Grenze International Journal of Engineering and Technology, June Issue

Enhanced Subjective Question Generation using Transformer: A Cross-Modal Approach

Prasad Chate[1], Snehal Rathi[2], Pallavi Rege[3], Sonal Jamdade[4], Gaurav Desai[5] and S. Atharva Chaudhari[6]
[1,2,3-5,6]Vishwakarma Institute of Information Technology, Pune, Maharashtra, India
Email: {prasad.22110212, snehal.rathi, pallavi.rege, gaurav.22110202, atharva.22110675}@viit.ac.in
[4]Bharati Vidyapeeth Deemed to be University College of Engineering, Pune
Email: srjamdade@bvucoep.edu.in

Abstract*—Automated Question Generation AQG Suites have burgeoned as a developing domain in applications and natural language processing with emphasis on education, corporate training, and research. Such remarkable advances have been made in generating objective questions, while subjects or open-ended questions generation is among the few tasks very difficult to bring about. Mostly, because such questions encourage students to think critically and to understand concepts. This study proposes creating such a system that would derive subjective questions from different modalities such as text, image, pdf, and video using the google t5 transformer. The proposed system uses an optical character recognition (OCR) for extraction of texts from images, an automatic speech recognizer for converting video sound files into text, and using pdf parsing techniques in processing the document files. Such techniques include dependency parsing, semantic role labelling, and named entity recognition (NER)-all of them ensures that the created question is contextually relevant and the pertaining concepts already found in the question can be of a very high contribution value. Develop an AQG system that will produce subjective questions from multi-modal input (text, images, PDF, video). Use the Google T5 transformer model to create activated subjective questions. Integrate OCR, ASR, and PDF parsing for more effective processing of multi-modal data. Apply advanced NLP techniques (dependency parsing, semantic role labeling, NER) to ensure contextual accuracy in subjective question generation. Enable flexible generation of questions according to user preference on type and diversity for subjective question generation.

***Keywords*—Automated Question Generation, Subjective Questions, T5 Model, NLP.**

I. INTRODUCTION

Automated Question Generation is a significant development in education, corporate training, and research, especially in generating subjective or open-ended questions to elicit critical thinking and deeper understanding. Although the overall objective question production has done a fair bit of development in completeness, producing valuable subjective questions is still a tough task owing to their complexity and requirements for contextual precision.

Here is an innovative system now proposed for using the Google T5 transformer model to generate questions subjective from different input modalities, including text, images, sounds, and PDFs. It presents techniques like optical character recognition (OCR), automatic speech recognition (ASR), and PDF parsing in its efforts to source relevant text and context from different platforms.

Grenze ID: 01.GIJET.11.2.147_20

The latest NLP techniques such as dependency parsing, semantic role labeling, and named entity recognition work together to ensure that the produced questions are contextually relevant. Beyond classrooms, most likely include corporate training as well as self-assessment, still have challenges regarding ambiguity about the data and system scalability across different languages and domains.
Automated Question Generation -AQG- is landmark advancement in education, corporate trainership, and research to create subjective or open-ended questions intended to provoke critical thinking and deeper conception. Indeed, with much improvement towards completeness in objective question production, the generation of very valuable subjective questions remains a struggle owing to their complex nature and requirement for contextual specificity. A new system is being developed that will deploy the Google T5 transformer model for generating subjective questions from multi-modality inputs, namely, text modalities, images, video and PDF files. This system comprises techniques like optical character recognition (OCR), automatic speech recognition (ASR), and parsing PDF documents to extract relevant text and context from disparate sources. All advanced NLP techniques like dependency parsing, semantic role labeling, and named entity recognition are out there to ensure that the questions asked are relevant in context. Powers more than the colleges most likely include corporate training developed and self-assessment, although the challenges are still with ambiguous data, and scalability for the system across many languages and various domains.

II. LITERATURE SURVEY

Automated Question Generation (AQG) is such a modern field in NLP that claims to create and grow today at a fruitful pace in various areas of application in education, corporate training, research, and content curation. The advancement in AQG concerning other kinds of approaches such as rule-based, neural network-based, transformer-based, and key contributions and challenges faced by them has been highlighted here.

