RSP Science Hub

International Conference on intelligent COMPUting TEchnologies and Research (i-COMPUTER) 2023

Histobot: Question Generation System Using Deep Learning Techniques

Docca Pranav [1], Badri Prasad V R [1]
[1]Department of Computer Science and Engineering, PES University Bangalore, India.

Emails: doccapranav@gmail.com, badriprasad@pes.edu

Abstract

From the analysis of recent researches of automatic question generation using deep learning techniques, we examined papers between 2022 and early 2023 from the examination of recent research on automatic question production using deep learning techniques. Our study comes after the survey report that broadens the analysis of earlier evaluations of AQG content that surfaced between late 2014 and early 2019. We examined the researched works that were included, looking at things like the (1) framework for question generation and (2) generating method. We discovered that contemporary methods frequently rely on generative frameworks that deploy Transformer-based models and GPT-n series, which are more efficient in terms of analysis and performances. We discovered that question creation has gained popularity recently and has significantly improved the educational field. Yet, it can be challenging to produce automatic questions and create the necessary question patterns, structures, and forms. Our additional research advises testing out more practical, efficient models and strategies for autonomous question generating.

1. Introduction

A vital role in educational fields is played by frameworks of questions. The question format has the capacity to affect learning in addition to its framing function. The benefits of well-crafted questions include:

- The opportunity to practice retrieving information from memory;

- Getting feedback from students on their misunderstandings.

- Increasing student focus on the course/central subject's idea

- Evaluating students' performance in light of this idea.

Though having Notwithstanding these benefits, producing handcrafted questions is a challenging process that requires resources, information, and skill. The demand for questions has expanded with the quick development of online learning. Multiple-choice questions (MCQs), which make up the majority of real-time exam questions, have a poor quality, as shown by the previous analysis of real-time test questions. which outlined using the steps to conduct a systemic review of those systems. Techniques for automatic question generating (AQG) have become a viable option for increasing productivity of questions at a lower cost.

About the framework of the algorithms' development of the structured queries from the information sources. Due to its widespread employment in quizzes (Dhawaleswar Rao CH et al., 2020; Bowei Zuo et al., 2022) (Rao, H, Saha, et al. Zou et al.),

learning activities, and functioning as a data augmentation method for training Question Generation (QG) systems, the focus of this survey is on AQG from comprehension paragraphs. Our project is a survey paper that involves reading and evaluating prior research on AQG from the comprehension paragraph. (Vie et al., 2017) (Vie et al.) strives to adjust contemporary teaching methods to the knowledge of the pupils. Lastly, we have been able to automate the control of question characteristics such as difficulty of questions and level of logical reasoning, which can guide the design of a good quality question exam.

2. Literature survey

Researchers have used a variety of teaching strategies and automatic question creation techniques to produce questions automatically from provided comprehension passages or texts. The survey begins by gaining an overview of the prior research work through the use of numerous regularly utilized approaches and automatic question production schemes.

"Automatic Multiple Choice Question Creation from Text: A Survey," by Dhawaleswar Rao CH et al., (Rao, H, Saha, et al.) In this study, the author suggests developing MCQ questions automatically while doing a thorough investigation into how to evaluate questions culled from text for use in the educational sector. In this research, the authors reviewed the literature to determine what approaches or strategies might be used to generate multiple-choice questions (MCQs) and encountered some difficulties that needed to be overcome. With the current process approach being distinguished from the existing process approach. The multiple-line factor MCQ test candidates' ability to handle the idea using various accepted evaluation techniques.

"Automatic True-False Question Creation for Educational Use," by Bowei Zuo et al., (Zou et al.) In this study, the author suggests developing true-false questions automatically while conducting in-depth research to make them scalable on an educational scale. They have examined how to develop inquiries from

• **Template-based framework:** language learning and understanding the process with fixed text from the input.

• **Generative-based framework:** create T/F questions that are more flexible and complicated.

In this research, the authors studied the literature to determine the various methods for creating questions, and then concentrated on the pedagogical approach to question selection with novel masking features to evaluate the performance and flexibility. while coming up with the difficult questions.

"Automatic Fill-the-blank Question Generator for Student Self-assessment," by Girish Kumar et al., (Kumar, Banchs, D'haro, et al.) In order to assist students in identifying specific themes on which they are having trouble focusing, the author of this study has suggested a teacher-centric approach to the system. With this strategy, professors are able to construct and modify pop-up exam questions using lecture notes. This concept enables students to learn alongside teachers and gain from our system's close supervision.

