IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, May 2024

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

Prof. I. M. Shaikh, Pathak Yogiraj, Pawar Sharad, Tilekar Chetan, Pujari Vaishnavi
Department of Computer Engineering
Pune Vidyarthi Griha's College of Engineering & Shrikrushna S. Dhamankar Institute of Management, Nashik

Abstract: *In any educational course curriculum, courses are defined by specific learning objectives. To assess whether students have achieved these objectives, teachers conduct various assessments. However, creating diverse question papers that align with these objectives and meet university assessment standards is a challenging task for educators. Currently, there are no standardized methods to ensure the quality of these question papers, highlighting the need for a system that can automatically generate question papers based on teacher specifications within seconds. Researchers suggest using various tags to define questions, such as cognitive level, difficulty level, question type, and content/topic. We propose an autonomous question paper generation system that addresses this need. This system allows educators to input a set of questions and specify the complexity of each one. Using machine learning techniques, the system assigns marks to each question based on Bloom's taxonomy. These questions, along with their assigned marks, are then stored in a database, ensuring a consistent and efficient approach to question paper creation.*

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

I. INTRODUCTION

Creating question papers that encompass a variety of topics and adhere to a course's learning objectives is a challenging task for professors. Currently, there are no established procedures to guarantee the quality of test questions. Consequently, there is a need for a system that can quickly and automatically generate test questions based on teacher-provided specifications. Researchers recommend using various tags to define questions, such as cognitive level, difficulty level, question type, and content/topic. To address this need, we propose an automated question paper generation system. This system allows users to input a list of questions, each tagged with the relevant attributes. In today's world, education is crucial for achieving success, and examinations play a vital role in the educational process. Properly structured examination papers are essential for accurately assessing students' knowledge. Traditionally, question papers have been generated manually by officials, a method that can be prone to bias, repetition, and security issues. Our proposed system offers an automated process for question paper generation that is fast, streamlined, randomized, and secure. Automation ensures efficient use of storage space and addresses concerns related to bias and security. Additionally, we have developed a new algorithm that guarantees complete randomization of questions and prevents repetition. This proposed system can be highly beneficial for educational institutions, providing a reliable and efficient method for creating high-quality examination papers.

II. LITERATURE SURVEY

[1] "Exam Question Classification Based on Bloom's Taxonomy: Approaches and Techniques" (Karima Makhlouf, Lobna Amouri, Nada Chaabane and Nahla EL-Haggar) - The 2020 IEEE Eighth International Conference provided a thorough examination of prevalent approaches and techniques for classifying exam questions based on Bloom's Taxonomy. This comprehensive review delved into various methodologies for achieving the goal of question classification, presenting multiple pathways for implementation. Some of these techniques were noted for their complexity. Over the recent years, numerous researchers have dedicated efforts to automating the classification of exam questions according to Bloom's Taxonomy. The paper offers an in-depth overview of the standard methods and techniques employed in question classification (QC).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, May 2024

[2] 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, - This paper centers on leveraging question types to influence the task of question generation. Within our multi-task framework, we introduce the task of predicting question types to unveil the inherent connection between the context-answer pair and the question being generated. Additionally, the author incorporates a metric learning mechanism to enhance the semantic relevance between the generated question and the context. The model exhibits superior performance compared to baseline systems, particularly evident on the SQUAD benchmark. Experimental results underscore the importance of process of question generation.

[3] 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 paper presents a secure automated system for generating question papers and assessing computerized subjective answers. In the process of automatic question paper generation, the administrator establishes a database containing questions categorized with Bloom's taxonomy. Upon clicking the generate button, questions are automatically generated. To ensure secure transmission to authorized personnel, cryptography is employed. In the evaluation of answers, the administrator constructs a database with questions and associated keywords. The implementation of this system aims to mitigate human errors, significantly reduce the time required for question generation, and expedite the evaluation of answers.

[4] 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 paper introduces an algorithm designed for the creation of question paper templates that adhere to specified constraints. The algorithm, detailed in the paper, utilizes four constraints: question paper format, syllabus coverage, difficulty level coverage, and cognitive level coverage based on Bloom's taxonomy. The algorithm's flexibility is demonstrated as it can easily accommodate any user-defined constraints.

