AGeES: Automatic Multiple Choice Question (MCQ) Generation from Extractive Summary of Video Lectures Using BertSum

Karthika Devi Muthusamy Sellamuthu[(✉)], Barath Srinivasan Basavaraj,
Lalit Arvind Balaji, Bharath Mohan, and Baskaran Ramachandran

Department of Computer Science and Engineering, College of Engineering Guindy,
Anna University, Chennai 600025, Tamil Nadu, India
`{mskarthikadevi2014,baaski}@annauniv.edu`,
`bbarathsrinivasan@gmail.com`, `lalitarvind.balaji@gmail.com`,
`bharathfsca@gmail.com`

Abstract. Multiple-choice questions (MCQs) from online lectures can benefit students in a number of ways. MCQs are a quick and accurate technique to evaluate knowledge and understanding since they can rapidly and accurately reinforce important topics discussed in the lectures. They are adaptable and time-effective, and they encourage students to study the content covered in class. Existing research focuses on generating MCQs from textbook data which are typically well formatted and structural in the way the information is conveyed, unlike Youtube or any other educational videos where typically, the information is conveyed in an informal or casual manner without upholding the appropriate formal structures or grammar of the sentences. The present technologies for generating MCQs have a number of limitations which include the ability to create MCQs from only highly informative textual content, employing a singular approach for sentence selection, and inefficient keyword extraction for scientific data. The proposed work AGeES overcomes the challenges of generating efficient MCQs from less in-formative content retrieved from video lectures. To enhance the effectiveness of MCQ generation, the proposed method uses a dual approach for sentence selection and works well with generation for scientific data as well. In conclusion, the proposed work overcomes the structural differences that hinder the system's ability to isolate factual information from the video. As a result, this methodology generates coherent MCQs by extracting transcripts from video lectures followed by a dual sentence selection approach.

Keywords: Topic Modeling · BERT · Keyword Extraction · T5 transformer · Sense2Vec · coreference resolution · MCQ Generation

1 Introduction

The Covid-19 pandemic has indeed brought a sharp change in the education system throughout the world. The imposition of the lockdown led to the halt of the physical mode of classes and then made online classes the new norm. Although online classes have been successful in keeping education alive in these difficult times, they cannot completely replace physical classrooms.

However, online lectures are still preferred even though the pandemic has taken a step back. This is mainly due to the fact that online lectures provide better flexibility and comfort to both the students as well the teachers. In addition to the flexibility and comfort factors, the online mode of classes tends to be much more cost-effective and time-saving, as compared to the physical mode. These factors combinedly highlight the fact that online lectures will still continue for a number of years in the future.

In fact, online lectures are not just limited to the conventional education system. An immense number of workshops across the world are held online. A number of online learning platforms, like Coursera, Udemy, Udacity, etc., have existed for quite a long time now. This makes us understand the importance and the widespread utility of online lectures.

A vast proportion of the student community pays no heed to online lectures. This not just wastes the time of the faculties who invest their whole energy in teaching but will also bring down the quality of education over time. Consequently, a steep drop can be observed in the grades of students.

In addition to a drop in their grades, the sense of procrastination in the students is associated to bring in emotional and psychological pressure. Their productivity takes a big hit and this automatically brings down their motivation levels in accomplishing things. Moreover, students are inclined towards a bigger problem of poor time management, which will make their lives difficult in the future. Eventually, the students will not be finished products, by the time they graduate from the institution.

A solution to this problem will be to ensure that the students pay attention to the online lectures. It is utterly impractical to constantly go behind the students to demand their attention, especially at times when they cannot be observed. An interactive session can prove to be useful in having the students engaged and not procrastinating. However, this is a massive task in the present era because the attention level of youngsters has decreased due to the advent of technology. Moreover, it depends on the faculties in holding interactive online lectures.

A better idea for this problem will be conducting a small assessment based on the lectures. In fact, the assessments can be conducted immediately after the online lecture ends. Although it is up to the faculties to record these scores, this will ensure that students stay attentive throughout the lectures. In addition to this, these assessments will help teachers to get a better idea about how well they teach and the areas where they can improve.

