International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Automatic Question Paper Generation With Marks Allocation Using Bloom's Taxonomy

V. R. Lele, Associate Professor, Dept. of Computer Engineering, KCT's Late G N Sapkal College Of Engineering, Nashik, Savitribai Phule Pune University, Maharashtra.

Siddhesh Andhale, Student of, Department Computer Engineering, KCT's Late G N Sapkal College Of Engineering, Nashik, Savitribai Phule Pune University, Maharashtra.

Sahil Bramhankar, Student of, Department Computer Engineering, KCT's Late G N Sapkal College Of Engineering, Nashik, Savitribai Phule Pune University, Maharashtra.

Saurabh Nalkar, Student of, Department Computer Engineering, KCT's Late G N Sapkal College Of Engineering, Nashik, Savitribai Phule Pune University, Maharashtra,

Prathamesh Nemade, Student of, Department Computer Engineering, KCT's Late G N Sapkal College Of Engineering, Nashik, Savitribai Phule Pune University, Maharashtra,

---

Abstract - In the field of education, creating evaluations that are in line with the goals of the courses is a significant difficulty for instructors. The task becomes harder due to the requirement for diversity in question papers and the need to follow university assessment norms. In order to tackle this problem, a system that can quickly create question papers according to teacher guidelines is desperately needed. To successfully explain and define questions, researchers recommend using a variety of tags, such as cognitive level, difficulty, question type, and content/topic. Our suggested remedy for this is the use of an independent mechanism for creating question papers. Users can input a series of questions and add varying degrees of complexity to each question here. Using Bloom's taxonomy, the system uses machine learning techniques to evaluate and mark questions in order to expedite this process. Subsequently, the indicated questions are effectively recorded in a database together with their corresponding levels of complexity. In addition to saving teachers time, this creative method aims to guarantee that the question papers are of a high calibre and reflect the course's overall learning objectives. accurately give the question's marks and also be able to use machine learning and Bloom's Taxonomy approach to create various question papers.

Key Words: Question paper generation, Machine learning, Bloom's taxonomy, Natural Language Processing (NLP)

1.INTRODUCTION

Creating well-crafted question papers that correspond with the learning objectives of a course presents a formidable obstacle for instructors. This problem is increased by the lack of regulated processes to guarantee the caliber of exam questions. A system that can quickly and automatically create exam questions based on teacher-entered parameters is becoming more and more necessary in order to address this. To fully define questions, researchers advise using several sets of tags, such as cognitive level, difficulty level, question type, and topic.

Our suggested approach presents an automated test question creation mechanism in response to this difficulty. Users can enter a list of questions into this system. In the modern world, education is a crucial path to success, underscoring the significance of tests in the educational process. Question papers have traditionally been generated manually by authorities. They create the paper by hand. Nevertheless, this approach presents certain limitations, such as possible partiality, duplication, and security issues. We aim to tackle these problems with our suggested Automated Question Paper Generation technique, which provides a quick, efficient, safe, and randomized method. This system runs entirely autonomously, which allays worries about bias, storage capacity, and security. Furthermore, an innovative approach has been implemented to guarantee total randomization of questions, hence avoiding duplications. In conclusion, the suggested Automated Question Paper Generation system seeks to transform the conventional method by offering a secure, effective solution that can be tailored to the various needs of educational institutions.

2. AIM

In the modern educational environment, the traditional process of creating question papers is done by hand. But the realization of this conventional process's intrinsic shortcomings and inefficiencies has made a creative method necessary. A technology that can automate the laborious manual processes involved in creating question papers is desperately needed. In order to improve productivity and respond to the changing needs of education in a more efficient and flexible way, this automation trend is essential.

3.OBJECTIVES

- **Minimize Time:** The principal goal is to decrease the amount of time spent on the manual creation of question papers, with the intention of optimizing conventional techniques to enhance productivity.
- **Implement Automation:** By using computerized equipment to satisfy the unique needs of users, the current manual procedure is to be automated. This guarantees the safe keeping of important information over a longer period of time, making access easier.
- **Error Reduction:** The goal is to reduce errors as much as possible during the question paper creation process in order to improve accuracy and overall quality.
- **Flexibility and Efficiency**

