



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 6, June 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.165



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Question Generation Using Natural Language Processing

Dr. P L. Rajarajeswari¹, M. Devadharshini², P. Kaviya.³, C.S.Moshmi⁴

Associate Professor, Department of Computer Science and Engineering, Saranathan College of Engineering, Trichy,
Tamil Nadu, India¹

Students, Department of Computer Science and Engineering, Saranathan College of Engineering, Trichy,
Tamil Nadu, India²⁻⁵

ABSTRACT: Humans are naturally inquisitive about learning new things, and curiosity is what drives discoveries and learning. Teachers question the students and assess their performance in a manner similar to how they examine the students. Sometimes humans are not very good at asking appropriate questions due to their inconsistent brains in particular situations, which is why the notion of a system with the assistance of which would be able to produce the questions from a text automatically was proposed. The technology, also known as the Automated Question Paper Generator System, is quick, secure, and capable of producing questions at random. This suggested system may be given an input in the form of a document, a pdf file, or just plain text. The goal is to automatically generate accurate and pertinent questions using the provided keywords and textual information. This project uses an NLP-based system to automatically generate questions. The questions were framed using the keywords that were found to be essential terms in the instructional content using the NLP approach.

KEYWORDS: Question generation, Natural language processing, Deep learning.

I. INTRODUCTION

The goal of question creation from natural language text is to produce questions using text as the source, which may be useful for educational purposes. Question creation, which is the opposite of question-answering, has the ability to offer The long-term objective of the emerging Question Generation (QG) research community, which unites researchers from a wide range of disciplines including, but not limited to, Intelligent Tutoring Systems, Psycholinguistics, Discourse and Dialogue, Natural Language Generation, and Natural Language Understanding, is to automatically generate high-quality questions.

II. LITERATURE REVIEW

[1] Question Generation Using WordNet by Nguyen-Thanh LE & Niels Pinkwart, published by Humboldt-Universität zu Berlin, Germany. Asking Questions, according to studies, is a wonderful method to learn. Asking targeted, detailed questions can help identify knowledge gaps in novices who are often unable to clarify their questions. Other researchers used prompts as a form of question to encourage students to explain themselves, suggesting that prompts can be a useful instructional tool for enhancing conceptual knowledge. They employ questions to encourage their creativity in this study, with the goal of using the offered questions to construct new arguments for a certain debate topic. How can students be helped to develop more arguments on a given discussion topic by asking semantically related questions? In this study, a method is described for using WordNet to develop questions on a conversation topic and analyse the following research. This paper presents the findings of an evaluation study designed to examine the research topic in question.

[2] Automatic Question Generation Systems by Sheetal Rakangor, Dr. Y. R. Ghodasara published by International Journal of Scientific and Research Publications. This paper looks at how to create automatic objective questions using natural language processing (NLP) or statistical patterns, as well as a review of past work in the field. MCQ (Multiple Choice Question) style questions are typically used when there are a high number of students taking tests like GATE,

CAT, and NET. MCQs are straightforward to assess, and because they are assessed using automated programmes, results can be announced in a matter of hours, with the assessment process being totally transparent. The researcher looked into the test technique to see if questions might be generated automatically with the help of a computer programme. As a result, the workload of the educator will be minimized. NLP, is a topic of study in which many researchers have published their findings and where accuracy is still being investigated.

[3] Automatic question generation and answer assessment by Bidyut Das, Mukta Majumder, Santanu Phadikar & Arif Ahmed Sheik, published by Research and practice in enhanced technology learning. Learners can use online learning to learn on the internet using a computer or other digital device. The three major types of online learning based on the learning materials are textual learning, visual learning, and audio-video learning. Online learning necessitates two things: learning resources and assessment of learners who use such resources. Students can learn from a number of internet sources using the learning tools provided. On the other hand, manual questions from the learning materials are required for the learner's assessment. To the knowledge, no generic assessment approach has been provided in the literature to evaluate learners' learning gaps from e-reading documents. As a result, automated question production and evaluation procedures can aid in the assessment system's automation. The first contribution is to compile a list of relevant datasets. Various purposes are studied to vary aims and constraints of using these datasets critically. The second contribution is to examine and summarize present and potential question generation technologies, as well as the evaluation procedures that go with them.

[4] Automatic generation of multiple choice questions using dependency based semantic relations by Naveed Afzal and Ruslan Mitkov, published by Soft Computing technologies. This work presents an unsupervised dependency-based technique for extracting semantic linkages that can be applied to automatic multiple choice question generation (MCQs). Multiple choice tests, or MCQs, are a popular choice for large-scale assessments because they make it easy for test takers to finish exams and for examiners to understand the results. Creating MCQs by hand is a costly and time-consuming procedure, yet it is commonly necessary on a large scale and in short iterative cycles. The problem of automated MCQ production is addressed via unsupervised relation extraction, a technique used in a number of related natural language processing challenges. The goal of unsupervised relation extraction is to find the most important named entities and phrases in a document and then recognise semantic relationships between them.

