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 1, April 2024

Automatic Question Paper Generation and Weightage Assigning using Bloom's Taxonomy

Prof. B.S Chaudhary, Janhvi Jagtap, Minakshi Bhavar, Vaishmavi Wagh, Kiran Pal
Department of Computer Engineering
Loknete Gopinathji Munde Institute of Engineering and Research Center, Nashik, India
bhushan7887@gmail.com, janhvijagtap2005@gmail.com, minakshibhavar91@gmail.com,
vaishnaviwagh234@gmail.com, pal070226@gamil.com

Abstract: *Bloom's Taxonomy is a classification of learning objectives within education that educators set for students. The cognitive domain within this taxonomy is designed to verify a student's cognitive level during a written examination. Educators may sometimes face the challenge in analysing whether their examination questions comply within the requirements of the Bloom's taxonomy at different cognitive levels. This paper proposes an automated analysis of the exam questions to determine the appropriate category based on this taxonomy. This rule-based approach applies Natural Language Processing (NLP) techniques to identify important keywords and verbs, which may assist in the identification of the category of a question. This work focuses on the computer programming subject domain. At present, a set of 100 questions (70 training set and 30 test set) is used in the research. Preliminary results indicate that the rules may successfully assist in the identification of the Bloom's taxonomy category correctly in the exam questions.*

Keywords: NLP, Bloom's Taxonomy.

I. INTRODUCTION

With the advancement of technology in the education sector, there is a growing need for automating various processes to make them more efficient and effective. One such process that can benefit from automation is the generation of question papers for assessments and exams. Bloom's taxonomy is a framework that categorizes different levels of learning objectives, from basic knowledge recall to higher-order thinking skills like analysis and synthesis. By using Bloom's taxonomy as a guide, educators can create well-rounded question papers that assess students' understanding and application of concepts at various levels. Automatic question paper generation using Bloom's taxonomy involves using software tools to analyze learning objectives, curriculum content, and student performance data to generate customized question papers tailored to individual student needs.

Blooms taxonomy:

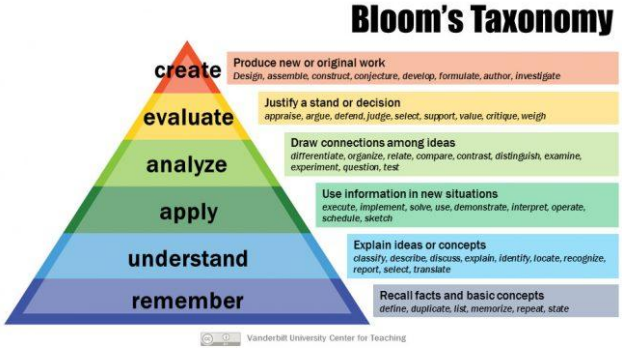


Fig1:- Blooms Taxonomy

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 1, April 2024

This approach ensures that the assessments are aligned with the learning objectives and provide a comprehensive evaluation of students' knowledge and skills. By automating the question paper generation process, educators can save time and effort in creating assessments, while also ensuring that the questions are of high quality and effectively measure students' learning outcomes. This approach can also help in reducing bias and ensuring fairness in assessments by providing a standardized and objective method of evaluation. automatic question paper generation using Bloom's taxonomy is a valuable tool for educators to create meaningful assessments that promote student learning and achievement. By leveraging technology in this way, educators can enhance the assessment process and provide valuable insights into students' progress and performance.

Uses advanced algorithms to generate exam questions based on Bloom's Taxonomy that are relevant to the course content and aligned with learning outcomes. Incorporates a diverse range of question types, such as multiple-choice, short answer, and essay questions, to ensure the comprehensive evaluation of students' knowledge and skills. Allows educators to customize and modify generated question papers according to their preferences and curricula. Provides an intuitive user interface that allows educators to generate question papers efficiently and quickly. Includes a knowledge base that is updated regularly with relevant and accurate information to generate high-quality questions. Is reliable, secure, and scalable to handle question papers for various courses, classes, and academic levels. Offers data analysis and reporting capabilities to enable educators to evaluate the effectiveness of the question papers and the assessment of student learning. project's scope is to provide an efficient, effective, and flexible solution for generating exam questions using Bloom's Taxonomy that saves time and enhances the quality of student learning and assessment.