4.PROBLEM STATEMENT

Understanding how important education is as a means of achieving success, educational institutions view the evaluation process as an essential tool for gauging students' performance. Exam preparation takes a lot of time and effort on the part of teachers, who usually create question papers by hand that are in line with the syllabus. This labor-intensive manual process, carried either by paper setters or examiners, takes a long time and occasionally leads to the award of inaccurate marks. We suggest putting our approach into use to address these issues. A cornerstone of our methodology is Bloom's Taxonomy (BT), a commonly accepted framework for developing and assessing exam questions. The system recognizes important keywords and verbs in queries by utilizing Natural Language Processing (NLP) techniques along with a rule-based methodology. The Naïve Bayes algorithm is then employed to classify the questions. The main objective is to create an independent system that can grade questions according to Bloom's taxonomy. This approach seeks to minimize the amount of time spent on question paper production while simultaneously attempting to minimize human mistake. The main goal is to automate the current manual system by integrating electronic technology and fulfilling the demands of educational establishments. This automatic method makes sure that important data is stored for a long time and makes it easy to access and manipulate. The suggested system's capacity to produce question papers with assigned marks is a noteworthy feature. Additionally, the system offers the ease of emailing the generated paper in CSV format straight to the examiner. All things considered, the system represents a complete answer to optimize and modernize the process of creating question papers for educational examinations.

5.LITERATURE SURVEY

“Exam Question Classification Based on Bloom’s Taxonomy: Approaches and Techniques” (Karima Makhlouf, Lobna Amouri, Nada Chaabane and Nahla EL-Haggar) - The Eighth IEEE International Conference in 2020 included a comprehensive analysis of popular methods and strategies for categorizing test questions according to Bloom's Taxonomy. This thorough analysis explored different approaches to accomplish the objective of question classification and provided several implementation routes. Some of these methods were renowned for their intricacy. Many scholars have worked hard in recent years to automate the Bloom's Taxonomy- based classification of exam questions. This paper provides a comprehensive overview of common approaches and strategies used in question classification (QC).

R. Sun, X. Zhou and F. Fang,” Neural Question Generation Using Question Type Guidance,” 2021 17th International Conference on Computational Intelligence and Security (CIS), 2021, pp. 328-332, - The main focus of this paper is how to use question types to impact the question creation process. We incorporate the task of question type prediction into our multi-task framework to reveal the underlying relationship between the generated question and the context-answer pair. The author also includes a metric learning method to improve the semantic relevance between the context and the generated question. The model performs better than baseline systems, which is especially noticeable on the SQUAD test. The significance of question type guidance in the question generating process is highlighted by the experimental results.

R. Ragasudha and M. Saravanan,” Secure Automatic Question Paper Generation with the Subjective Answer Evaluation System,” 2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN), 2022, pp. 1-5 -This study describes a safe, automated approach for creating test questions and evaluating artificially intelligent, subjective responses. The administrator creates a database with questions classified using Bloom's taxonomy in order to initiate the automatic question paper generating procedure. Questions are generated automatically when you click the "generate" button. Cryptography is used to provide secure delivery to authorized persons. The administrator creates a database containing questions and related keywords for the purpose of evaluating the responses. Using a keyword matching technique, the system evaluates responses automatically during the test. The purpose of putting this system in place is to minimize human mistake, cut down on the amount of time needed to generate questions, and speed up the assessment of responses.

V. M. Kale and A. W. Kiwelekar,” An algorithm for question paper template generation in question paper generation system,” 2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE), 2013, pp. 256-261 - This work presents an algorithm for generating question paper templates that follow predefined guidelines. The paper's algorithm makes use of four constraints: Bloom's taxonomy-based cognitive level coverage, syllabus coverage, difficulty level coverage, and question paper structure. The algorithm's adaptability is shown by how well it can handle any user-specified restrictions.

6.SYSTEM ARCHITECTURE

- **Data Collection:** Compile a dataset containing questions, grades, and Bloom's Taxonomy levels for each question. The machine learning models will be trained on this dataset.
- **Preprocessing:** Clean and preprocess the dataset by addressing missing values, eliminating superfluous information, and converting the textual data into an analysis-ready format.

