



IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2320-6710



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 7, July 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

Automatic Question Generation from Text Using Machine Learning

Kunal Paudmal, Rushikesh Bhagwat, Rohan Bhapkar, Krishna Jadhav, Prof. Aparna Hambarde

Department of Computer Engineering, KJCOEMR, Pune, India

Abstract: Curiosity is the fuel for discoveries and learning, and we humans are always curious about learning novel things. Asking different types of questions is a tendency of a good learner and so the students ask a variety of questions from their teachers. Similarly to examine the students, the teachers also ask questions from the students and also evaluate their performances. Because of their inconsistent minds in certain situations, sometimes humans are not very skilled in asking good questions and because of that we have come with the idea of a system with the help of which we would be able to generate the questions from a text automatically. The system is known as the Automated Question Paper Generator System, which is a fast and secure system and which can also generate questions in a random way. The document or pdf file or simply text can be provided as an input to this proposed system. NLP technology is used for preparing the proposed system.

KEYWORDS: Question Paper Generation, stop words, NLP, Key phrase Extraction,

I. INTRODUCTION

Traditionally question papers were generated manually. Preparing any exam paper is a very challenging task for the educators because they have to check whether there is any repetition in the pattern or not, and other than that security is one of the major concerns for them, also due to lack of teaching staff in any institute, creation of paper is not at all an easy task. So, here we are proposing an Intelligent Automatic Question Paper Generator System which provides storage of the data, fast operations, and high security for all its tasks. It can be helpful to many educational as well as NGO institutes. Tasks performed by this system

Are automated and because of that storage space, security is not a concern anymore. Proposed system works upon Natural Language Processing and is fast due to computer based automation, streamlined, randomized and unbiased, secure and generates unique questions and overcomes the problems of Human based paper generation.

Qualities of a good question

- **Appropriate:** A good question is appropriate. It focuses on memorizing only the content included in your lesson and aligns well with the overall learning objectives.
- **Concise :** A good question is usually concise and crisp. This leaves out any unnecessary information that students have to spend time understanding correctly.
- **Understandable :** A good question is formulated in clear, easily understandable language without any ambiguity. Students should understand what the question requires, even when they do not know the answer.
- **Purposeful :** A question without a specific purpose has no value. The purpose helps evaluate the question on the basis of some defined benchmarks. A good question can seek both innate and specific knowledge.
- **Guiding But Not Leading :** A good question guides the students to understand the concept in the image. But at the same time, it does not give them specific answers.
- **Stimulates Thinking :** A good question requires learners to think and remember through the concepts taught. It does not protect them by asking explicitly.
- **Single-Dimensional :** A good question focuses on one dimension at a time. If there are multiple ideas to evaluate, it is better to split them into multiple questions.

Difference Between Manual Process And Automated Process Of Generating Question Paper

Manual Paper Generation	Automated Paper Generation
It is a human process.	It is an automated process.
In this process repetition in paper may occur.	It is a random as well as unbiased process
Less Security.	Higher security.
It is a slow process.	It is a faster process.
It has less varieties of questions.	It has more varieties of questions.

II.LITERATURE SURVEY

Automatic question paper generation has been an area of interest for many researchers and a lot of research has been done for the question paper generation system. Researchers are very interested in this field and they are showing their interest too. The reason for the interest is basically the time taken by the professor to write the thesis. And every researcher has undergone the university processes. AUTOQUEST was one of the first automated QG systems proposed to support novices to learn English. Kunichika et al, who proposed a method for generating questions based on both syntactic and semantic information (space, time, and agent) so that he could generate more types of questions (where, when and who). Most recently, Mostov and Chen proposed an approach for generating deep questions based on a situation model. It can generate questions about what, how and why. Several approaches to automatic QG with multiple selection of reading materials have been proposed. Coniam deleted every nth-word in the text to be a test item, and distractors were identified by choosing the same part of speech and similar word frequency to a tagged corpus. Mitkov and Ha extracted key terms that are noun phrases with a frequency that exceeds a certain threshold.

Limitations to the existing paper generator include:

- Inadequacy of storage space.
- May be harmed.
- Indigent environmental credentials.
- Incapable document transportation.
- Collaboration restricted.

- Problems related to modification.

III. PROPOSED SYSTEM

The main task is to create questions automatically based on the given text, i.e. sentences or groups of sentences. The basic input is a positive or declarative sentence or group of sentences. The output should be different kinds of questions depending on the type of sentences. Basically we can say that the input would simply be the text whereas output would be a number of questions. The basic goal of this system is to offer a solution to the problem of first breaking down the paragraph and then turning it into questions.

