Automatic Question Paper Generation Using Bloom's Taxonomy

[1]Anushka Chavan, [2]Avinash Lone, [3]Tejas Shejwal, [4]Nidhi Chordiya, [5]Prof. Manjusha Gaikwad

Department of Computer Engineering
MET's Institute Of Engineering
Nashik.

***Abstract*- In any educational course curriculum, the courses are defined with learning objectives. Teachers conduct assessments to know if students have achieved certain learning objectives or not. Teachers generate variety of question papers as per the universities' assessment requirements. It is very challenging for the teachers to make question papers with varied questions and which meet learning objectives of the course. There are no standardized methods to ensure quality of question paper. Hence there arises a need to have a system which will automatically generate the question paper from teacher entered specification within few seconds. Researchers recommend different sets of tags such as cognitive level, difficulty level, type of question, content /topic for defining a question etc. In this system, we proposed an autonomous question paper-generation system. In our system we allow users to input a set of questions. We also allow the user to provide complexity for each of these questions. After this, the system will assign marks to each question based on Bloom's taxonomy using machine learning and then the questions are stored in the database along with their marks.**

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

Published in IJIRMPS (E-ISSN: 2349-7300), Volume 11, Issue 6, Nove-Dec 2023

INTRODUCTION

Today businesses and governments are largely reliant on information and communication technology to communicate and making contacts. E-tendering is increasingly being adopted through the world. E-tendering in its simplest form is described as the electronic publishing, communicating, accessing, receiving and submitting of all tender related information and documentation via the internet. Thereby replacing the traditional paper-based tender processes and achieving a more efficient and business process for parties involved. The basic principles of the tendering process have been applied to many business areas, such as purchasing goods, seeking service providers, business consulting, or the selection of main contractors for construction work [1].

Inadequate security brings opportunities for fraud and collusion by parties inside and outside of the tendering process. In this paper first, a general framework for legal and security requirements for a typical e-tendering system will be identified. Secondly, the three stages of development and implementation for an electronic tendering system and security issues related to each stage will be discussed.Today businesses and governments are largely reliant on information and communication technology to communicate and making contacts. E-tendering is increasingly being adopted through the world. E-tendering in its simplest form is described as the electronic publishing, communicating, accessing, receiving and submitting of all tender related information and documentation via the internet. Thereby replacing the traditional paper-based tender processes and achieving a more efficient and business process for parties involved. The basic principles of the tendering process have been applied to many business areas, such as purchasing goods, seeking service providers, business consulting, or the selection of main contractors for construction work [1].

Inadequate security brings opportunities for fraud and collusion by parties inside and outside of the tendering

process. In this paper first, a general framework for legal and security requirements for a typical e-tendering system will be identified. Secondly, the three stages of development and implementation for an electronic tendering system and security issues related to each stage will be discussed.
Making question papers that include a variety of topics and adhere to the course's learning objectives is extremely difficult for the professors. No established procedures exist to guarantee the calibre of the test questions. Therefore, it becomes necessary to have a system that would quickly and automatically produce the test questions from the teacher-entered specifications. Researchers advise using various sets of tags to define questions, such as cognitive level, difficulty level, question type, substance or topic, etc. In this system, we suggested a system for automatically generating test questions. Users are able to enter a list of questions into our system. In today's age, education is the most important way of achieving success. When we discuss education, it is imperative to mention tests and examination. Examinations prepare students in their quest for knowledge. So, having a proper examination paper and format is quite necessary. Now the traditional method of generating question paper has been manual. In this method certain officials chalk out the question paper. But this method can be ineffective at times owing to bias, repetition and security concerns.
We have proposed an Automated process of Question Paper Generation which is fast, streamlined, randomized and secure. Every task performed by this system is automated so that storage space, bias and security is not a concern anymore. Furthermore, we have proposed a new algorithm which ensures total randomization of questions and avoids repetitions. The proposed system can be helpful to many educational institutes.

1. Aim

It is impossible to discuss education without mentioning tests and examinations. Exams help students in their quest for knowledge. As a result, having a proper examination paper and format is essential. The traditional method of producing question papers is now manual.so it is needed to have a system which can automate this manual system.