- **Feature Engineering:** Take note of pertinent details from the questions, like the frequency of words, the construction of sentences, and the semantic meaning. The machine learning models will be trained using these features.
- **Training the Models:** Utilize machine learning techniques, like Random Forest or Linear Regression, to train models that, using the attributes that were retrieved, can predict the scores and Bloom's Taxonomy levels for new questions.
- **User Interaction:** Create an interface that lets users enter questions or themes to be used in the creation of question sheets. Additionally, the system must to offer choices for defining the preferred degree of difficulty or Bloom's Taxonomy distribution.
- **Question Paper Generation:** Question papers can be automatically generated using the trained models in response to user input. The system will ensure that Bloom's Taxonomy levels are distributed appropriately and assign marks to each question based on projected values.
- **Verification and Refinement:** Check the generated question papers for quality and accuracy. If necessary, improve the models' performance or make changes to the papers that are produced.
- **Output:** Give the user access to the created question paper, outlining each question with its corresponding mark and Bloom's Taxonomy level.

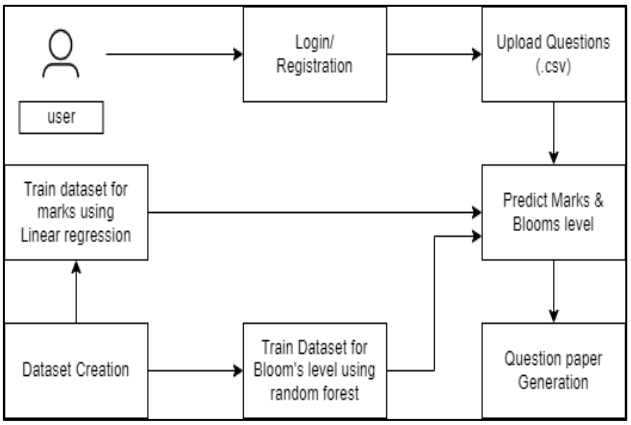

Fig -1: System Architecture Diagram

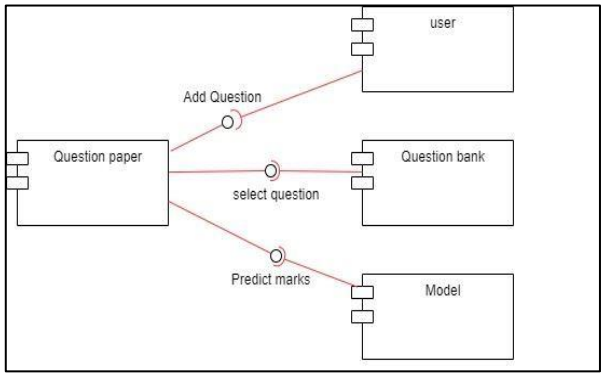

Fig-2: System Modules Diagram

7.MODULES

- **User Registration:**
  This module makes it easier for new users to register. It gathers user data, carries out verifications, and establishes user accounts inside the framework.
- **User Login:**
  Here, we're putting into practice a feature that allows users to log in to our application or system and use other capabilities. For users, this module manages the authentication procedure. It has features like session management, password validation, and user login.
- **Data set Trained:**
  The dataset that is used to train the machine learning models is managed by this module. It has features for storing, preprocessing, and importing datasets. Additionally, data validation and data assurance might be involved.
- **Bloom's Taxonomy Prediction:**
  This module uses machine learning algorithms, including Random Forest, to estimate the Bloom's taxonomy level for new queries. It determines the appropriate Bloom's taxonomy level by using the content of the query as input.
- **Mark Prediction:**
  With the use of machine learning techniques like linear regression, this module forecasts the grades for every question. To estimate the marks, it takes into account variables including the question's content, Bloom's taxonomy level, and possibly other characteristics.
- **Question paper generation:**
  In order to create the question paper, we submit the csv file to the system. The system then reads the questions from the file and generates the paper. This module creates test questions according to predetermined standards, like the quantity of questions and the arrangement of Bloom's taxonomy levels. To produce a well-balanced question paper, it chooses pertinent questions from the question database.