There are different types of problems that exist in the current manual system of question paper generation. These problems are:-

- Low security as paper is not secured using any mechanism.
- Patterns or repetitions may occur in paper.
- Slow as human labor is involved.
- Less variety of different types of questions.

Thus there are many problems associated with manual human based question paper generation, therefore to overcome this we came up with 'Automatic Intelligent Question Paper Generator'.

IV. PROPOSED SOLUTION

Automatic Question Paper Generator

In today's ambition, examination plays an important role in ensuring students' academic progress and the era of information technology is now being replaced by the productive application of technology. In the "information society" the production of utility from knowledge is important for the development of society. For various examinations conducted in any academic course in one year, teachers are required to make a variety of question papers as per the guidelines and evaluation requirements of the autonomous college. It is a tough task for the teachers to cover all the characteristics of the course results and avoid duplication of questions in successful exams. Many times, this entire element may degrade the standard of the question paper. Automated creation of exam paper provides a platform to create a streamlined exam paper and at the same time automation is easy to incorporate many elements that determine the quality of the question paper. An efficient question paper is generally suitable for ordinary students, but it also engages the demanding item for clever students. Thus automatically creating a question paper from a description entered by a teacher uses a word-labeled question bank that is currently required by the hour. Here we are implementing a system which accordingly creates the question paper from this semantically labelled question bank.

It is necessary to categorize queries as different class of queries need different methods for automatic generation of Question-type

- **Yes/No Questions:** also known as polar questions, whose expected answer is either yes or no.
- **Wh-Questions:** Wh-questions use interrogative words, such as why, when, who, where, what, which, etc.... to request information.

V. PROPOSED SYSTEM IMPLEMENTATION

The Automatic Question Paper Generator System is developed using the Python programming language.

In a fully functional system, there is a repository of syllabus, questions and pattern of question papers. It takes a simple text, a document or a pdf file as an input and provides a list of questions as an output. Process that is followed for generating questions automatically has the following steps -

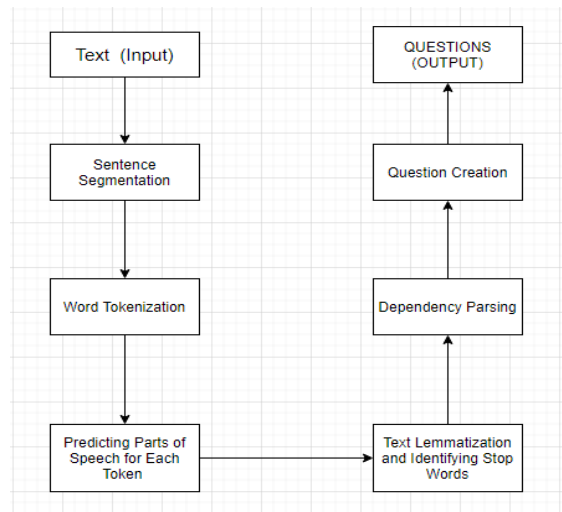


Fig 1. Working Model

Description of all the steps are as follows: Step 1: Sentence Segmentation: The very first step in the pipeline is to break the text in the document into separate sentences. Now that we've divided our document into sentences, we can process them one at a time.

Step 2: Word Tokenization: The second step in our pipeline is to break this sentence into tokens or separate words, which is called tokenization. In tokenization, we'll just split apart words whenever there's a space between them. And we'll also treat punctuation marks as separate tokens since punctuation also has meaning.

Step 3: Predicting Parts of Speech for Each Token: Next, we'll look at each token and try to guess its part of speech whether it is a noun, a verb, an adjective. Knowing the role of each word in the sentence will help us start to figure out what the sentence is talking about.

Step 4: Text Lemmatization & Identifying Stop Words: Lemmatization means figuring out the most basic form or lemma of each word in the sentence. Stop words are usually identified by just checking a list of known stopwords.

Step 5: Dependency Parsing: The next step is to figure out how all the words in a sentence relate to each other. This is called dependency parsing. So far, we've treated every word in our sentence as a separate entity. But sometimes it makes more sense to group together the words that represent a single idea or thing. We can use the information from the dependency parse tree to automatically group together words that are all talking about the same thing.

Step 6: Question Creation: At this point, we already have a useful representation of our sentence. We know the parts of speech for each word, how the words relate to each other and

Which words are talking about named entities? Now we'll create 'Wh' questions from this information.