A. Rule-Based Approaches

Rule-based technologies are the oldest ones under AQG. These systems apply predefined rules of linguistics and are extremely computationally efficient without any of the usual problems of scaling and specificity:

- The main reason for lack of performance is because rule-based question generation by Ali Meftah et al. demands rule crafting which is by far the most manual effort. [1].
- Neural-based and rule-based techniques thus have been summarized by Pratik Gupta et al., the main disadvantages being simplicity of rule-based approaches and their potential for generalisability [2].
- They would be Contextually Specific Questions for Educational Games but sadly, could not become flexible with most of the game formats. [3].

B. Neural Network-Based Approaches

At the same time, the task of neural network models indeed has experienced advances in automatic question generation in terms of automating the pattern learned and generating diversified questions, but it is still one of the more demanding and constraining processes in terms of high computational demands. Anuj Goyal et al. have compared retrieval and generation models, and the conclusion is that generative models work on a sequence-to-sequence paradigm with attention features; hence, they are best suited for the generation of more diverse questions [4]. Lili Mou et al. proposed a model that is GAN based for mapped questions but needs too much tuning for it to fit other contexts [5]. Sudha Morwald et al. gave an idea of generation of question with explanation [6]. GEN along with porous priors is a system that learns implicitly through feedback during a question formation process and iteratively improves it employing the user amendments as told by Hugo Rodrigues et al. [7].

C. Transformer-Based Approaches

AQG is tremendously scaled to context-aware accuracy by transformers. Prakhar Gupta with Anish Acharya proved the fact that context is associated with the quality of questions posed by using finely tuned models of T5 and BERT. [8]. The S. Rathi et al. describe systems that are multimodal in nature and develop an AQG-from any input-be it text, image, pdf, or video. Generated using OCR, ASR, and fine-tuned transformers, the objective was to generate both objective and scenario-type questions and hold that capability for serving educational and industrial purposes. [9]. Ayan Kumar Bhowmick developed a modular framework for the generation of multiple-choice questions using transformer-based models, which has been proved highly effective in educational consumption. [10]. Asahi Ushio et al. described the QG-Bench, a multi-lingual benchmark for the purpose of question generation, which shows how transformer-based systems adapt very well to different domains and languages. [11]. Jincheng Zhou has suggested a complete end-to-end system named QOG for generating

question-option pairs which is based on the sequence-to-sequence language models for more efficient and scalable question generation. [12].
Prakhar Gupta and others have examined the possibilities of prompt patterns for improving the generation of questions regarding their educational relevance. [13].

III. METHODOLOGY

A. Data Extraction

The system already has many versatile options for processing a vast range of input formats, beginning with text files and extending to images and videos. This enables its data capture to be comprehensive. Although specific tools such as PyPDF2 will be used to extract content directly from PDF files for text-based documents, OCR would be required for image type PDFs to read the text within the images. Tesseract OCR comes to assist in such cases: to process images from which a supposed textual description is derived. In the case of exposing videos, they would be sampled after a given time interval and output images from the sampling would be processed with OCR to extract the text from the visual contents within the video. Such multi-modal extractions are shown in figure 1, and will thus be capable of feeding any kind of text variation encoded in multiple formats, thus making the device robust across many data inputs.

B. Preprocessing

The preprocessing stage readies the material to security question generation after it has been taken from the original text. This includes cleaning and organizing the text so that the input is structured and meaningful, in particular tokenization, where the text is broken down into individual sentences or tokens. Further, grammatical constructs such as parts of speech tags and noun phrases would be identified using linguistic processing techniques.
For example, if the input text is "Paris is the capital of France," noun phrases such as "Paris" and "France" can be extracted from the text. They are required for formulating objective-type questions such as fill-in-the-blank questions. Also, irrelevant or vague content would not make it into the process as a contribution to event generation; it would contribute to improvement by keeping things vital to that process regarding event generation to questions, thus improving the relevance and the quality of those resulting questions. The global pre-processing workflow is captured in the system flow chart in Figure 3.

