms to process quiz data, evaluate user responses, and generate intelligent learning

feedback. These algorithms work together to automate quiz generation, perform semantic answer evaluation, and analyze user performance patterns.

The primary algorithms used in the system include LSTM-based Optical Character Recognition (OCR), TF-IDF vectorization, Cosine Similarity, K-Means clustering, and Transformer-based generative models. Each algorithm performs a specific function within the system architecture and contributes to the overall intelligent tutoring process.

5.1 LSTM-Based Optical Character Recognition (OCR)

Optical Character Recognition (OCR) technology is used to convert scanned exam papers or images containing text into machine-readable digital text. The proposed system uses Tesseract OCR, which utilizes Long Short-Term Memory (LSTM) neural networks for accurate text recognition.

LSTM networks are a type of Recurrent Neural Network (RNN) designed to process sequential data. They are capable of learning long-term dependencies within text sequences and are particularly effective for recognizing characters in scanned documents.

The OCR module processes uploaded exam papers in the following steps:

1. Image preprocessing (noise removal and contrast enhancement)
2. Character segmentation
3. Feature extraction
4. Text recognition using LSTM networks

After processing, the extracted text is structured into quiz questions and stored in the system database.

Advantages of LSTM-based OCR include:

1. High accuracy in recognizing printed and scanned text
2. Ability to handle sequential character patterns
3. Improved recognition of complex document layouts

6. PERFORMANCE ANALYSIS

The performance analysis compares the traditional quiz platforms with the proposed AI-driven platform. In traditional systems, quiz generation is usually performed manually, which requires significant time and effort from instructors. In contrast, the proposed AI platform automatically generates quizzes using advanced algorithms and OCR techniques, making the process faster and more efficient. Traditional platforms evaluate answers using simple keyword matching, which may not accurately capture the meaning of a student's response.

Table 2 - Performance Comparison Between Traditional and Proposed AI Platform.

Metric	Traditional Platform	Proposed AI Platform
Quiz Generation Time	Manual	Automatic
Answer Evaluation	Keyword matching	Semantic similarity
Analytics	Basic score	Advanced analytics dashboard
Learning Improvement	Limited	Adaptive recommendations

7. RESULTS AND DISCUSSION

The experimental evaluation of the proposed AI-driven quiz generation and evaluation platform demonstrates its effectiveness in automating quiz creation, evaluating user responses, and providing detailed performance analytics. The system was tested using a dataset consisting of past government recruitment exam papers including Police Bharti, Railway Recruitment Board examinations, and banking aptitude tests.

The prototype platform successfully converts scanned question papers into structured quiz formats using the OCR module. The OCR component based on LSTM networks achieved high accuracy in extracting textual information from scanned documents. After text extraction, the system processes the questions using Natural Language Processing techniques and converts them into structured quiz questions that can be attempted through the web interface.

The automated evaluation system was tested using both objective and descriptive questions. For objective questions, the system performs direct answer matching with stored correct answers. For descriptive questions, the system applies semantic similarity techniques such as cosine similarity to compare user responses with reference answers. This approach improves answer evaluation accuracy compared to traditional keyword-based methods.

The analytics dashboard plays a crucial role in visualizing user performance. The dashboard displays several metrics including:

1. Subject-wise accuracy
2. Average quiz completion time
3. Performance improvement trends
4. Weak topic identification

These analytics allow students to easily identify subjects in which they need improvement. For example, if a student consistently performs poorly in reasoning questions but performs well in mathematics, the system highlights reasoning as a weak area. This helps students adjust their preparation strategy more effectively.

To analyse performance patterns across multiple users, the system applies the K-Means clustering algorithm. Based on quiz accuracy, time taken, and subject scores, users are categorized into three performance groups:

1. High Performers
2. Moderate Performers
3. Needs Improvement

This clustering approach enables the platform to generate personalized learning recommendations. For example, students classified under the “Needs Improvement” category may receive additional practice quizzes and AI-generated explanations for difficult topics.

8. CONCLUSIONS

The proposed AI-driven quiz generation and evaluation platform provides a modern approach to improving competitive exam preparation through the integration of artificial intelligence and data analytics. Traditional exam preparation systems primarily rely on static mock tests and manual evaluation methods that provide limited feedback to students. In contrast, the proposed system introduces an intelligent learning environment that automates quiz generation, performs semantic answer evaluation, and provides detailed performance analytics.

The platform utilizes Optical Character Recognition (OCR) technology to convert scanned exam papers into digital quizzes automatically. This feature eliminates the need for manual question entry and significantly reduces the time required to generate quizzes. By integrating Natural Language Processing (NLP) techniques such as TF-IDF vectorization and cosine similarity, the system is able to evaluate descriptive answers more accurately compared to traditional keyword-matching approaches.

Another important component of the system is the performance analytics module, which tracks student performance across different subjects and quizzes. Using machine learning techniques such as K-Means clustering, the system categorizes learners into different performance groups based on accuracy, time taken, and subject-wise scores. This enables the platform to generate personalized recommendations and adaptive learning paths for users.

9. REFERENCES

- [1] R. Smith, 'An Overview of the Tesseract OCR Engine', ICDAR Conference.
- [2] Vaswani et al., 'Attention Is All You Need', NeurIPS 2017.
- [3] J. R. Anderson, 'Intelligent Tutoring Systems', AI in Education.
- [4] OpenAI Documentation, GPT API Reference 2024.
- [3] J. R. Anderson, “Intelligent Tutoring Systems,” *Artificial Intelligence in Education*, 1985.
- [4] OpenAI, “GPT API Documentation,” OpenAI Technical Report, 2024.
- [5] J. Paladines and J. Ramírez, “A Systematic Literature Review of Intelligent Tutoring Systems with Dialogue in Natural Language,” *IEEE Access*, vol. 8, pp. 164246–164267, 2020.
- [6] K. VanLehn et al., “Learning How to Construct Models of Dynamic Systems: Evaluation of the Dragoon Intelligent Tutoring System,” *IEEE Conference on Advanced Learning Technologies*, 2017.
- [7] A. Devlin et al., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *NAACL*, 2019.
- [8] T. Brown et al., “Language Models are Few-Shot Learners,” *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.
- [9] R. Smith, D. Antonova, and D. S. Lee, “Adapting the Tesseract Open-Source OCR Engine for Multilingual OCR,” *International Workshop on Multilingual OCR*, ACM, 2009.
- [10] A. Rogers, O. Kovaleva, and A. Rumshisky, “A Primer in BERTology: What We Know About How BERT Works,” *Transactions of the Association for Computational Linguistics*, 2020.
- [11] N. Patwardhan et al., “Transformers in the Real World: A Survey on NLP Applications,” *Information Journal*, 2023.
- [12] D. Kulshreshtha et al., “Few-shot Question Generation for Personalized Feedback in Intelligent Tutoring Systems,” *arXiv preprint*, 2022.
- [13] C. C. Lin et al., “Artificial Intelligence in Intelligent Tutoring Systems toward Sustainable Education,” *International Journal of Educational Technology*, 2023.
- [14] M. As'ad et al., “Intelligent Tutoring Systems with Generative AI Integration,” *AI in Education Journal*, 2024.
- [15] W. Feng et al., “Research on the Construction and Application of Intelligent Tutoring Systems Based on Generative Models,” *Educational Technology Research*, 2025.

- [16] S. Robby et al., "Implementation of Optical Character Recognition Using Tesseract," *Procedia Computer Science*, 2019.
- [17] A. Graesser et al., "Auto Tutor: A Tutor with Dialogue in Natural Language," *Institute for Intelligent Systems Research*, 2005.