III. FEATURES

1. Automated Question Classification: Utilizes Bloom's Taxonomy to categorize questions by cognitive level, applying rule-based methods and NLP techniques to identify key components and extract relevant keywords and verbs.

2. Mark Allocation: Uses the Naïve Bayes algorithm to assign marks based on question classification, ensuring appropriate mark distribution and reducing inconsistencies and errors.

3. Error Reduction: Automates classification and marking to minimize human errors, ensuring consistent question quality and reliable assessments.

4. Efficient Data Storage: Provides long-term, secure storage of questions and classifications, enabling easy access and manipulation of data for future use or analysis.

5. Convenient Distribution: Generates question papers in CSV format for easy modification and printing, with email functionality to send the papers directly to examiners or educational authorities.

IV. DESIGN

Education is essential for success, and assessing learner performance is a critical part of this process. Traditionally, teachers manually create question papers, which is time-consuming and error-prone. To address this, we propose an automated question paper generation system. Leveraging Bloom's Taxonomy for crafting and evaluating questions, the system uses rule-based methods and Natural Language Processing (NLP) techniques to identify key terms. Questions are classified using the Naïve Bayes algorithm, allowing the system to assign marks based on Bloom's Taxonomy. This reduces human error and streamlines the question paper creation process, saving significant time.The goal is to automate the manual system with computerized tools, meeting educational institution needs and ensuring long-term data storage for easy access and manipulation.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, May 2024

V. ALGORITHM

1. Acquire a dataset comprising questions, their corresponding marks, and Bloom's Taxonomy levels for training machine learning models.
2. Cleanse and preprocess the dataset by eliminating irrelevant information, handling missing values, and transforming textual data into an analysis-friendly format.
3. Extract pertinent features from questions, including word frequencies, sentence structure, and semantic meaning, to train machine learning models.
4. Train models, utilizing algorithms like Linear Regression or Random Forest, to predict marks and Bloom's Taxonomy levels for new questions based on extracted features.
5. Create a user interface enabling users to input topics or questions for question paper generation. The system should allow specification of difficulty levels or Bloom's Taxonomy distribution.
6. Employ trained models to automatically generate question papers based on user input. Allocate marks to questions according to predicted values, ensuring a balanced Bloom's Taxonomy distribution.
7. Review generated question papers for accuracy and quality. Make refinements to papers or adjust models if needed for enhanced performance.
8. Present the generated question paper to the user, displaying each question alongside allocated marks and Bloom's Taxonomy level.

VI. IMPLEMENTATION

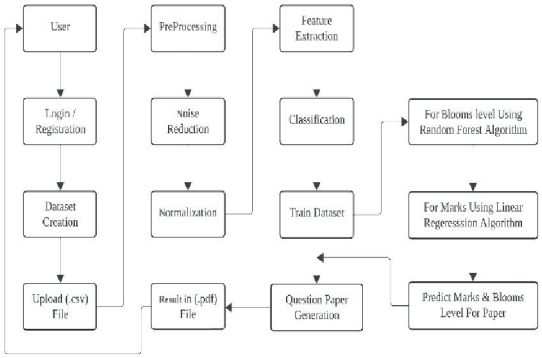


Fig 1: System Architecture

1. User Registration: This module facilitates the registration process for new users by collecting user information, performing validations, and creating user accounts in the system.

2. User Login: This module handles the authentication process, enabling users to log in to the system or application and access other features. It includes functionalities such as user login, password validation, and session management.

3. Dataset Training: This module manages the dataset used for training machine learning models. It includes functionalities for importing, pre-processing, and storing the dataset, as well as data validation to ensure data integrity.

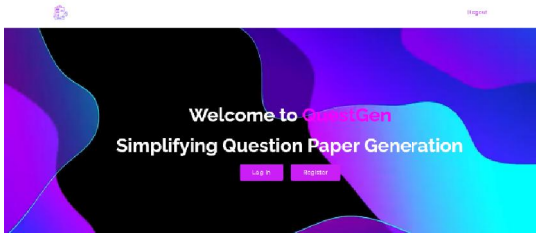
4. Bloom's Taxonomy Prediction: This module uses machine learning algorithms, such as Random Forest, to predict the Bloom's taxonomy level for new questions. It analyses the question content and assigns the appropriate Bloom's taxonomy level.

