A Survey on Development Approaches for Automated MCQ Generator Using Natural Language Processing

ADITYA SANGWAI[1], ADITYA AGRAWAL[2], PRASAD PATIL[3], MANDAR KULKARNI[4], PROF. DR. ANUPAMA PHAKATKAR[5]

[1, 2, 3, 4] *Computer Engineering Department, Pune Institute of Computer Engineering, Pune, India.*

[5] *Assistant Professor, Pune Institute of Computer Engineering, Pune, India*

Abstract- Within the field of education, it is widely acknowledged that posing questions to learners at the end of a lesson is an effective teaching strategy. For cost-saving reasons, especially when there are multiple candidates, the majority of educational institutions and have made multiple-choice questioning (MCQ) the mainstay of their testing procedures. In Natural Language Processing (NLP), the task of autonomously generating multiple-choice questions is both beneficial and challenging. It involves using textual information to automatically generate relevant and accurate queries. Teachers find it stressful and challenging to manually generate meaningful, significant, and relevant questions, despite its importance. In our project, we describe an NLP-based method for producing MCQs on its own. Natural language processing (NLP) is an artificial intelligence field that studies how humans and computers interact with natural language. Our approach places a strong emphasis on using natural language processing (NLP) to set up multiple choice questions (MCQs). This improves the process of creating and modifying MCQs and creates a useful question bank that academics can use later on with their students. This will ensure that the multiple-choice questions (MCQ) contain options and questions relevant to the learning objectives.

Indexed Terms- Multiple Choice Questions, Natural Language Processing, Distractor Generation, Summary Generation, Automated Question Generation

I. INTRODUCTION

Very Common type of Assessment for testing someone's knowledge is Multiple Choice Questions (MCQs) test. Respondents are asked to choose the best option from a list of options in multiple choice questions (MCQs), a fairly common type of evaluation. The three components of an MCQ are the distractors, the target word, and the stem. The sentence that forms the question's stem (also called an item) is the target word (also called a key). The creation of an automated multiple-choice question generator has gained popularity as a research topic in recent years. Three main procedures have been followed by automatic multiple-choice question (MCQ) Identify applicable funding agency here. If none, delete this. systems in the literature that we have observed: selecting sentences (or stems), selecting target words, and creating distractions. Very Common type of Assessment for testing someone's knowledge is Multiple Choice Questions (MCQs) test. Respondents are asked to choose the best option from a list of options in multiple choice questions (MCQs), a fairly common type of evaluation. The three components of an MCQ are the distractors, the target word, and the stem. The sentence that forms the question's stem (also called an item) is the target word (also called a key). The creation of an automated multiple-choice question generator has gained popularity as a research topic in recent years. Three main procedures have been followed by automatic multiple-choice question (MCQ) Identify applicable funding agencies here. If none, delete this. systems in the literature that we have observed: selecting sentences (or stems), selecting target words, and creating distractions. The generation of automated multiple-choice questions has become a focal point of research in recent years. This technological advancement seeks to streamline the process of creating MCQs, offering efficiency and consistency in assessment design. In the literature, three distinct procedures have been identified in the development of automatic MCQ systems: selecting sentences or stems, identifying target words, and generating appropriate distractors. The selection of sentences involves choosing the text that will form the basis of the question, often referred to as the stem. The stem sets the context and presents the scenario for the MCQ. Concurrently, the identification of target words is crucial, as these represent the correct answers within the options provided. Additionally, the creation of distractors is a key aspect of automated MCQ generation, involving the formulation of plausible but incorrect alternatives to challenge the knowledge of respondents. This surge in interest and research in automated MCQ generation reflects a growing need for scalable and efficient assessment tools. Such systems not only save time and resources but also have the potential to enhance the overall quality and diversity of questions presented in assessments. As technology continues to evolve, the development of sophisticated automated MCQ generators is likely to play a pivotal role in shaping the landscape of knowledge evaluation methodologies.

II. LITERATURE SURVEY

In [1], authors delve into the realm of advanced technologies for automatic question paper generation. They employ language models such as BERT, Blooming Taxonomy, and Summarizer algorithms, alongside NLP techniques like Dependency Parsing, POS tagging, and NER. The application developed as part of their research has demonstrated efficacy in producing relevant questions, receiving positive user feedback. The study underscores the transformative potential of natural language processing in crafting tools that enhance organizational operations, particularly in the educational domain. By seamlessly integrating sophisticated algorithms with practical applications, the research contributes to the ongoing discourse on leveraging technology to elevate modern classroom assessment methodologies.