2. Objectives

1. Reduce the time: To reduce the time required for manual question paper generation. Make traditional methods seamless.
2. Automation: To automate the existing manual system with the help of computerized equipment, fulfilling their requirements, so that their valuable data/information can be stored for a longer period with easy access.
3. Reduce Errors: To reduce the error in the question paper.

PROBLEM STATEMENT

As an education is a key to success, assessment process is an essential activity in educational institutions to test performance of the learners. Teacher has to spend more time and effort before examinations. The question papers are framed by the teacher according to the syllabus. In traditional system, paper Setters/Examiners used to make/set question paper manually. This is very time consuming process. While making the question paper, sometimes incorrect marks could be assigned by the examiner. To overcome this problem, we are proposing our system. One of the most popular guidelines used to write and evaluate the exam questions is Bloom's Taxonomy (BT). This rule-based approach applied Natural Language Processing (NLP) techniques to identify important keywords and verbs. To classify, the questions Na¨ıve Bayes algorithm is used. This system can assign marks to the questions and generate question paper as per format.

LITERATURE SURVEY:

" **Exam Question Classification Based On Bloom's Taxonomy: Approaches and Techniques**" (Karima Makhlouf, Lobna Amouri, Nada Chaabane and Nahla EL-Haggar)2020 IEEE Eighth International Conference.[1]
Shown a comprehensive review on the commonly used approaches and techniques for exam Questions Classification Based on Bloom's Taxonomy. It has shown multiple ways to accomplish a goal. Some techniques are hard to implement. Since the past recent years, many researchers have worked on the automation of exam question classification into BT. This paper gives a through overview of the standard approaches and techniques used in question classification (QC).

”**Automatic Generation of Question Paper from User Entered Specifications Using a Semantically Tagged Question Repository**,” (G. Nalawade and R. Ramesh)2016 IEEE Eighth International Conference on Technology for Education (T4E), 2016, pp.148-151[2]

There are no standardized methods to ensure quality of question paper. Hence there arises a need to have a system which will automatically generate the question paper from teacher entered specification within few seconds. Researchers recommend different sets of tags such as cognitive level,difficulty level, type of question, content /topic for defining a question etc.The existing tools are rigid and support very basic or limited tags. The proposedsystem will automatically generate a question paper from semantically tagged question repository. This system offers flexibility by supporting all four tags and allows entry of every property in the form of ranges i.e. lower bound and upper bound. The question paper is generated in xml format and as Microsoft Word document.

”**An algorithm for question paper template generation in question paper generation system**,”(V. M. Kale and A. W. Kiwelekar) 2013 The International Conference on Technological Advances in Electrical,Electronics and Computer Engineering (TAEECE), 2013, pp. 256-261 [3]

In this paper, we present the design of an algorithm to generate question paper template which satisfies the above mentioned constraints. The algorithm is illustrated in the paper by using four constraints namely question paper format, coverage of syllabus, coverage of difficulty levels and coverage of cognitive levels according to Bloom’s taxonomy. The algorithm presented is extensible to support any number of user defined constraints.

”**Neural Question Generation Using Question Type Guidance**,” (R. Sun, X. Zhou and F. Fang)2021 17th International Conference on Computational Intelligence and Security (CIS), 2021, pp. 328-332, [4]

In this paper, author focus on how to exploit question type to guide the question generation task. The question type prediction task is introduced into our multi-task framework to discover the underlying relationship between the context-answer pair and the question. In additional, author also applies metric learning mechanism to improve the semantic relevance between the generated question and the context. The model achieves a better performance than the baseline systems on SQuAD benchmark. Experimental results demonstrate that the guidance of question type is significant for question generation.