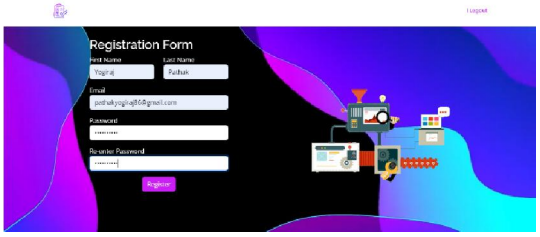
5. Mark Prediction: This module predicts the marks for each question using machine learning algorithms, such as Linear Regression. It takes into account factors like question content and Bloom's taxonomy level to estimate the marks.

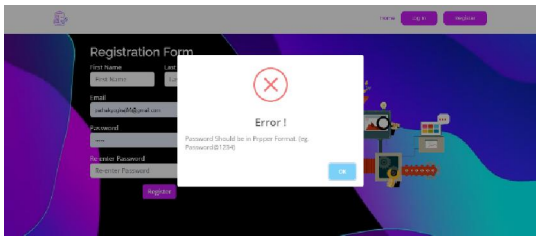
6. Question Paper Generation: This module generates question papers based on specified criteria, such as the number of questions and the distribution of Bloom's taxonomy levels. It reads questions from an uploaded CSV file and selects appropriate questions from the database to create a well-balanced question paper.

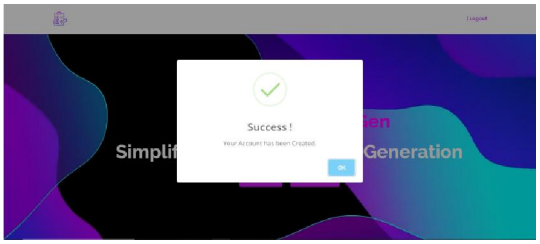
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, May 2024