VI. ARCHITECTURE

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 pdf file.

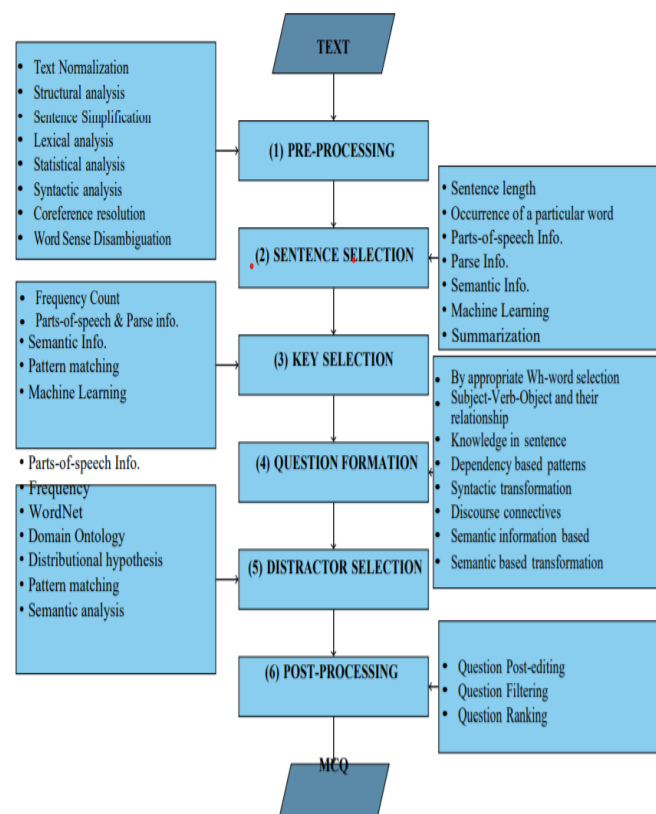


Fig 2.

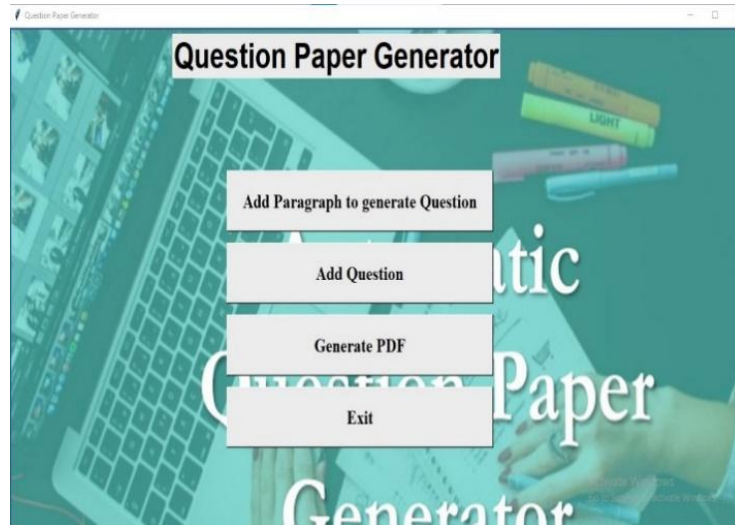
VII. RESULTS

There are many advantages of this approach that are described below:-

1. Compared to other algorithms, 'Artificial Intelligence Algorithm (using NLP)' is one of the best algorithms which we have used in automatic question paper generators and has few advantages over traditional methods. When compared to other question paper generating methods such as Backtracking Algorithm/Random algorithm, Artificial Intelligence Algorithm has higher processing speed, generation success and quality ratio.
2. We have used Python as the main language to implement the generator due to its flexibility, ease of use and extensive

collection of libraries for natural language processing

3. Its simple interface improves the ease of data updating and retrieving.
4. Generates and prepares the Question Paper in seconds. Questions can easily be edited.
5. Question Type can be Memory-based, Logic-based, Knowledge-based and Application-based.