8. CLASS DIAGRAM

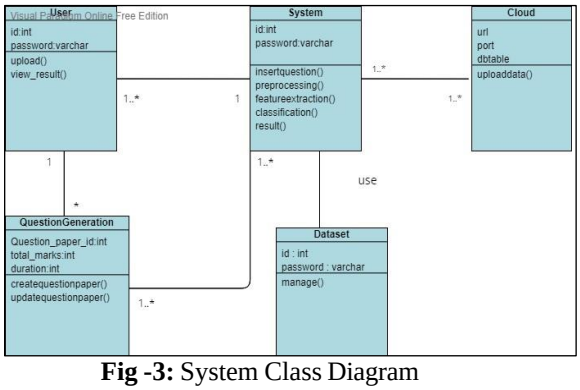

Fig -3: System Class Diagram

9.STATE DIAGRAM

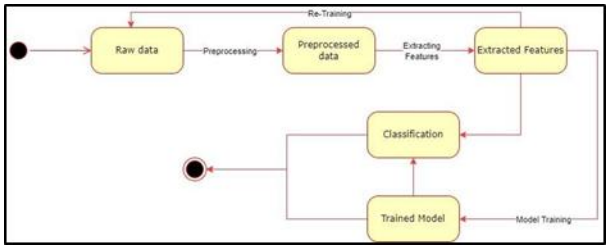

Fig -4: System State Diagram

10. UML DIAGRAM

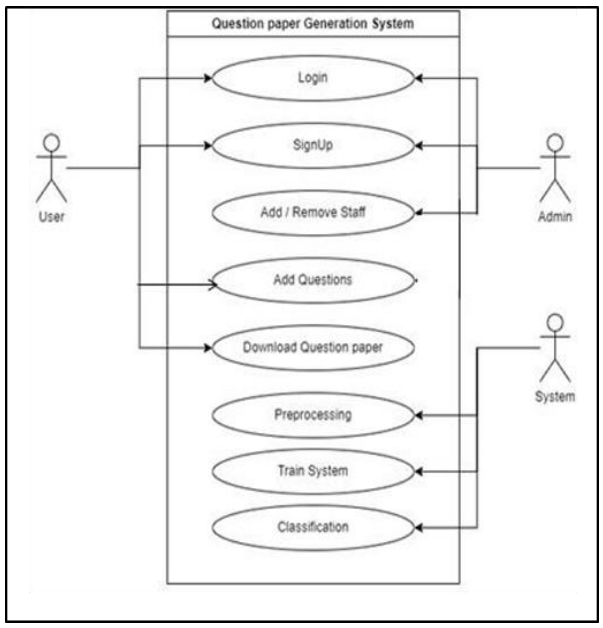

Fig -5: System UML Diagram

11. ALGORITHM

- Dataset creation
- Pre-processing of data
- Transform your questions into numerical form
- Assign grades
- Include the algorithm for Random Forest
- Creation of the final question sheet
- Stop

12. BLOOM'S TAXONOMY

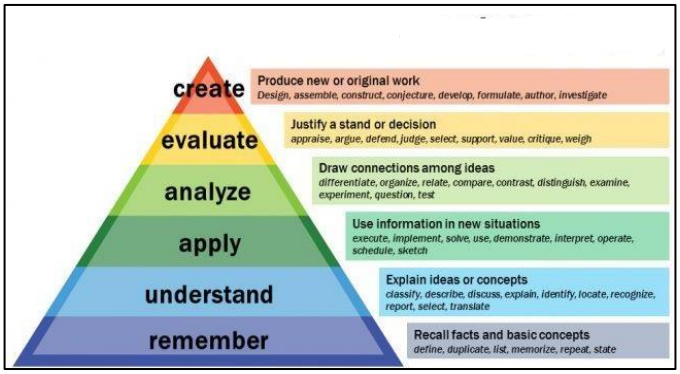

Fig -6: Bloom's Taxonomy levels

1) Remember:

a) **Definition**: Retrieve, retrieve, or identify pertinent information from long-term memory (e.g., remember the parts of a bacterial cell, or the dates of significant historical events in US history).

b) ***Keywords****: cite, define, describe, identify, label, list, match, name, outline, quote, recall, report, reproduce, retrieve, show, state, tabulate, and tell.*

2) Understand:

a) Definition: Show understanding using one or more explanation techniques (e.g., categorize mental illnesses, draw comparisons between religious rituals in two distinct religions).

b) Keywords*: abstract, arrange, articulate, associate, categorize, clarify, classify, compare, compute, conclude, contrast, defend, diagram, differentiate, discuss, distinguish, estimate, exemplify, explain, extend, extrapolate, generalize, give examples of, illustrate, infer, interpolate, interpret, match, outline, paraphrase, predict, rearrange, reorder, rephrase, represent, restate, summarize, transform, and translate.*