[5] A Systematic Review of Automatic Question Generation for Educational Purposes by Ghader Kurdi, Jared Leo, Bijan Parsia, Uli Sattler, Salam Al-Emari, published by International Journal of AI. Several published evaluations of authentic test questions (mostly multiple choice questions (MCQs)) reveal poor quality, which Tarrant et al. (2006) attribute to a lack of assessment development skills. The need to replace assessment questions on a regular basis to ensure their validity, as their value diminishes or vanishes after a few rounds of use (due to being shared among test takers), as well as the rise of e-learning technologies such as massive open online courses (MOOCs) and adaptive learning, which necessitate a larger pool of questions, exacerbate the problem. The difficulty that test developers have in developing a large number of high-quality questions have given rise to AQG approaches. AQG is working on developing algorithms for generating questions from structured and unstructured information sources. The AQG has been studied since the 1970s. AQG is growing more significant as MOOCs and other e-learning technologies become more popular.

III. PROPOSED SYSTEM

In the currently proposed system of question generation, various types of questions were generated.

- 1.) MCQ
- 2.) TRUE/FALSE
- 3.) FILL IN THE BLANKS

MCQ

The Article/Documents are given as an input to the system for generating questions. The given input is summarized using the text summarizer and the sentence is mapped by finding the main context of the input. Then the keywords are extracted for deciding which sentences can be mapped. With the keywords derived, distractors are generated using wordnet and conceptnet. The generated distractors are ranked using BERT in such a way they are in connection to the



given input phrase and the keywords. Finally with the keywords and ranked distractors the system generates a number of MCQ questions.

TRUE/FALSE

The article/documents are given as the input for generation of true or false questions. The input is summarized, for true sentence generation, the compound and complex sentences are converted into simple sentences. For generation of false sentence, input text needs to be falsified. For falsifying the input there are many ways,

- i.) Add or remove negation
- ii.) Change adjective
- iii.) Change a named entity
- iv.) Change main verb
- v.) Split compound or complex sentences to simple sentences
- vi.) Changing noun phrase or verb phrases

For modifying the noun phrase or verb phrase, consistency parsing is used; this parser analyzes the sentences by breaking them down into sub-phrases. The sub-phrases belong to verb phrases or noun phrases. This parser identifies and removes the verb phrase or noun phrase in the sentence. Then open AI GPT-2 is used to predict the probability of the sentence. This open AI GPT-2 generates multiple sentences with different options of verb phrases or noun phrases. These various sentences are ranked using BERT. The most dissimilar sentence is taken as the false sentence for generation of false question.

FILL IN THE BLANKS

The article/documents are given as the input for the fillup generator, using textblob the python library. It figures sentences in a given text, and for each sentence it would first find the proper noun and delete them and replace it with the blank. If proper noun is not present in the sentence it would search for the normal noun, delete them and replace it with the blank and the questions will be generated. If no nouns are present in the sentences, it would return none.

