AUTOMATIC QUESTION GENERATION SYSTEM

Prof. D. R Nanda Devi[1], Yaramada Akshara[2], Sindhuja Nadgoud[3], Veda Charitha Bandikallu[4]

[1]Assistant professor, MTECH, Department of Computer Science and Engineering, G Narayanamma Institute of Technology and Science (For Women), Telangana, India
[2, 3, 4]Final Year Student, Department of Computer Science and Engineering, G Narayanamma Institute of Technology and Science (For Women), Telangana, India

---

Abstract - *This paper focuses on automatic generation of all possible questions from a content of interest. The main thing is to induce fact- grounded questions about a given content from its associated information. Queries are generated by exploiting the named existent data, the keywords and thus the predicate argument structures of the sentences (along with linguistics rules) present within the given body of textbook. The named entities and the semantic tags are used to identify applicable corridor of a judgment in order to form appropriate questions about them. This paper aims to automate the process of question generation grounded on question templates and Bloom's Taxonomy using Natural Language Processing. The end is to produce an interface for the user to generate questions.*

***Key Words*: Bloom's Taxonomy, Text Blob, NLP, Grammar Semantics, Linguistic Rules, NLTK, POS tagging**

1. INTRODUCTION

The manual generation of questions is a tedious process in all educational institutions. One of the basic works of the teachers would be preparing questions of various levels for every exam. This is a time-consuming work. Examinations play a vital role in assessing a student's performance. Having a smart question generation system enhances the growth of students. Design of the question paper should be such that it encourages the students by building up one's levels of understanding. Bloom's taxonomy is a hierarchy that defines the classification of learning methods in schools and colleges[4]. Generally, Bloom's taxonomy has six levels of classification. It includes create, evaluate, analyze, apply, understand and remember.

2. EXISTING SYSTEM

2.1 Paper-Based Question Generation System

Manual preparation of question paper is known as paper-based system. The existing system for Question Paper Generation requires human staff to chalk out questions that appear in the question paper. Teachers or professors select the questions according to the syllabus and pattern of paper as prescribed in the Bloom's Taxonomy. This is a very tedious process.

2.2 Existing automated Question Generation System

Automatic question paper generation has been a field of interest for numerous researchers and a lot of research has been done for question paper generation system based on Bloom's Taxonomy. The reason behind the interest is basically the time consumed in generating the paper manually. Some of the existing systems are algorithms to shuffle and randomize questions from the existing database, which can be favorable to students and may not test their skills entirely. In a system called "Design of Adaptive Question Bank Development and Management System", which was an adaptive system but the data entered is assumed to be error free which could affect the overall precision of the system.

Another system tried to deploy a question generation system with Multiple Choice Questions (MCQs) and Fill in the Blanks (FIBs) which randomly pick a sentence and weigh the important keywords from it and convert it to a MCQ/FIB[5]. The major disadvantage of this system was that the choices given in MCQs were too obvious or the keyword weighing was not appropriate.

3. PROPOSED SYSTEM

The proposed system aims at providing questions of all levels on any type of given input text through a GUI to the user where the user can select whichever question, he/she wants to and select the required questions to download in excel format. This system is aimed at Grammar Rule based question generation where various topics are mapped to linguistic rules[1]. The appropriate keywords from blooms taxonomy level hierarchy are incorporated into the questions.

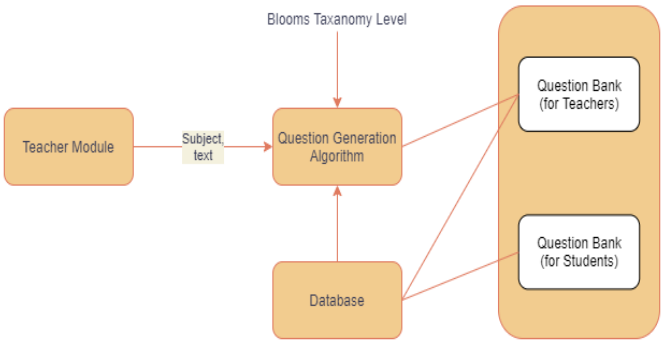

Fig -1: Architecture of proposed system

Fig -1 Architecture of proposed system; exposes the structure of the system. It displays the Teacher module, Algorithm module, Database module and the Question Bank module. An input paragraph is sent from the teacher module to the question generating algorithm, where the questions are generated based on bloom's taxonomy. The database module consists of previous year questions which can be accessed by both teacher and student modules.