3) Apply:

a) Definition: use knowledge or expertise in a novel setting (e.g., perform a multivariate statistical analysis using a data set you haven't worked with before, or apply Newton's second law to solve an issue).

b) Keywords*: apply, calculate, carry out, classify, complete, compute, demonstrate, dramatize, employ, examine, execute, experiment, generalize, illustrate, implement, infer, interpret, manipulate, modify, operate, organize, outline, predict, solve, transfer, translate, and use.*

4) Analyze

a) Definition: Dissect a piece of writing into its

component parts and ascertain how those parts relate to each other and/or to a larger structure or purpose (e.g., analyze the relationship between different characters in a play; analyze the relationship between different institutions in a society; or break the material down into its constituent parts).

b) ***Keywords****; analyze, arrange, break down, categorize, classify, compare, connect, contrast, deconstruct, detect, diagram, differentiate, discriminate, distinguish, divide, explain, identify, integrate, inventory, order, organize, relate, separate, and structure.*

5) Evaluate:

a) Definition: make decisions based on standards and criteria (e.g., identify fallacies or inconsistencies in a process or product, assess whether a scientist's conclusions are supported by observed data, decide between two approaches to solve a particular problem, assess a product's quality based on disciplinary criteria).

b) ***Keywords****; appraise, apprise, argue, assess, compare, conclude, consider, contrast, convince, criticize, critique, decide, determine, discriminate, evaluate, grade, judge, justify, measure, rank, rate, recommend, review, score, select, standardize, support, test, and validate.*

6) Create:

a) Definition: combine components to create a fresh, functional, or harmonious whole; compose a new piece of music, write a play, create a new set for a theater show, build a different hypothesis based on criteria, and reorganize pieces into a new pattern or structure.

b) Keywords*: arrange, assemble, build, collect, combine, compile, compose, constitute, construct, create, design, develop, devise, formulate, generate, hypothesize, integrate, invent, make, manage, modify, organize, perform, plan, prepare, produce, propose, rearrange, reconstruct, reorganize, revise, rewrite, specify, synthesize, and write.*

13. ALGORITHM/TECHNOLOGY

- To train machine learning models, obtain a dataset that includes questions, their related marks, and Bloom's Taxonomy levels.
- Purify and preprocess the dataset by removing unneeded data, dealing with null values, and formatting textual data so that it can be analyzed easily.
- To train machine learning models, extract relevant data from questions, such as word frequencies, sentence structure, and semantic meaning.
- Create training models that use extracted data to predict marks and Bloom's Taxonomy levels for new questions using methods such as Random Forest or Linear Regression.
- Provide an interface that allows users to enter questions or subjects for the creation of question papers. The system ought to enable the specification of Bloom's Taxonomy distribution or difficulty levels.
- Use user input to automatically construct question papers using learned models. Give questions marks based on expected values to maintain a balanced Bloom's Taxonomy distribution.
- Check the quality and accuracy of the question papers that are generated. For improved performance, edit papers or modify models as necessary.
- Show the user the created question paper, with each question shown with its corresponding mark and Bloom's Taxonomy level.

14. RULES DEVELOPMENT

In this study, exam questions are categorized using a rule-based approach according to Bloom's Taxonomy. The guidelines are based on a training set of seventy questions covering different topics. The application of the rule is governed by two criteria: the necessity to identify patterns in question structures and the situation in which a term shares more than one category. Part-of-speech tagging reveals patterns, such as the fact that most questions begin with a verb. For example, the following pattern can be seen in the question "Write down the output of the following program": In order to classify questions, the algorithm provides two rules: Rule 1 and Rule 2. For instance, Rule 2 gives weight to "synthesis," but Rule 1 gives weight to "knowledge."

15. RANDOMIZATION ALGORITHM

- Given a database with N questions in it:
- Make a list called "L" with N entries.
- Create the number "n" at random such that 1<=n <= N.
- Go back to Step 3 if n is already in list L.
- If not, record n in list L.
- Choose a question where the flag is true from the database that corresponds to n.
- Make sure the flag for the chosen question is set to false.