The type of assessment will play a vital role in the way in which the students will respond. An assessment consisting of descriptive-type questions will prove to be ineffective. Expecting students to come up with answers to such questions is quite impractical. Due to this, an assessment with objective-type questions will fulfill the purpose.

A major problem, however, lies with the implementation of this solution. Within a limited timeframe, it is impossible for the faculties to come up with assessments of their own. Including questions about the concepts not discussed in the lecture can also prove to be a big problem. This will create a huge ruckus among the students and can even bring these assessments to an end.

In this methodology, we tend to address all these shortcomings and build a system that automatically generates multiple-choice questions based on the lectures. The aim

is to generate this e-assessment within a short span of time so that the students can immediately attend the test.

The proposed system intends to extract the audio from the online video lectures and generate transcripts for this audio. This is followed by the provision of the transcript to two modules – one which takes care of text pre-processing and the other is an implementation of the BERTSum model, which takes care of performing text summarization. The pre-processed text is then fed as input to a module that deals with the selection of important sentences to frame questions. Both these modules give a set of keywords as their output. We filter this set of keywords, which is basically the set of answers for the multiple-choice questions that we will generate further. The next task will be to generate incorrect options for the MCQs aka distractors. Similar to the selection of important sentences in the transcript, we make use of two methods in the process of generating distractors. This is followed by the module which deals with the formation of questions. The module makes use of the important sentences (that were found earlier) and the generated distractors, as input. Evaluation of this assessment will take place once the students submit their answers.

2   Related Work

AGeES aims to develop an automatic Multiple Choice Question generator based on the video lectures. Many research papers have dealt with the topic of automated MCQ generation, sentence selection, distractor generation, and question formation.

Chen Liang et al. [4] presented an adversarial training framework consisting of a generator G and a discriminator D for Distractor Generation, along with a cascaded learning framework to improve performance. However, the model lacks an appropriate user interface. Meanwhile, CH et al. [3] discussed a partially subject-independent pipeline for automatically generating middle school-level multiple-choice questions from textbooks, which includes preprocessing, sentence and key selection, and distractor generation, utilizing various techniques such as entity recognition, WordNet, and neural embeddings. Experimental results show that the system is capable of generating high-quality questions, although it struggles with complex, multi-line questions.

Ma et al. [10] proposed a model for extractive and abstractive text summarization that combines BERT's architecture with topic-embedding information to improve contextual information capture. This approach produces high-quality summaries through NTM inferring, using a combination of token embedding, segment embedding, position embedding, and topic embedding. The two-stage model shares information to generate salient summaries while reducing redundancy. The analysis shows that the model can generate consistent summaries with high quality, but may struggle with longer articles containing multiple topics. On the other hand, Nwafor et al. [15] presented an NLP-based system for automatic MCQG in CBTE. This system uses NLP techniques to extract keywords from lesson materials, which are then used to generate exam questions. The system was found to be effective at extracting keywords and generating exam questions, but MCQs based solely on extracted keywords were not found to be efficient for exams.

Mukta Majumder et al. [11] form MCQs by performing PTM or Parse Tree Matching with test sentences, employing topic modeling to filter the sentences according to

topics. NER is used to identify the keywords and gazetteer lists to generate distractors. The system has a tendency to exclude time and date-related information and selects incomplete sentences.

Xian Wu et al. [17] utilize a contextual encoder and attention mechanism to generate semantic representations for text materials, while also introducing two modules to guarantee incorrectness and generate diverse distractors using beam search. However, the model is limited in its ability to generate distractors that require multi-sentence/hop reasoning. Meanwhile, Dhanya et al. [6] propose a system that automates the processes of sentence and key selection, question formation, and distractor generation by using Google T5's sequence-to-sequence approach, tesseract for text extraction, context recognition, and Sense2Vec for distractor generation. While the system produces high-quality work and reduces human intervention, the percentage of relevant auto-generated incorrect options is low.

Mehta et al. [13] introduced a system based on Google's BERT Model to create auautomated questions, generate summaries using BERTSUM, and generate distractors using the WordNet approach. However, the WordNet approach may not be effective in all cases. Maniar et al. [12] proposed an approach that uses transformers to paraphrase the input text before generating MCQs, which are graded using image processing. However, the model may generate multiple MCQs from the same line if the input text has fewer sentences than the desired number of questions. Ming Liu et al. [14] presented a mixed similarity strategy for generating Chinese multiple-choice distractors using a statistical regression model that considers appearance, pronunciation, and semantic meanings. Although the proposed strategy outperforms the common distractor generation strategies, it faces difficulties in extracting semantic distance features of the characters, which are not available in the knowledge base.