C. Question Generation Using Transformers

That's indeed a very specific and generated discussion concerning the T5 model series employed in the Automatic Question Generation system, more so the subjective question generation, which is not the easiest thing in material understanding, along the lines of "why," "how," "explain," or "discuss," which ask in a very intuitive way for understanding view and detail input dissection.
An encyclopaedia-type question is developed in a model that allows for the use of the encoder-decoder framework together with the attention mechanism for the subjectivity-related contextualization of generated relevant subjective questions.
With T5, context-relevant subjective questions are generated in the Automatic Question Generation system, mainly dependent on the subjective question generation. Subjective questions necessitate the comprehension of relations alongside complexity to frame questions like "why," "how," "explain," or "discuss," which are highly contextually driven and input-analyzed.
It is for such an understanding between quite complex relations and the formulation of the questions like "why," "how," "explain," "discuss," into which understanding ties the context and the detailed input analysis that the Automatic Question Generation (AQG) model is using the T5 (Text-to-Text Transfer Transformer) here mainly for generating subjective-type questions.
The strength of this T5 definitely lies mainly with its encoder-decoder model and attention mechanism in producing contextually rich and relevant subjective questions.
Steps for Subjective Question Generation Using T5 Transformer
The input is actually text in the format of a 'text-to-text' task that can have specific instructions to have it generate subjective questions through the T5 model. It foreshadows the task, making the model look for ways to create open-ended questions.

Input Format for T5:
Text: "The Eiffel Tower is an iconic landmark in Paris, known for its architectural brilliance and historical significance."

Encoder Processing
This input is processed by T5 encoder and converted into a compact dense vector representation. Through that, the meaning of semantically and syntactically captures the input text in a way such that it would allow the model to know what this input represents in relation to key concepts and other such detail.
The encoder identifies critical elements like "Eiffel Tower," "iconic landmark," and "Paris."

Attention Mechanism
The attention mechanism of the model presents itself in holding concentration on all or some important passages of input text and framing questions contextually fitting, for example; "iconic landmark" and "architectural brilliance."

Decoder Question Generation
The T5 decoder generates the output in the form of subjective questions. Based on the context, the model creates open-ended questions that encourage explanation or elaboration.
Generated Question: "Why is the Eiffel Tower considered an iconic landmark in Paris?"

Fine-Tuning for Subjective Question Generation
The T5 model is fine-tuned using large datasets containing input-output pairs of text and subjective questions. This training ensures that the model generates high-quality questions tailored to various contexts.

Training Pair Example:
Input: "Isaac Newton discovered the law of gravity, revolutionizing the field of physics."
Output: "How did Isaac Newton's discovery of the law of gravity revolutionize physics?"
Examples of T5-Generated Subjective Questions
Input Text: "The Industrial Revolution marked a significant turning point in history, leading to major advancements in technology and economy."
Generated Question: "What were the key technological and economic impacts of the Industrial Revolution?"

D. Objective questions using NLP

Parts-of-Speech (POS) Tagging
Identify grammatical constructs in the input text, such as nouns, verbs, and adjectives, to locate key elements for question generation.
Noun Phrase Identification
Extract noun phrases, which are often the main subjects or objects in the text, to serve as candidates for blank spaces in fill-in-the-blank questions.
Dependency Parsing
Analyze the grammatical structure of sentences to determine relationships between words (e.g., subject, predicate, object) and select meaningful segments to transform into blanks.
Generating Fill-in-the-Blank Questions
Replace key noun phrases or other relevant parts of the text with blank spaces to create fill-in-the-blank questions.

E. Question Selection

A selection mechanism ensures that questions are generated according to the preferences of and diversity between the users.
The users then choose the number and types of questions (for example, subjective or objective) that they would like generated. Randomization techniques can derive even questions from the same input text: event questions ensure that the generated question sets remain new and nonredundant.
The user flexibility regarding this issue is also reflected in the user interface of the system as exemplified in Figures 2, 4, 5.