16. NAÏVE BAYES ALGORITHM

1. **Initialize:** Compile a dataset that has labels for each feature.
2. **Data Preparation:** Assist with missing values, preprocess the data, and make sure its format is appropriate for analysis.
3. **Feature Extraction:** Take pertinent features out of the dataset.
4. **Data Splitting:** Divide the dataset into sets for testing and training.
5. **Model Training:** Compute the probability that the model requires using the training dataset.

6. **Probability Calculation:** Determine the conditional probability of every feature for every class.
7. **Prior Probability:** Determine each class's previous probability.
8. **Prediction:** To determine the posterior probability of each class given the features, apply the Bayes theorem.
9. **Class Selection:** As the predicted class, choose the one with the highest posterior probability.
10. **Model Evaluation:** Make use of the testing dataset to evaluate the model's performance.
11. **Iterate and Refine**: Iterate and improve the model as needed in response to input and performance.

17. ACTIVITY DIAGRAM

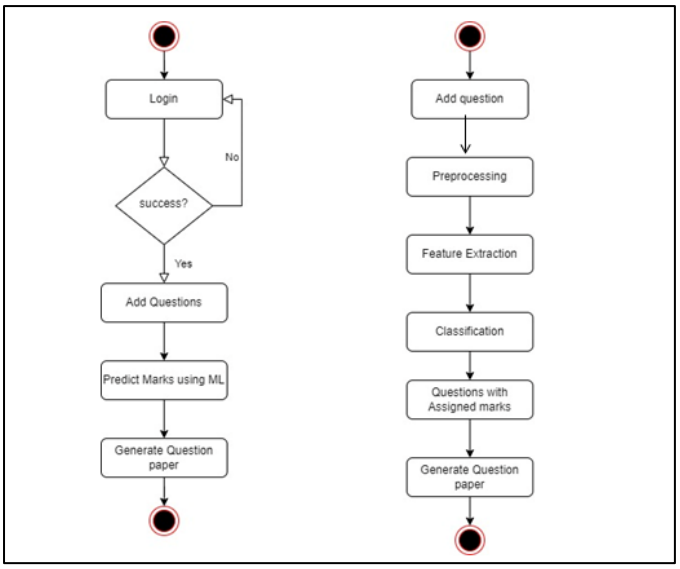

Fig -7: Activity Diagram

18. CONCLUSION

In this project, we recommend using Bloom's Taxonomy to create reliable and significant evaluation questions. We present a system that creates question sheets and allots marks to individual questions. Our technology examines and classifies questions into various cognitive levels by utilizing Bloom's Taxonomy. All question papers will be methodically sorted according to problem kinds and difficulty levels using text mining.

19. REFERENCES

[1] R. Ragasudha and M. Saravanan discussed "Secure Automatic Question Paper Generation with the Subjective Answer Evaluation System" at the 2022 International Conference on Smart Technologies and Systems for Next Generation Computing.

[2] J. Zhang discussed the "Online Examination System of Automatic Paper Composition Based on Hybrid Algorithm" in the China Computer Communication journal in 2021.

[3] B. Zhao, H. Ning, and R. B. Zhang conducted research on the "Automatic Test Paper Generation System based on Genetic Algorithm" in the Applied Science and Technology journal in 2021.

[4] In the paper titled "Neural Question Generation Using Question Type Guidance," authors R. Sun, X. Zhou, and F. Fang shared their insights at the 2021 International Conference on Computational Intelligence and Security.

[5] "Exam Question Classification Based on Bloom's Taxonomy: Approaches and Techniques," authors Karima Makhlouf, Lobna Amouri, Nada Chaabane, and Nahla EL-Haggar presented insights at the 2020 IEEE Eighth International Conference.

[6] Z. C. Wang and X. Y. Ouyan focused on the "Design of Intelligent Test Paper Generation System Based on Improved Genetic Algorithms" in the Computer Knowledge and Technology journal in 2020.

[7] G. Nalawade and R. Ramesh discussed the "Automatic Generation of Question Paper from User Entered Specifications Using a Semantically Tagged Question Repository" at the 2016 IEEE Eighth International Conference on Technology for Education.

[8] V. M. Kale and A. W. Kiwelekar presented an algorithm for question paper template generation in their paper at the 2013 International Conference on Technological Advances in Electrical, Electronics, and Computer Engineering.

[9] Q. F. Cheng, H. T. Liu, and X. M. Yang explored the "Design of Test Paper Strategy Based on Randomized Algorithm" in the Science Mosaic journal in 2011.