Normalization of the Data:

The process of automatic question paper generation can be normalized using Bloom's taxonomy by ensuring that questions cover a range of cognitive levels as defined in the taxonomy.

Bloom's taxonomy categorizes cognitive skills into six levels:

1. Remembering: This level involves recalling facts or information.
2. Understanding: This level involves comprehending the meaning of information.
3. Applying: This level involves using knowledge to solve problems or apply concepts.
4. Analyzing: This level involves breaking down information into parts and examining relationships between them.
5. Evaluating: This level involves making judgments or decisions based on criteria and standards.
6. Creating: This highest level involves generating new ideas, products, or ways of viewing things.

To normalize the automatic question paper generation process based on Bloom's taxonomy, each question generated should be mapped to one of these cognitive levels to ensure a balanced distribution across all levels within the question paper.

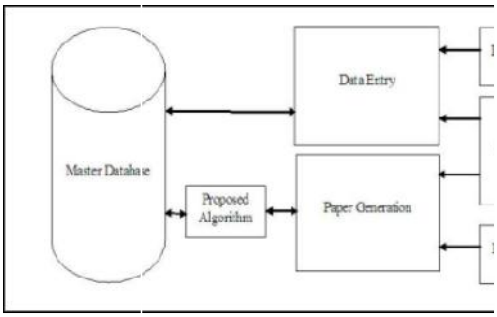
For example, when creating an automatic system for generating questions for a test in mathematics:

- Remembering - A simple multiple-choice question asking students to recall a mathematical formula
- Understanding - A problem-solving scenario that requires students to demonstrate understanding of a concept
- Applying - A real-life application problem that requires students to use mathematical principles in a new context
- Analyzing - A question that asks students to break down a complex problem into its component parts and identify patterns
- Evaluating - An open-ended question where students must analyze different solutions and justify their choice
- Creating - A challenge where students have to devise an original mathematical problem for their peers

By normalizing the distribution of questions across these cognitive levels, educators can ensure that the assessment covers various aspects of learning and provides valuable insights into student capabilities beyond simple recall.

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 1, April 2024

II. METHODOLOGY



Consider the block diagram from Fig1.1

Bloom's Taxonomy is a classification of learning objectives within education that educators set for students. The cognitive domain within this taxonomy is designed to verify a student's cognitive level during a written exam- ination. Educators may sometimes face the challenge in analysing whether their examination questions comply within the requirements of the Bloom's taxonomy at different cognitive levels. This paper proposes an automated analysis of the exam questions to determine the appropriate category based on this taxonomy. This rule-based approach applies Natural Language Pro- cessing (NLP) techniques to identify important keywords and verbs, which may assist in the identification of the category of a question. This work fo- cuses on the computer programming subject domain. At present, a set of 100 questions (70 training set and 30 test set) is used in the research. Preliminary results indicate that the rules may successfully assist in the identification of the Bloom's taxonomy category correctly in the exam questions.

Data mining: Data Mining is the computing process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics and database systems.

Blooms Taxonomy: Blooms Taxonomy is a classification of the different objectives and skills that educators set for their students (learning objectives). The taxonomy was proposed in 1956 by Benjamin Bloom, an educational psychologist at the University of Chicago.

Natural Language Processing: The field of study that focuses on the interactions between human language and computers is called Natural Language Processing, or NLP for short.

Randomized Algorithm: A randomized algorithm is an algorithm that employs a degree of randomness as part of its logic. Formally, the algorithm's performance will be a random variable determined by the random bits; thus either the running time, or the output (or both) are random variables.

2.1 Objective

1. To make question papers with varied questions and which meet learning objectives of the course.
2. To generate the question paper from teacher entered specification within few seconds.
3. To cover all aspects of the course objectives and avoid duplication of questions in the subsequent exams.