F. Integration with Flask Framework

The user interface is made quite simple and in Figure 2, it shows how easy it is for a user to set preferences with the results that come from it.
The sequence diagram, as seen in Figure 1, will lay out the complete end-to-end interaction workflow for a user in the system from uploading files to having questions generated. In addition, the generated questions are also presented quite systematically, for instance, there are examples of objective-type questions along with options for multiple choice shown in figure 4, and for subjective questions also.

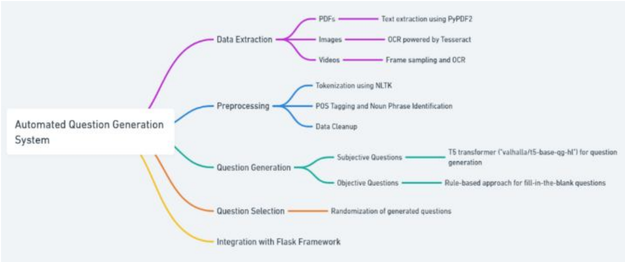

Figure 1 Flow Diagram

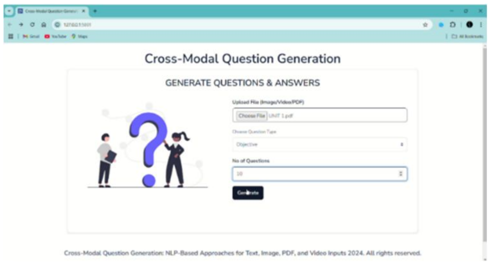

Figure 2. User Input UI

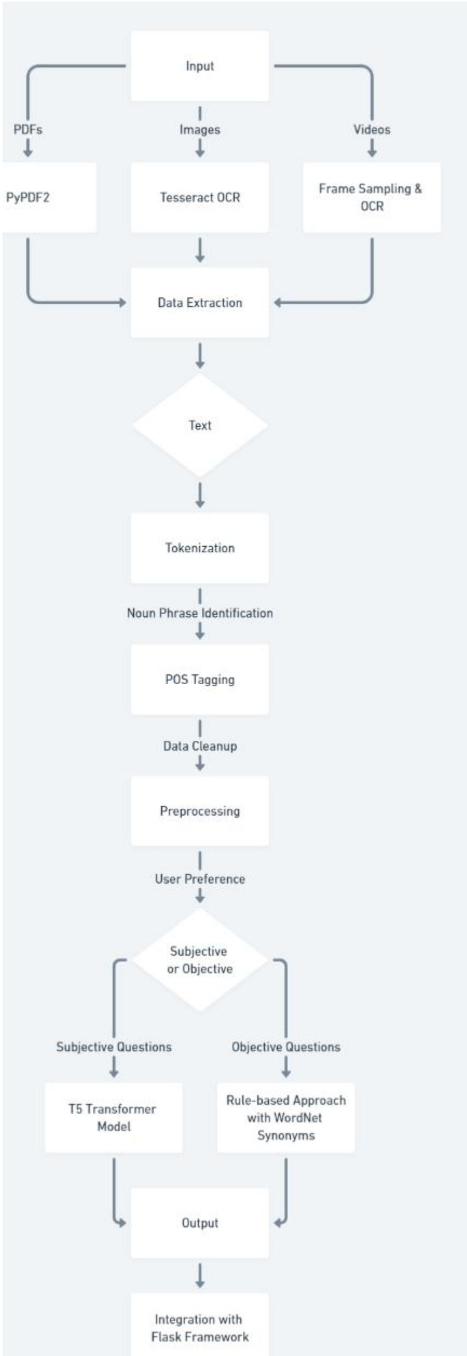

Figure 3

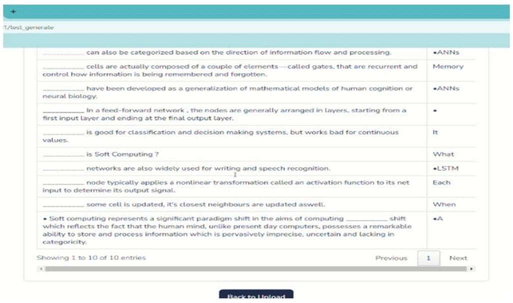

Figure 4 Generated Objective Questions

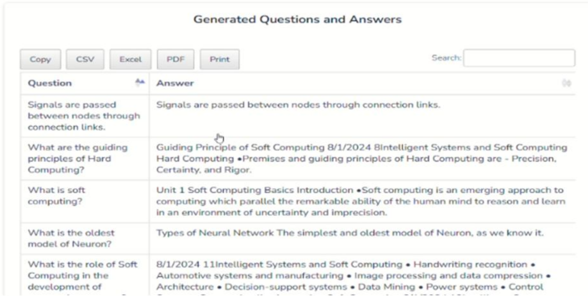

Figure 5 Generated subjective Questions

IV. Conclusion

AQG has a system that boasts more and generates subjective and objective questions from a variety of media: from text to PDFs, images, and video-including proudly strutting forward for Automatic Question Generation. It employs state-of-the-art NLP tools like T5 transformer model for the subjective part and rule-based methodologies for the objective to guarantee contextual relevance, grammatical correctness, and diversity in output. It integrated OCR and video frame processing with other modalities which would suit educational as well as corporate research and training needs. Whether suffering from ambiguity of the data or scaling in multilingual and specific domains, the system would still find its way to crossing a bar of high relevance and flexibility. Future enhancements-such as multilingual interest-or large dataset training, could only go on to revolutionize the AQG field further and greatly contribute to personalized learning and assessment by possible standards.

References

[1] Meftah, A., Amari, N., & Aouameur, D. (2018). Automated generation of contextual questions. International Journal of Computer Applications, 179(16), 1-5

[2] Gupta, P., Mehta, I., Rao, V. A., Ekbal, A., & Bhattacharyya, P. (2019). A comprehensive review of question generation: Approaches and associated challenges. Information Fusion, 60, 67-85.