IV. ACCURACY OF THE PROJECT

Accuracy

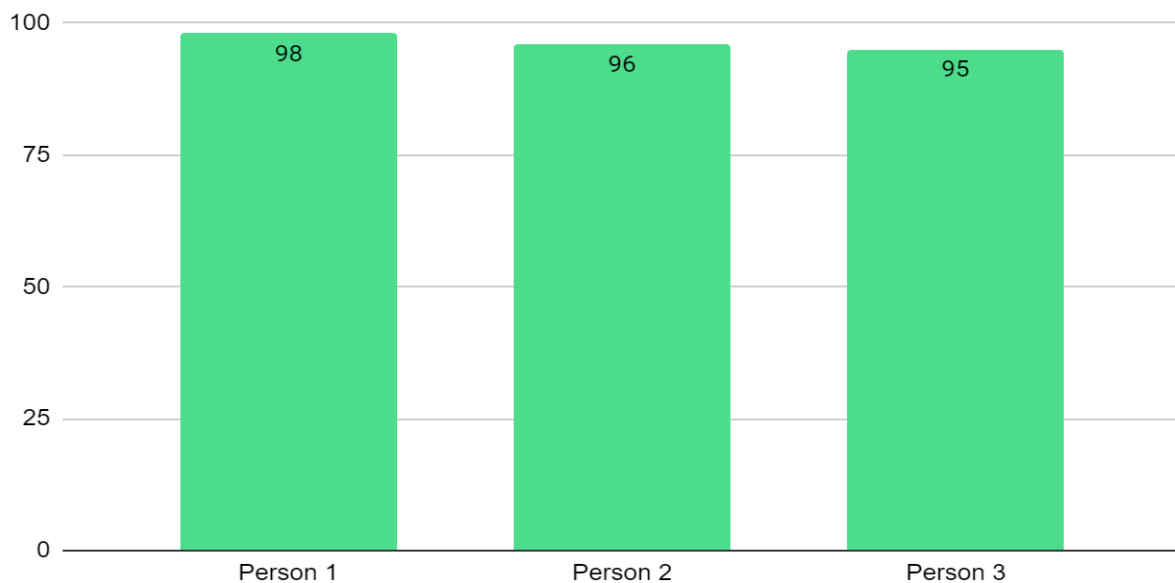


Fig 1: Bar graph representation of the accuracy of the project.

V. CONCLUSION

A method for generating questions utilizing NLP as a source of semantic information is described in this project. The idea is to use produced questions to encourage students to brainstorm and participate more actively in debates. A pilot study compared system-generated questions to questions that were manually developed by scholars in the argumentation and question generation research areas. The findings suggest that there isn't much of a difference between human-generated and system-generated questions: human raters couldn't discern the difference. It should be emphasized, however, that it had to manually select a small number of questions from a large number (hundreds) of machine generated questions.

REFERENCES

- [1] Neutral Question generation using NLP Qingyu Zhou, Nan Yang, Furu Wei, Chuanqi Tan, Hangbo Bao & Ming Zhou. https://link.springer.com/chapter/10.1007/978-3-319-73618-1_56
- [2] Question generation for question answering Nan Duan, Peng Chen <https://aclanthology.org/D17-1090/>
- [3] A Systematic Review of Automatic Question Generation for Educational Purposes Ghader Kurdi, Jared Leo, Bijan Parsia, Uli Sattler & Salam AlEmari <https://link.springer.com/article/10.1007/s40593-019-00186-y>
- [4] Question Generation Using WordNet. Nguyen-Thinh Le, Niels Pinkwart https://www.researchgate.net/publication/273630846_Question_Generation_Using_WordNet
- [5] Automatic question generation and answer assessment: a survey. Bidyut Das, Mukta Majumder, Santanu Phadikar & Arif Ahmed Sekh <https://telrp.springeropen.com/articles/10.11879>
- [6] Nasrin Mostafazadeh, Ishan Misra, Jacob Devlin, Margaret Mitchell, Xiaodong He, and Lucy Vanderwende. 2016. Generating natural questions about an image. <http://www.aclweb.org/anthology/P16-1170>
- [7] Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In Stan Szpakowicz Marie-Francine Moens, editor, Text Summarization Branches Out: Proceedings of the ACL-04 Workshop. Association for Computational Linguistics, Barcelona, Spain, pages 74–81. <http://aclweb.org/anthology/W/W04/W04-1013.pdf>
- [8] Matthew Richardson, Christopher J.C. Burges, and Erin Renshaw. 2013. MCTest: A challenge dataset for the open-domain machine comprehension of text. <http://www.aclweb.org/anthology/D13-1020>
- [9] David Lindberg, Fred Popowich, John Nesbit, and Phil Winne. 2013. Generating natural language questions to support learning on-line. <http://www.aclweb.org/anthology/W13-2114> 80 Kenneth Mark Colby, Sylvia Weber, and Franklin Dennis Hilf. 1971. [https://doi.org/10.1016/0004-3702\(71\)90002-6](https://doi.org/10.1016/0004-3702(71)90002-6)
- [10] Michael Denkowski and Alon Lavie. 2014. Meteor universal: Language specific translation evaluation for any target language <http://www.aclweb.org/anthology/W14-3348>
- [11] Srinivasan Iyer, Ioannis Konstas, Alvin Cheung, and Luke Zettlemoyer. 2016. Summarizing source code using a neural attention model. <http://www.aclweb.org/anthology/P16-1195>
- [12] Ruslan Mitkov and Le An Ha. 2003. Computer Aided generation of multiple-choice tests. In Jill Burstein and Claudia Leacock, editors, Proceedings of the HLT-NAACL 03 Workshop on Building Educational Applications Using Natural Language Processing. <http://www.aclweb.org/anthology/W03-0203.pdf>
- [13] Alexander M. Rush, Sumit Chopra, and Jason Weston. 2015. A neural attention model for abstractive sentence summarization. <http://aclweb.org/anthology/D15-1044>
- [14] Joseph Weizenbaum. 1966. Eliza—a computer program for the study of natural language communication between man and machine. <https://doi.org/10.1145/365153>



INNO  SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

 **doi**[®]
CROSS **ref**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details