III. LITERATURE SURVEY

The existing Learning Management Systems (LMS) support very basic level or limited tags such as question types . Even the most preferred LMS, Moodle allows creating only subjective/objective type of questions. Thus automatically generating question paper from a teacher's entered specification using a semantically tagged QR is the need of the hour today. The system to semi automatically tag the questions of a repository is in place. Here we are proposing a system which automatically generates the question paper from this semantically tagged question repository. Since the existing systems are rigid and lack the flexibility of supporting all types of tags, the generated question paper may not be totally aligned with its given objectives. Our system supports all four tags and also flexible enough to provide an interface that allows user to enter specifications for each tag/property in the form of lower and upper bounds. Each property is specified with a range indicating that value should not be lower than minimum value and not exceed the maximum value of the range. Also, it is rule base system which takes all the combinations of the tags and generates output based

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 1, April 2024

on the rule applicable. The output is generated in xml format and in word document. A literature survey was started to understand the need for automatic generation of QP. As mentioned in, many existing LMS support tagging feature but users may not utilize this feature fully. The comparative study shows that Moodle is best LMs to support large number of users and also for any educational institution. But it allows user to define only question type. Hence the questions in the repository may have only basic tags or no tags at all. So it becomes overhead for teachers to tag these questions before using them. Properly tagged questions can be efficiently retrieved from repository. Hence it is very much necessary to tag the questions before adding them to repository. Most teachers and researchers recommend four types of tags such as topic, question type, cognitive level and difficulty level. A system which offers generation of question paper using user given input parameters considers only fixed range of values. Our system not only supports upper and lowers bounds for inputs but also supports more granular level of topics than chapters and more question types as compared to only three types offered by this system. We are using automatically tagged question repository as input instead of untagged questions. Automatic Question Paper Generator System described in has great interface, but supports only question type tag. Hence the question paper generated would have only one difficulty level

3.1 Existing Paper-based Systems

The existing system for Question Paper Generation requires human staff to chalk out questions that appear in the question paper. These teachers or professors select the questions according to the syllabus and pattern as prescribed by the curriculum. The question paper then may be referred to a higher authority who has the final say in these matters. B. Limitation of Paper-based Systems As most human working processes, this system suffers due to bias. There might be some questions which are repeated in many question papers as the professor has a personal inclination towards them. So there is no guarantee of pure randomly generated question paper. Other problems that may plague this system are non-availability of staff and resources, natural calamities and accidents. Also, the security of the system can be easily compromised if leverage over the person responsible for generating question papers is obtained. Other limitations include: -

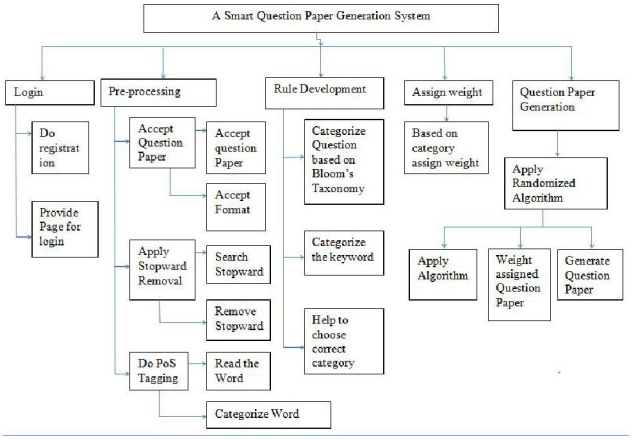
a) Lack of storage space
b) Prone to damage
c) Inefficient document transportation
d) Supply costs
e) Poor environmental credentials
f) Limited collaboration
g) Editing problems

3.2 Existing Question Paper Generation Systems

The research paper „Framework for Automatic Examination Paper Generation System" has provided a thorough insight into the process of automated paper generation. As the manual generation of a balanced question paper by an individual is quite complex, the blending of technology into teaching and learning process is inevitable. A simple and efficient way for an examination paper generation is provided. A three tier model is provided in this framework. Generation of Examination Papers is governed by the Syllabus Engine, Pattern Composer and Question Aggregator. The generated question paper is based on the pattern or skeleton of the course. Another component called Bank Management takes care of User Rights and Privilege assignment. Questions are entered through the Question Aggregator. The attributes related to questions are type, marks and complexity. All these attributes are efficiently used during Question Paper Generation. The paper generator selects a question according to the pattern and complexity. This engine also introduces marking systems wherein any selected question is marked so that it might not be selected again. This prevents repetition of questions in subsequent papers. Finally, generated papers are stored as pdfs.