SYSTEM ARCHITECTURE

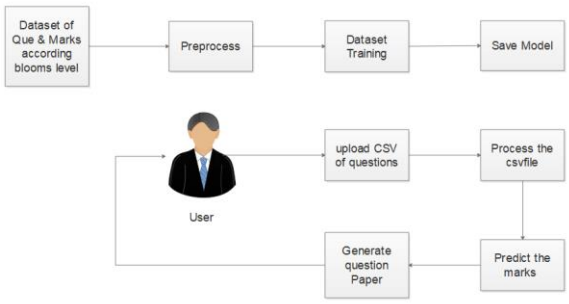

Fig -1: System Architecture Diagram

Algorithm

1. Start
2. Initialization
3. Dataset creation & allocation of marks according to blooms level
4. Preprocess
5. Data Training Process
6. Saving the model for further utilization
7. User login
8. User will upload the CSV of questions
9. Processing the CSV file in system

10. Predicting the marks
11. Generating the question paper
12. END

Bloom's Level Explanation

a) Knowledge-level:

Also known as recalling of data (Bloom, 1956). Scott (2003) refers it as 'rote learning' or 'memorization'. This level serves as the beginning level of the hierarchy). It is a level where students memorize facts or recall the knowledge they learn before. The questions for programming in this category have the criteria of recalling specific input from previous lessons, explaining describing computing terms, methodology and process, stating relevancy description for a subject area, concept or term and listing explicitly information from questions (Scott, 2003). Examples:

a) List all the nodes in the leftsubtree of node J.
b) Describe the key properties of abinary tree.
c) Define method in JAVA.

b) Comprehension-level:

Bloom (1956) describes this level as grasping the meaning of information. The capacity to interpret, translating, extrapolating, classifying, explaining are the concepts of these levels. The questions for programming (Thompson et al., 2003) in this category could be translating algorithm (e.g.; write output of a program), explaining the processes and flows of program and providing examples to illustrate a concept or an algorithm. Examples: a) What is the outputof the following code segment? b) Define in words what happens in the following C++ code.

Application-level:

Application is defined by applying the concept to a certain scenario (Starr et al., 2008). The questions for programming in this category have the following criteria: understand the concept and use it to a new algorithm and modifying controls. Examples: a) Declare a variable, employees to represent the records of 120 employees. b) Modify the given 'for' loop into 'while' loop.

c) Evaluation-level:

This is a final level where judging, criticism, supporting or defending own stand involves. Thompson et al. (2008) discuss this level in Bloom's Taxonomy for CS Assessment.[11] According to them, programming question is interpreted by checking codes if the code fits the requirement for testing strategy. This level also includes discussion quality of codes based on standards or execution criteria. Example: a) Explain the concept of inheritance and give the sample of code to illustrate your answer.

d) Synthesis-level:

If a student achieves this level, the student should be able to synthesis and associate ideas or concepts by rearranging components into a new whole (a product, plan, pattern or proposal) (Bloom, 1956). Scott (2003) suggests programming questions for this level should instruct student to write codes based on previous level by writing a complete program or create new alternative methods or algorithm to solve a problem. Examples: a) Write the definition of the function Output Time if the statements from the lines 22 to 34 were to be performedin a function. b) Write a program that helps the user to input the masses of the bodies and the distance between the bodies. The program then outputs the force between the bodies.

e) Analysis-level:

This level requires students to breakdown information into simpler parts and analyse each of it. This may signify drawing a relationship, assumptions, distinguish or classifying the parts. According to Thompson et al. (2008), programming questions should contain the following: subdivide programming algorithm into classes, methods; establish elements to achieve objective; recognize components of a development and distinguish non related factor or needs. In addition, it should be able to explain what absolutely happens to

memory when the codes are executed line by line. Examples: a) Outline how class Book List could be implemented using an array. b) Given the following postfix notation: 12 9 + 9 *5 3 / - = Using the stack implementation to evaluate the above postfix

ALGORITHM/TECHNOLOGY

Pre-Processing:

Text pre-processing is a method in natural language processing to prepare the computer understand the structure or content of the text. It will allow us to prepare the text more precise and easy to use for later process. Text preprocessing associate processes such as *stopwords removal, stemming, lemmatization and POS tagging*. In this work, stopwords removal is applied to the question in order to make the text more readable for later process. Following this, each and every word will then be tagged using a tagger. In this research, NLTK tagger (Bird et al., 2009) is used to tag the exam questions. To emphasize the tagging process, consider the following sentence:

"Outline how class ArrayList could be implemented usingan array.",

The tagged output is: Outline/VB how/WRB class/NN ArrayList/NN could/MD be/VB implemented/VBN using/VBG an/DT array/NN./.