In [2], the researchers explore the application of the BERTSUM model, RAKE algorithm, and WordNet approach. Notably, their research reveals that the BERTSUM model exhibits a commendable potential to generate multiple-choice questions (MCQs) characterized by a higher level of confusion among the options compared to alternative models. However, the findings also identify an opportunity for improvement, particularly in terms of enhancing the meaningfulness of question-answer pairs generated by the system. This nuanced evaluation sheds light on the strengths and areas for refinement in the proposed automatic question and distractor generation methods, contributing valuable insights to the field of question generation research.

In [3], authors introduce a novel approach to question generation. Their model employs a Reinforcement Learning Framework, Copying Mechanism, and Knowledge Graphs to enhance the question generation process. Notably, the model incorporates copy and coverage mechanisms, along with a pointer network, to effectively generate questions from textual content. The evaluation of their model is comprehensive, utilizing automatic metrics such as BLEU, ROUGE-L, and METEOR. Additionally, human evaluation is conducted to assess the syntactic correctness, semantic correctness, and relevance of the generated questions. This Generator-Evaluator framework contributes a sophisticated and thoroughly evaluated methodology to the evolving landscape of question generation from text.

In [4], authors present a cutting-edge approach that integrates state-of-the-art techniques. The method harnesses the power of T5 for text generation, coupled with meticulously designed rule-based algorithms and enriched by human expert evaluation. This holistic system not only automates the generation of structured multiple-choice questions but also empowers educators with precise control over grammar topics and question content. The inclusion of human expert evaluation ensures a refined and high-quality output, addressing nuances that algorithms alone may overlook. Beyond its immediate applications, the research underscores the broader transformative potential of artificial intelligence in reshaping English grammar education, signaling a paradigm shift in pedagogical practices and assessment methodologies. As technology continues to play an increasingly pivotal role in education, this innovative approach stands as a beacon, showcasing the symbiosis of advanced algorithms and human expertise in advancing language learning and assessment.

In [5], authors explored the dimensions of Question Generation (QG), Question Answering (QA), and Discourse Generation (DG) within the context of

MCQ-based questionnaire generation. The processing pipeline is implemented using the T5 (Text-to-Text Transfer Transformer) language model, showcasing the model's versatility in handling different aspects of question generation. In particular, the Discourse Generation task adopts a Text-to-Text format, resulting in a notable 14.91 ROUGE-L improvement on the DG-RACE dataset. The discussion introduces cosine similarity (0.71 average) as a complementary metric for Discourse Generation. Furthermore, the research recommends the exploration of improved metrics to measure Question Generation (QG) performance, aiming for an enhanced state-of-the-art reference in the field. The study contributes to advancing the capabilities of Text-to-Text Transfer Transformer models in the comprehensive generation of Multiple-Choice Questions.

In [6], authors presented a novel methodology that seamlessly combines traditional linguistic methods with advanced machine learning techniques for automated question generation. The model employs Support Vector Machines (SVM), Naive Bayes, Recurrent Neural Networks (RNN), and Transformers in a data-driven learning process, minimizing reliance on human experts and showcasing self-sufficiency. Notably, the system exhibits continuous improvement through reinforcement learning, expanding transformation rules and adapting to user feedback. The introduction of a hierarchical pattern system enhances versatility, enabling the generation of diverse question types based on similarity matches at various linguistic levels. Rigorous experimentation and comparisons highlight the framework's superior performance, consistently outperforming existing systems in correctness and garnering favorable user evaluations. This research underscores the efficacy of their machine learning-driven approach in advancing the state-of-the-art in automatic question generation.

In [7], authors presented an innovative approach to Automatic Question Generation (AQG) for IELTS reading comprehension. Integrating Natural Language Processing (NLP) and the K-Nearest Neighborhood (KNN) algorithm, the authors employ a systematic process involving news article scraping, tokenization, tree structure conversion, and extraction. Refinement steps include pronoun elimination, POS tag conversion, and numeric conversions for linguistic diversity. Quantitative data and distance measurement introduce a machine learning element, enhancing question relevance. The methodology emphasizes linguistic enhancement techniques like synonym substitution and syntax improvement for challenging IELTS questions. Overall, their model showcases the synergistic power of NLP and KNN in advancing automated question generation in reading comprehension.

In [8], authors presented a comprehensive approach leveraging three key linguistic tools: a Syntactic Parser, Named Entity Recognizer (NER), and Part of Speech Tagger. The authors propose an innovative method for automated question generation, employing the TREC 2007 dataset and evaluating their approach with a focus on Recall and Precision metrics. Their strategy involves a target-driven approach for data filtering and streamlined sentence processing through syntactic information. They classify sentences based on their constituent elements, such as subject, verb, object, and preposition, laying the foundation for question generation. The authors have successfully implemented automated question generation using predefined interaction rules, with plans to expand the rule set for increased versatility. Notably, their work highlights the need for future enhancements, emphasizing improvements in sentence classification robustness and the integration of semantic information and word sense disambiguation to elevate the overall effectiveness of their question generation approach.