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 1, April 2024

IV. ARCHITECTURE DIAGRAM



First the authorized user will register, after the successful registration the user will be provided by the login-Id and password. After logging-In the user can input the text file either by manually
entering it or by providing the path of the file. After the input is provided, the questions will be
generated accordingly. Questions can be reviewed and manipulated by the user. A question paper can be generated by selecting desired questions generated by the software. After the generation of question paper, there will be two options: first it will be emailed to the authorized user or admin and the secondly it will be downloaded as a file.

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 IJIRMPS230356 Website: www.ijirmps.org Email: editor@ijirmps.org 4
10. Predicting the marks
11. Generating the question paper
12. END

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.

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 1, April 2024

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 twodifferent 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 databaseStep 1: Create a List 'L' of N elements

Step 2: Generate a random number 'n' such that 1<=nStep 3:)f n ∈ L Go to Step 2

else Store n in the List L

Step 4: Select a question from database corresponding ton, whose flag==true

Step 5: For the question, set flag=false

V. CONCLUSION

Bloom's Taxonomy is a classification of learning objectives within education that educa- tors set for students. To automate the process of categorizing examination question according to Bloom's Taxonomy based on its cognitive levels. The formation of rules may improve the accuracy of the result .In this system smart model for question paper generation will be im- plemented as real time application. The proposed work describe a smart system that progress from the traditional method of paper generation to an smart process by providing the control access to the resources.

REFERENCES

[1]. kithsiri jayakodi,madhushi bandara,Indika pereca.An automatic classifier for exam qua- tions in engineering :A process for Blooms taxonomy. IEEE international conference on Teching,Assessment and learning for Engineering(TALE).PAGE:195-202,2015.

[2]. Swart,Author James. Evaluation of final examination papers in engineering : A case study using Blooms Taxonomy . Education, IEEE Transaction on 53,no.2,p.257- 264,2010.

[3]. Lister, R., and Leaney, "Introductory programming,criterian-referencing , and Bloom.proceedings of the 34th SIGCSE technical symposium on computer science education",Reno,Nevada,USA.ACM Press,2003.

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 1, April 2024

[4]. Yusof,Norazah, Chai Jing Hui.Determination of Blooms cognitive level of question items using artificial neural network. In intelligent system design and application (ISDA),2010 10th International conference on,pp. 866-870,2010.

[5]. Ibrahim EldesokyFattoh ,"Automatic Multiple Choice Question Generation System for Semantic Attributes Using String Similarity Measures in journal Computer Engineer- ing and Intelligent Systems" www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online) Vol.5, No.8,2014

[6]. Ahmed EzzAwad and Mohamed Yehia Dahab,"Automatic Generation of Question Bank Based on Pre-defined Templates",in International Journal of Innovations Ad- vancement in Computer Science IJIACS ISSN 2347 8616 Volume 3, Issue 1 April

[7]. Ms.Deepshree S. Vibhandik, Prof. Mrs.Rucha C. Samant,"An Overview of Automatic Question Generation Systems",in International Journal of Science, Engineering and Technology Research (IJSETR), Volume 3, Issue 10, October 2014 .

[8]. Bloom, Benjamin S.Longman," Taxonomy of educational objectives: The classification of education goal .Cognitive domain. Handbook",1956.

[9]. Ibrahim EldesokyFattoh,"Semantic Based Automatic Question Generation using Artifi- cial Immune System",in Computer Engineering and Intelligent Systems www.iiste.org ISSN 2222-1719 (Paper) ISSN 2222- 2863 (Online) Vol.5,No.8,2014

[10]. AnnamaneniNarendra, Manish Agarwal and Rakshit shah LTRC,"Automatic Cloze- Questions Generation", IIIT-Hyderabad, India,Sept 2013.

[11]. Yvonne SKALBAN , Le An HA , Lucia SPECIA , Ruslan MITKOV,"Automatic ques- tion generation in multimedia-based learning",Dec-2012.

[12]. Manish Agarwal, Rakshit Shah and PrashanthMannem,"Automatic Question Genera- tion using Discourse Cues",June-2011.

[13]. MingLiuandRafaelA.Calvo,"G-Asks:AnIntelligentAutomaticQuestiongeneration System for Academic WritingSupport",2012.

[14]. Andreas Papasalouros, KonstantinosKanaris, KonstantinosKotis,"Automatic genera- tion of multiple choice questions from domainOntologies",July-2008.