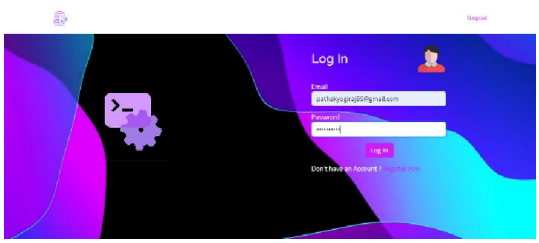
Some Projects Snapshots: -

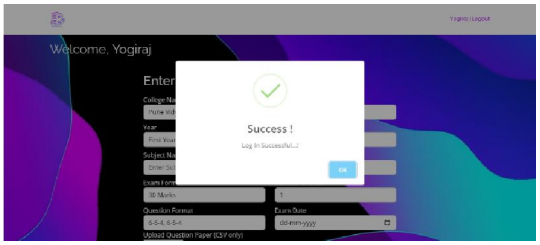


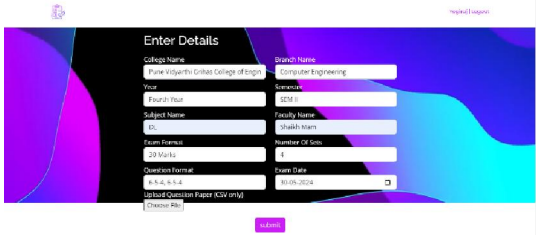


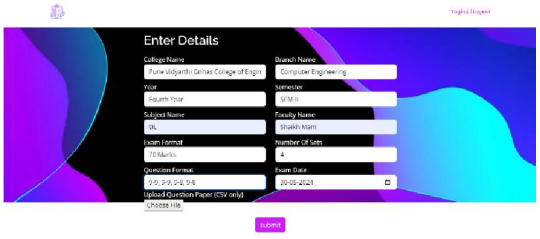


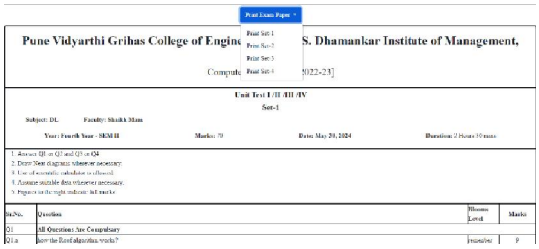


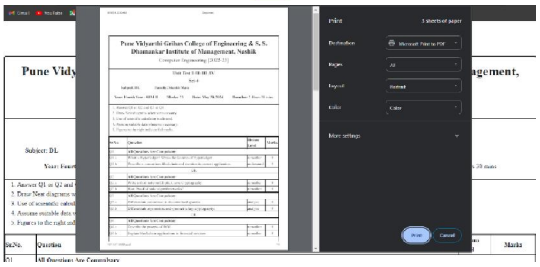


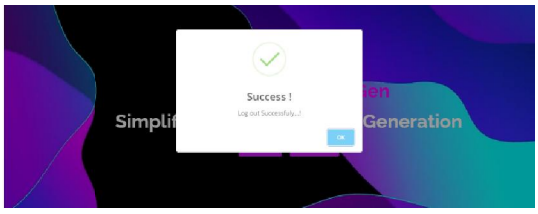




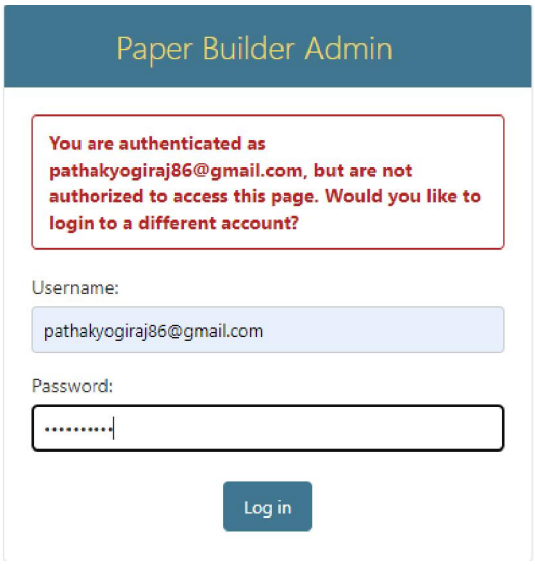


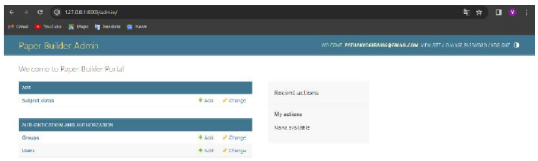






IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, May 2024





VII. SYSTEM REQUIREMEENTS

Software Requirements

To ensure the smooth functioning of the proposed system, specific software requirements must be met. The system is designed to be compatible with both Windows and Linux operating systems, providing flexibility in deployment. A web server such as Nginx or Apache is essential for hosting the system's web-based user interface, enabling users to interact with the system seamlessly. Data storage and management are facilitated by SQLite, a lightweight and efficient database solution. Python 3.7 or later version serves as the primary programming language, while deep learning frameworks like TensorFlow or PyTorch are utilized for implementing machine learning algorithms. Additionally, the system relies on various libraries such as NumPy, Pandas, SciPy, NLTK, and Keras to support different functionalities.

Hardware Requirements

The proposed system also has specific hardware requirements to ensure optimal performance. A quad-core processor with a speed of 2.5 GHz or faster is recommended to handle the computational tasks efficiently. A minimum of 4 GB RAM is necessary to support system operations and data processing effectively. Adequate storage space, preferably 1 TB or more, is essential for storing datasets and system files. Network connectivity, either through Ethernet or Wi-Fi, is required for system access and communication. Additionally, an NVIDIA GPU is highly recommended, particularly for tasks involving deep learning model training, to accelerate computations and enhance system performance. By meeting these software and hardware requirements, the proposed system can be deployed effectively, providing reliable and efficient solutions for educational assessment and question paper generation.

VIII. ALGORITHM DETAILS

Linear regression is used to predict marks based on question features like content and Bloom's taxonomy level. The model is trained on a dataset, fits a line to minimize errors, and encodes categorical features numerically. Performance

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, May 2024

is evaluated using metrics such as MSE and R-squared. CountVectorizer converts text into numerical data by tokenizing, normalizing, and building a vocabulary, resulting in a matrix of word frequencies for machine learning algorithms. Random Forest uses multiple decision trees trained on data subsets with random feature selection and bagging to improve accuracy and handle missing values. Bloom's Taxonomy categorizes educational objectives into six levels Remembering, Understanding, Applying, Analysing, Evaluating, and Creating guiding instructional design and assessment.