"Chatbot: An Automated Conversation System for The Educational Area," by Anupam Modal et al., (Modal et al.) The author of this study has suggested a retrieval-based approach employing and They included semantic and contextual elements along with the bags-of-words (BOW) paradigm, which is essential for chatbot construction. They merged Term Frequency (TF), Inverse Document Frequency (IDF), and TF-IDF algorithms with NLP algorithms in this.

The following are some potential chatbot restrictions that are discussed:

• **Fixed Rule-based:** Current chatbots are created using simple machine learning algorithms, a set of preset rules, and template-based matching.

• **Grammatical Mistakes:** It is quite challenging to train a model that understands the query past the grammatical mistakes that the query contains.

• **Closed-Domain:** The majority of chatbots are built to respond to queries from a closed domain or to only those queries that are specifically defined in the database.

"Conversational Question Answering Over Knowledge Base Using Chat-Bot Framework" by Sai Sharath et al., (Sharath, Banafsheh, Rekabdar, et al.) The author of this study has suggested employing RADA (Robust-Answer Driven-Assistant) on KB as an attention-based approach (Knowledge bases). To implement the question-answer structure, RADA upgraded the QA (Question Answering)

model into a chatbot framework.

"A review of recent advances in adaptive assessment" by Vie et al., (Vie et al.) In this study, the author has suggested a technique whereby computerized adaptive assessments (CATs) choose the subsequent question based on the student's previous responses. When using

- **Q-matrices:** a linear complementary problem typically related to simulated data.

- **Cognitive diagnosis models :**are restricted laten class models that take student performance on a knowledge-based learning test into account.

- **The partial ordering** of discrete information within subsets is the foundation of the abstract theory known as space theory.

"Automatic Multiple Choice Question Generation from Thai Text" by Atiwong Suchato et al., In this study, the author has suggested a method that divides input text into clauses based on the part of speech (POS) and utilizes sentence-breaking spaces to do so. A question that accepts phrases as nouns in the query was created from the Thai text using distractors and Thai dataset. The proposed solution includes the use of ranking algorithms with bespoke dictionaries and linear regression models.

"CSFQGD: Chinese Sentence Fill-in-the-Blank Question Generation Dataset for Examination," by Tianlin Zhang et al., (Zhang et al.) The approach of creating fill-in-the-blanks from reading comprehension using a dataset of Chinese sentences has been suggested by the author in this study. The dataset includes a variety of real test subjects, and it generates fill-in-the-blank questions using a variety of deep learning-based techniques, such as neural network-based models.

"Automatic Question Generation System," by Prof. D.R. Nanda Devi et al., (Prof, Devi, Akshara, et al.) In this study, the author has suggested a technique for developing fact-based queries about a piece of content using its connected data. In order to reduce human labour and time spent going over the question bank with related material, they have utilized linguistics rules to organize the sentences and forecast the argument in this case. They have also used semantic, POS tagging, and named entities to identify relevant passage.

"Automatic Question Generation System," by Dr. P. Prabitha et al., (Dr et al.) Using the Supervised Learning Approach, Naive Bayes method, noun filtering, summarization, selection of the tag content (what to ask), question type (what, why, how, and when), Porter Stemmer, and Phrase Mapper, the author has proposed a method in this study for generating numerous questions from brief documents.

"AI and Machine Learning Based E-Learning System for Secondary Education," by Wijayawardena G et al., (Wijayawardena et al.) In this study, the author suggests that a chatbot's technique for forecasting students' final grades and weak areas in the context of COVID-19 utilize the NL Toolkit (NLTK) and machine learning models. They urge students to study by doing this because Sri Lanka's secondary schools employ an E-learning programmed that has a number of difficulties and issues.

"Enhancing Paraphrase Question Generation with Prior Knowledge" by Jia Yuan Xie et al., (Xie et al.) The method of knowledge-aware paraphrase question has been proposed by the author in this study to rewrite the paraphrase question into a new paraphrase question. However, this method should have used an attention mechanism and a gate mechanism in addition to the explicit meaning of the original question to extract prior knowledge for the rewriting of the question. They created the questions using the SQuAD dataset and the discriminator module, which appropriately imported the questions.

"Generating Questions and Multiple-Choice Answers Using Semantic Analysis of Texts," by Jnu Araki et al., (Araki et al.) The author of this study has suggested a fresh technique to automated question production that enhances the work. Here, patterns taken from entities and events were applied as the second way, and semantic parsers were used as the third method to analyse a number of sentences.

"Chatbot For E-Learning Using Machine Learning," by Varsha R. Palandurkar et al., (Varsha et al.) The author of this study has suggested deploying this chatbot's college inquiry option, which includes all the information regarding courses, admissions, professors, and college infrastructures. They interact with college websites to obtain crucial information about their institution.