The tagger will assist to identify important nouns and verbs, which may be important in determining the question's category. In addition, the sentence pattern may help in the correct recognition of the question's category. After tagging, some rules will be applies according to question's structure.

Rules Development:

Through this research, a rule-based approach is accepted in determining the category of an examination question based on the Bloom's taxonomy. The rules are developed from a training set which consists of 70 examination questions in the various subjects. There are two conditionswhere the rules will be applied:

- The rules will distinguish the suitable keywordfor each question depending on its category.
- (elp to choose the appropriate category if the keyword shares more than one category. For example, Abstract may fall under Comprehension or Synthesis category.

After analysing all the questions in the training set, the questions' patterns show that most of them start with a verb. (however, only some of it begins with Wh-pronoun, a determiner, preposition or subordinating conjunction, anoun and an adverb. Before rules can be applied, specific patterns should be identified from the questions item. The following will determine how the pattern and rules are developed after POS tagging is applied. Question: Write down the output of the following program:

Question with tag: Write/VB down/RB output/VB following/JJ program/NN :/:

Pattern: /VB (1st word), /VB (3rd word)
Each verb in the question will be captured. The verb 'Write' appears as its keyword. Based on Bloom's Taxonomy, Write can be categorised into two categories: Knowledge and Synthesis.

FOR each sentence, read into an array. Split into words.

)IF pattern is found
)If the keyword "Write" is found
)IF found:
Apply Rule1: Assign weight toKnowledge
Apply Rule2 : Assign weight toSynthesis Choose the greater value
or positive value Assign questioncategory Store in database

FOR EACH(_match in pattern :
print join (keyword, category, question)

Based on the algorithm, the question can be applied to two different rules i.e. Rule 1 and Rule 2. Rule 1 states that the questions fall under the 'Knowledge' category meanwhile Rule 2 states that it can be categorized under the 'Synthesis' category. This raises a conflict as to which category the question should fall into. When this situation occurs, there is a need to introduce 'category weighting' to altering all subsequent blocks.

Randomization Algorithm

For N questions available in database Step 1: Create a List 'L' of N elements
Step 2: Generate a random number 'n' such that 1<=n Step 3:)f n ∈ L
Go to Step 2
else Store n in the List L
Step 4: Select a question from database corresponding to n, whose flag==true
Step 5: For the question, set flag=false

CONCLUSION

Using Bloom's Taxonomy for creating valuable and reliable assessment question papers, we have proposed a system in this work that generates question papers and assigns marks to each question. According to Bloom's Taxonomy, our system can read and categorize those questions, then place them in different levels of class. It is expected that our proposed system should be able to correctly assign the marks to the question and also able to generate multiple question papers using Bloom's Taxonomy concept and machine learning.

REFERENCES:

1] "Exam Question Classification Based On Bloom's Taxonomy: Approaches and Techniques"(Karima Makhlouf , Lobna Amouri , Nada Chaabane and Nahla EL-Haggar)2020 IEEE Eighth International Conference

[2] G. Nalawade and R. Ramesh, "Automatic Generation of Question Paper from User Entered Specifications Using a Semantically Tagged Question Repository," 2016 IEEE Eighth International Conference on Technology for Education (T4E), 2016, pp. 148-151.

[3] 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.

[4] 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.

[5] 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.

[6] Q. F. Cheng, H. T. Liu and X. M. Yang, "Design of Test Paper Strategy Based on Randomized Algorithm", Science Mosaic, vol. 34, no. 11, pp. 107-110, 2011.

[7] Z. C. Wang and X. Y. Ouyan, "Design of Intelligent Test Paper Generation System Based on Improved Genetic Algorithms", Computer Knowledge and Technology, vol. 16, no. 2, pp. 65-67, 2020.

[8] B. Zhao, H. Ning and R. B. Zhang, "Research and implementation of automatic test paper generation system based on genetic algorithm", Applied Science and Technology, vol. 48, no. 2, pp. 50-5, 2021.

[9] J. Zhang, "Online Examination System of Automatic Paper Composition Based on Hybrid Algorithm", China Computer Communication, vol. 33, no. 7, pp. 97-99, 2021.no. 2, pp. 45-54, 1988.