IX. APPLICATIONS

1. Educational Institutions: This system finds application across various educational institutions, including schools, colleges, and universities. It offers a streamlined approach to question paper creation, enhancing efficiency for educators.

2. Online Learning Platforms: E-learning platforms can leverage this system to generate quizzes and assessments for online courses. It ensures alignment with course objectives and maintains assessment integrity.

3. Corporate Training: Organizations conducting employee training and assessments can utilize this system to create standardized question papers. This ensures relevance and consistency in evaluating employee knowledge and skills.

4. Skill Assessment Programs: Government agencies and certification bodies can employ this system to develop standardized tests for assessing individuals' skills and knowledge. It ensures fairness and accuracy in certification processes.

5. Tutoring Services: Online tutoring platforms can integrate this system to offer practice tests for students. These tests aid in preparation for standardized exams or academic assessments, enhancing learning outcomes.

6. Customized Test Preparation: Students and learners can also benefit from this system by generating personalized practice tests. This supports self-assessment and exam readiness, aligning with individual learning objectives.

X. ADVANTAGES

1. Enhanced Efficiency and Time Savings: With the system's ability to generate question papers rapidly, teachers can save a significant amount of time and effort that would otherwise be spent on manual creation. This feature is particularly advantageous for educators handling multiple courses or assessments simultaneously.

2. Diverse Question Selection: The system offers a wide range of questions in each generated paper, ensuring uniqueness and reducing the likelihood of question sharing among students. This promotes fairness and integrity in the assessment process.

3. Alignent with Learning Goals: By allowing users to specify learning objectives, the system ensures that the generated questions are closely aligned with the educational objectives of the course. This alignment facilitates accurate assessment of students' mastery of the intended learning outcomes.

4. Consistent Assessment Practices: Through the utilization of machine learning and Bloom's taxonomy, the system maintains a standardized approach to assigning marks to questions. This consistency minimizes the potential for subjective grading and ensures fairness across assessments.

5. Streamlined Database Management: The system efficiently stores questions and their corresponding marks in a database, enabling easy retrieval and reuse for future assessments. This streamlined database management saves instructors valuable time in organizing and accessing assessment materials.

XI. DISADVANTAGES

1. Quality Assurance: The system's rapid question generation capability may lack assurance in producing consistently high-quality questions. Human review remains essential to ensure question validity, reliability, and contextual appropriateness.

2. Limited Creativity: Automated question generation may lack the nuanced creativity of human instructors, making it less suitable for courses requiring specialized or imaginative questions.

3. Dependency on Input: The quality of generated questions relies heavily on accurate and complete user input. Unclear or incorrect specifications may result in suboptimal question papers.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 5, May 2024

XII. CONCLUSION

Utilizing Bloom's Taxonomy as a framework for crafting robust assessment question papers, our system streamlines the process of generating and grading questions. By categorizing questions based on cognitive complexity levels, our system ensures comprehensive coverage aligned with educational objectives. Through the implementation of text mining techniques, the system adeptly organizes question papers into various problem types and difficulty levels, enhancing the quality and reliability of assessments.

XIII. FUTURE WORK

1. Our project's future scope is tailored towards enhancing examiner convenience through automated question paper generation. The application will feature a user-friendly interface enabling examiners to input parameters and swiftly generate question papers.
2. The system will be extended to support various file formats such as PDFs, Excel sheets, and Word documents for question paper templates. This flexibility empowers examiners to utilize their preferred formats, facilitating seamless generation of question papers.
3. Enhancements will be made to incorporate mathematical equations and operations into the question generation process. This addition enables examiners to include mathematical questions, thereby ensuring a comprehensive assessment of students' mathematical skills.
4. The system's interface will undergo improvements to enhance usability and accessibility for examiners. Clear navigation and intuitive design elements will simplify the question paper generation process.
5. Integration with LMS platforms will be explored to streamline the distribution and management of generated question papers. This integration ensures seamless interaction between the systems and existing educational platforms, further enhancing examiner efficiency.

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