"Automatic Question Generation Model Based on Deep Learning Approach," by Hala Abdel-Galil et al., (Abdel-Galil, Mokhtar, Doma, et al.) The researcher suggested using a GUI-based approach to

produce Wh-questions, or what they stand for: what, who, where, and when, and structure them into two alternative bases.

1) Question Bank template
2) Exam template

Following the implementation of this procedure, they found it to be quite beneficial in the field of education, and they applied the attention-based decoder, feature-rich encoder, and seq2seq technique.

"Towards Generalized Methods for Automatic Question Generation in Educational Domains", by Huy A. Nguyen et al., (Huy et al.) In order to generate questions for a hierarchical idea extraction model on the material utilizing cutting-edge technology and approaches, the author of this study has presented a way employing the T5 transformer model and the GPT-3 trained model in a data science course.

3. Conclusion

After reading multiple study papers, one thing is crystal clear: artificial intelligence is essential to the future of any industry. There are many advantages to deploying a chatbot in the educational sector, including the ability to reduce human labour in monotonous and repetitive jobs, offer round-the-clock customer service, quickly derive statistical inferences, and make predictions based on input data. All of the publications cited here reach the same result and emphasize the benefits of adopting chatbots across various industries. All of these research papers allow us to see the advantages, implementation strategies, and model architecture.
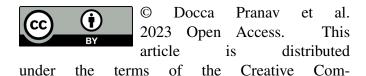
References

Abdel-Galil, Hala, Mai Mokhtar, Salma Doma, et al. "Automatic Question Generation Model Based On Deep Learning Approach". *International Journal of Intelligent Computing and Information Sciences (IJICIS)* (2021).

Araki, Dheeraj Jnu, et al. "Generating Questions and Multiple-Choice Answers using Semantic Analysis of Texts". *the 26th International Conference on Computational Linguistics: Technical Papers, (ICCL)* (2016): 1125–1136.

Dr, et al. "Automatic Question Generation System". *2014 International Conference on Recent Trends in Information Technology (ICRTIT)* (2014). 978-1-4799-4989-%202/14/%5C$31.00.

Huy, A, et al. "Towards Generalized Methods for Automatic Question Generation in Educational Domains". *2022 In book: Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption* (2022): 272–284.

Kumar, Girish, Rafael E Banchs, Luis Fernando D'haro, et al. "Automatic fill-the-blank question generator for student self-assessment". *2015 IEEE Frontiers in Education Conference (FIE)* (2015). 978-1-4799-8454-%201/15/%5C$31.00.

Modal, Anupam, et al. "Chatbot: An automated Conversation System for The Educational Domain". *International Jiont Symposium on Artificial Intelligence and Natural Language Processing* (2019). 978-1-7281-0614-4/18/%5C$31.00.

Prof, D, Yaramada Nanda Devi, Akshara, et al. "Sindhuja Nadgoud, Veda Charitha Bandikallu". *International Research Journal of Engineering and Technology (IRJET* (2021): 11–11.

Rao, Dhawaleswar, C H, Sujan Kumar Saha, et al. "Automatic Multiple Choice Question Generation from Text: A Survey". *IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES* 13.1 (2020).

Sharath, Sai, Japa Banafsheh, Rekabdar, et al. "Conversational Question Answering Over Knowledge Base using Chat-Bot Framework". *2021 IEEE 15th International Conference on Semantic Computing (ICSC)* (2021).

Varsha, R, et al. "Chatbot For E-Learning Using Machine Learning". *International Research Journal of Modernization in Engineering Technology and Science* (2022).

Vie, Jill-Jênn J, et al. "A Review of Recent Advances in Adaptive Assessment". *Learning Analytics: Fundaments, Applications, and Trends* (2017): 113–142.

Wijayawardena, G C, et al. "AI and Machine Learning Based E - Learning System For Secondary Education". *2022 IEEE 7th International conference for Convergence in Technology (I2CT)* (2022).

Xie, Wenhao Jia Yuan, et al. "Enhancing Paraphrase Question Generation with Prior Knowledge", (2022) Transactions on Audio". *Speech and Language Processing* (2022).

Zhang, Tianlin, et al. "CSFQGD: Chinese Sentence Fill-in-the-blank Question Generation Dataset for Examination". *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)* (2021). 978-1-7281-%206597-4/21/%5C$31.00.

Zou, Bowei, et al. "Automatic True/False Question Generation for Educational Purpose". *Proceedings of the 17th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2022)* (2022): 61–70.

Embargo period: The article has no embargo period.