VIII.CONCLUSION

The question selection difficulty has been formulated as a multi-constraint optimization issue, which aims to produce question papers that meet several constraints stated by the paper setter. Intelligent Question Paper Generator will generate a well formatted question paper in a matter of a few seconds i.e. saving a lot of time when compared to traditional systems. With the use of this system for question paper generation there are no chances of exam paper getting Leaked as paper is generated just before the test. It will save a lot of time for teachers and thus will improve efficiency. The implemented process strives to solve the problems listed above in a positive way. The implemented work narrates an automated system that heads away from the traditional process of paper generation to an automated process, by giving controlled entry to the resources that is attained by involving users and their roles in the colleges. We have also considered the importance of randomization in the process of paper generation. This system uses a logical algorithm which is absolutely randomized in nature and it also avoids duplication of questions. And therefore, the resultant automated system for Question Paper Generation will yield agrowth in phrase of random creation of question papers and also a secured platform.

IX.FUTURE SCOPE AND APPLICATION

This software can be widely used in educational institutes etc. It is a web based application so user location doesn't matter. Admin can access this software anywhere anytime. This system also evaluates the candidate's capability and skills efficiently. It is a fully automated system which results in fast results. In this system there is no need of transporting paper through police/security vans to all colleges. This system provides unbiased results. The usage of this system reduces human effort and saves time and resources to an extent.

This project can be further extended to generate yes/no type questions and fill in the blanks type questions. Further automatic answer copy checking systems can be added for checking the answer sheets and providing the result to the student.

X.ACKNOWLEDGMENT

It gives us great pleasure in presenting this project report titled "Automatic Question Paper Generator System" and we wish to express our immense gratitude to the people who provided invaluable knowledge and support in the completion of this project. Their guidance and motivation has helped in making this project a great success. We express our gratitude to our project guide **Prof. Aparna Hambarde** who provided us with all the guidance and encouragement throughout the project development. We would also like to express our sincere gratitude to the respective Project coordinators. We are eager and glad to express our gratitude to the **Head of the Computer Dept.** For his approval of this project. We are also thankful to him for providing us the needed assistance, detailed suggestions and also encouragement to do the project. We are extremely thankful to all staff and the management of the college for providing us all the facilities and resources required.

REFERENCES

- [1] Himanshu Jethwani, Mohd Shahid Husain, Mohd Akbar, ISSN 2347 - 7911, "International Journal for Innovations in Engineering, Science and Management" Available online at: www.ijiesm.com Volume 3, Issue 4, April 2015
- [2] Isha Sood, 7 Characteristics Of A Good Question. Online available at: <https://elearningindustry.com/characteristics-of-a-good-question-7>
- [3] Ming Liu, R. A. Calvo, A. Aditomo, L.A. Pizzato, "Using Wikipedia and Conceptual Graph Structures to Generate Questions for Academic Writing Support" Third Quarter 2012, pp. 251-263, vol. 5 DOI Bookmark: 10.1109/TLT.2012.5
- [4] T. Tokunaga, H. Nishikawa, and H. Obari, "Evaluation of automatically generated english vocabulary questions," Research and Practice in Technology Enhanced Learning, vol. 12, no. 1, p. 11, 2017.
- [5] R. Mitkov, L. An Ha, and N. Karamanis, "A computer-aided environment for generating multiple-choice test items," Nat. Lang. Eng., vol. 12, no. 2, pp. 177-194, Jun. 2006.
- [6] M. Liu, V. Rus, and L. Liu, "Automatic chinese factual question generation," IEEE Transactions on Learning



Technologies, vol. 10, no. 2, pp. 194–204, 2017.

[7] D. Metzler and W. B. Croft, “Analysis of statistical question classification for fact-based questions,” *Inf. Retr.*, vol. 8, no. 3, pp. 481–504, May 2005.

[8] M. Agarwal, “Cloze and open cloze question generation systems and their evaluation guidelines,” *International Institute of Information Technology, Hyderabad*, 2012.

[9] T. Goto, T. Kojiri, T. Watanabe, T. Iwata, and T. Yamada, “An automatic generation of multiple-choice cloze questions based on statistical learning,” in *Proceedings of the 17th International Conference on Computers in Education*. Asia-Pacific Society for Computers in Education, 2009, pp. 415–222.

[10] J. C. Brown, G. A. Frishkoff, and M. Eskenazi, “Automatic question generation for vocabulary assessment,” in *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, ser. HLT '05. Stroudsburg, PA, USA: Association for Computational Linguistics, 2005, pp. 819–826.

C.-Y. Chen, H.-C. Liou, and J. S. Chang, “Fast: An automatic generation system for grammar tests,” in *Proceedings of the COLING/ACL on Interactive Presentation Sessions*, ser. COLING-ACL '06. Stroudsburg, PA, USA: Association for Computational Linguistics, 2006, pp. 1–4.



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details