III. TAXONOMY

1) Factual Questions: Factual questions are designed to evaluate the respondent's knowledge of specific information or facts. These questions typically have a clear and correct answer, testing the individual's ability to recall and understand straightforward information.
2) Yes/No Questions: Yes/No questions require respondents to choose between two binary options, either affirming or negating a given statement. These questions are designed to assess a basic understanding of concepts and are particularly useful for gauging foundational knowledge.
3) True/False Questions: True/False questions present respondents with statements that they must categorize as either true or false. This format is

commonly used to assess the accuracy of knowledge and the ability to discern the correctness of statements.
4) Fill in the Blank with Options: Fill in the Blank with Options questions involve completing a sentence or phrase by selecting the correct option from a provided list. These questions assess not only knowledge of content but also the ability to apply that knowledge in context.
5) Exclusionary Questions: Exclusionary questions ask respondents to identify the option that does not belong to a specific category or group. These questions require a deeper understanding of classification and the ability to differentiate between items based on specific criteria.
6) Comprehension Questions: Comprehension questions are often part of a larger passage or text, requiring respondents to demonstrate their understanding of the material. These questions assess the ability to interpret and infer information from written content.
7) Cause and Effect Questions: Cause and Effect questions assess the respondent's understanding of relationships between events or actions. These questions require individuals to identify the causes and consequences of specific occurrences, demonstrating a deeper comprehension of causal connections.
8) Language Proficiency Questions: Language proficiency questions evaluate an individual's command of language, including grammar, vocabulary, and syntax. These questions aim to assess the respondent's ability to use language accurately and effectively, contributing to a broader evaluation of language skills.

CONCLUSION

In conclusion, our research highlights the significance of leveraging Natural Language Processing (NLP) for the autonomous generation of Multiple-Choice Questions (MCQs) in the field of education. Recognizing the effectiveness of posing questions at the end of a lesson, we address the common challenge faced by educational institutions in manually generating meaningful and relevant MCQs. By applying NLP techniques, our approach not only streamlines the process of question generation but also ensures the creation of questions and options that align closely with the learning objectives. This automated method not only saves time and effort for teachers but also contributes to the establishment of a comprehensive and useful question bank. As institutions increasingly adopt cost-effective testing procedures, our NLP-based approach emerges as a beneficial and practical solution for enhancing the quality of assessments and promoting effective learning outcomes. The ability to dynamically generate questions based on the latest educational content ensures that assessments remain current and reflective of the most up-to-date knowledge. Furthermore, the systematic integration of NLP techniques fosters a more personalized learning experience for students, as the generated questions can be tailored to individual learning styles and proficiency levels. This adaptability not only enhances engagement but also contributes to a more inclusive and student-centric educational environment. As we move towards an era of digital transformation in education, harnessing the power of NLP for MCQ generation stands out as a forward-looking strategy, poised to revolutionize the assessment landscape and elevate the overall educational experience.

REFERENCES

[1] Ruslan Mitkov, Leanha and Nikioskar Amanis. "A computer-aided environment for generating multiple-choice test items.", Research Group in Computational Linguistics, School of Humanities, Languages and Social Sciences, University of Wolverhampton, Wolverhampton WV1 1SB, UK, 2006.
[2] Itziar Aldabe and Montse Maritxalar. "Semantic Similarity Measures for the Generation of Science Tests in Basque.", 2014.
[3] Maha Al-Yahya. "Ontology-Based Multiple Choice Question Generation.", Information Technology Department, 2014.
[4] Manish Agrawal, Prashanth Mannem. "Automatic gap-fill question generation from textbooks.", 2011.
[5] Susanti, Y., Tokunaga, T., Nishikawa, H., and Obari, H., "Automatic distractor generation for multiple Choice English vocabulary questions.", Research and Practice in Technology Enhanced Learning, 2018.

[6] Nwafor, Chidinma Onyenwe, Ikechukwu. "An Automated Multiple-Choice Question Generation using Natural Language Processing Techniques." International Journal on Natural Language Computing, 2021.
[7] Ayako Hoshino and Hiroshi Nakagawa. "A real-time multiple-choice question generation for language testing: A preliminary study." EdAppsNLP 05: Proceedings of the second workshop on Building Educational Applications Using NLP, 2005.
[8] Husam Ali, Ylli as Chali, Sadid A. Hasan. "Automatic Question Generation from Sentences.", 2010.