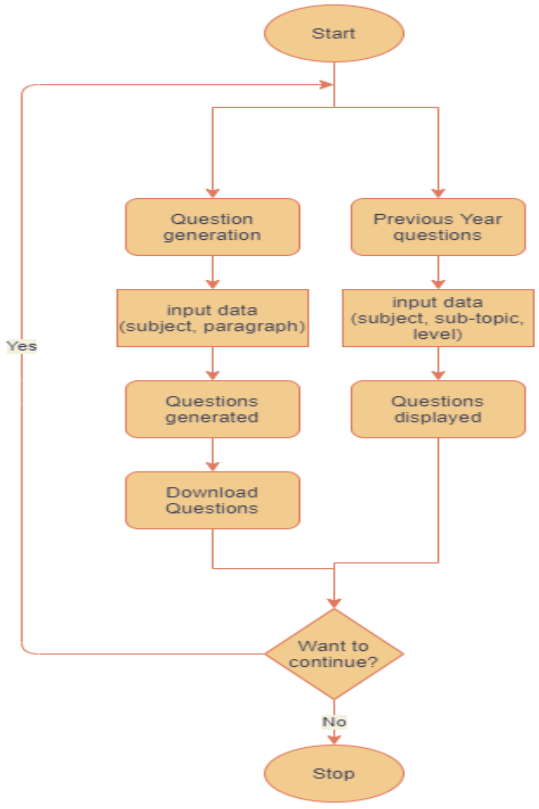

Fig -2: Flowchart of the proposed system

As per the Fig -2 Flowchart of the proposed system, user can either generate questions or access previous year questions. To generate questions, user has to select the subject and input a paragraph. After the questions are generated, user can select the required questions and download them in an excel sheet. To access previous year questions, user has to provide name of the subject, sub-topic and the level. All the questions filtered according to the requirements are displayed.

3.1 Implementation

The algorithm begins with identifying the input text given by the teacher via teacher module and generate questions from that text. The text is tokenized using parts-of-speech tagging and the stop words in the text (ex: is, her, '', etc.) are removed using the NLTK.stopwords module.

A text blob object is created for every line in the text and then segregate the words based on Parts of Speech Tagging into Nouns, Verbs, Adjectives, Pronouns, Prepositions and Adverbs with sub-categories like NNP, NN, VBG, etc. Now grammar rules are written for keywords to make the questions syntactically correct and the words are picked from each POS (Parts of Speech) bucket in order to generate the questions.

4. RESULTS

The current application generates the Bloom's taxonomy questions up to three-levels (create-L1, evaluate-L2, analyze-L3). The application allows the Users to choose a topic and level, based on their requirement of questions[6]. Semantics of the question is preserved by the usage of grammar rules whereas the syntax is preserved by using Subject Verb Predicate form.

Question	Level
Explain different testing startegies you apply?	L2
Explain overall view of testing strategy?	L2
What do you understand by the word testing?	L2
Desribe verification.	L2
Explain validation.	L2
Identify various misconceptions of testing?	L2
Examine the parameters for good testing?	L2
Disucuss different types of testing.	L2
Identify various misconceptions of testing?	L2
Explain Unit Testing in Detail.	L2

Fig -3: Sample Generated Questions

Fig -3 displays the questions generated based on the user input, where the subject is given as software engineering, topic as testing and level as L2- "Evaluate".

5. CONCLUSION

An automated question generation system focuses on Generation of Questions based on Bloom's Taxonomy. It significantly reduces human involvement by processes automation and helps instructors to spend less time on question paper design. This system allows for rapid data retrieval and manipulations to generate questions on the click with minimum effort. On top of that, the system offers

concise storage of questions, question bank can cover wide range of subjects and question types.

REFERENCES

[1] C. Unger and P. Cimiano," Pythia: Compositional meaning construction of question answering on the semantic web", 2011.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] Mona Nabil Demaidi, Mohamed Medhat Gaber & Nick Filer, Evaluating the quality of auto-generated questions, 2017.

[3] Shiyan Ou, Constantin Orasan, Dalila Mekhaldi, Laura Hasle, Automatic Question Pattern Generation for Question Answering, 2010.

[4] Shubham Dhainje, Renuka C, Komal Borse, Visha Bhamare, Automatic question generation, 2017.

[5] Tahani Alsubait, Bijam Parsia and Ulrike Sattler, Multiple Choice Question Generation, 2016.

[6] Ylias Chali & Sadid, Topic-to-Question Generation, 2015.