[3] French, R. M., Miller, P. A., Black, W. J., & Evens, M. W. (2012). Generating natural language questions to support learning in educational games. In S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia (Eds.), Intelligent Tutoring Systems (pp. 45-54).

[4] Goyal, A. K., Arora, S., & Vig, L. (2019). Question generation from paragraphs: A tale of two models. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 4530-4540).

[5] Mou, L., Meng, Z., Yan, R., Li, G., Xu, Y., Zhang, L., & Jin, Z. (2017). Championing adversarial techniques for crafting natural language queries. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (pp. 2921-2931). Association for Computational Linguistics.

[6] Morwal, S., Hazarika, D., Majumder, N., Ghosh, S., & Mihalcea, R. (2019). Question generation from sentences. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (pp. 6205-6210). Association for Computational Linguistics.

[7] Rodrigues, H., Nyberg, E., & Coheur, L. (2023). Using implicit feedback to improve question generation. arXiv preprint arXiv:2301.12345.

[8] Gupta, P., Acharya, A., Monti, E., Garg, V., & Khurana, S. (2021). Question generation from documents with pre-trained language models. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 1234-1245). Association for Computational Linguistics.

[9] Rathi, S., Chate, P., Desai, G., Gangji, O., Kale, V., & Kalbhor, A. (2023). Cross-modal question generation: NLP-based approaches for text, image, PDF, and video inputs. In Proceedings of the 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA) (pp. 1-6). IEEE.

[10] Bhowmick, A. K., et al. (2023). Automating question generation from educational text. In Proceedings of the 2023 Conference on Artificial Intelligence in Education (pp. 234-245).

[11] Ushio, A., Alva-Manchego, F., & Camacho-Collados, J. (2022). Generative language models for paragraph-level question generation. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (pp. 5678-5689). Association for Computational Linguistics.

[12] Zhou, J. (2024). QOG: Question and options generation based on language model. In Proceedings of the 2024 Conference on Natural Language Processing (pp. 789-798). Association for Computational Linguistics.

[13] Gupta, P. (2024). Exploring prompt patterns for generative artificial intelligence in education. In Proceedings of the 2024 Conference on Educational Data Mining (pp. 345-356). International Educational Data Mining Society.