Dmytro Kalpakchi et al. [2] have fine-tuned a pre-trained BERT2 for Swedish for distractor generation. Two linear layers with layer normalization and a softmax activation layer had been added on top of BERT2. They have also proposed an effective method to evaluate the generated MCQs. However, only half of the distractors generated using the model were plausible.

Animesh Srivastava et al. [1] proposed a pipeline that integrates natural language processing and image captioning techniques to generate questions, answers, and distractors for both textual and visual inputs. While the system has demonstrated remarkable performance, it is suggested that further improvements can be made to the captioning dataset to enhance the question generation model. Meanwhile, Selvia Ferdiana Kusuma et al. [16] introduced an ontology-based approach that can automatically generate 11 categories of questions. By breaking down all ontology information into categories and converting them into SPARQL queries, questions can be generated with an accuracy of 86%. However, the success of this method heavily relies on the completeness of the ontology information.

Jiaying Lu et al. [7] propose a reinforcement learning-based framework called GOB-BET. The framework makes use of pre-trained Visual Question Answering models as an alternative knowledge base to guide the distractor generation process. The performance degradation of existing VQA models is utilized for detecting the quality of generated distractors. The utility of the distractors that are generated is exhibited through data

augmentation experiments. The sparsity of training samples, however, proves to be a major challenge to the framework.

Ainuddin Faizan et al. [9]'s approach uses semantic annotation to find named entities in the slide content and utilizes property information of the entities to generate questions and find appropriate distractors using SPARQL queries. SPARQL is also used to retrieve further information about the entities in the form of RDF triples, which are then verbalized to form the question text. The model has instances where the resource is not identified and inaccurate distractors are produced.

Devi, M.K. et al. [5] utilize a neural network model to extract important information from the comic book cover and generate a brief summary of the story. The system is trained on a large dataset of comic book covers and uses unsupervised learning techniques to identify the key elements of the story. Devi, M.K. et al. [8] generate concise and novel descriptions using unsupervised learning and semantic analysis to generate concise and novel descriptions., with promising results for various applications.
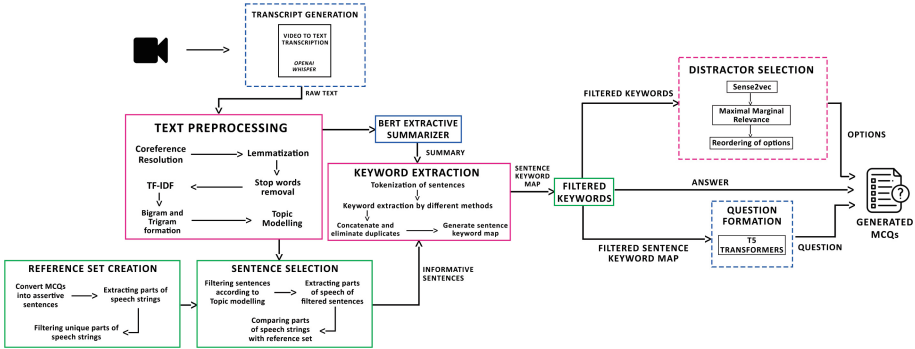
3   Methodology

Fig. 1. AGeES Architecture

3.1   Transcript Generation

MCQ generation generally uses textual content from books or journals because of its well-formatted nature. The proposed work intends to extract content from videos as a basis to generate MCQs which can be done in two ways, by extracting the transcript of the audio from the videos or extracting the textual content which is displayed in the video. AGeES focuses on generating MCQs for scientific videos which tend to have a lot of factual information which will help us overcome the challenges posed by the unstructured informal grammar forms used in these videos. These videos typically tend to read out whatever text is displayed on the screen so the transcript would be redundant even if we extracted the text displayed in the video so this work limits to only extracting the transcript of the audio belonging to the video (Fig. 1).

3.2   Preprocessing

The Preprocessing module involves two major processes: Coreference resolution and topic modeling. The transcript extracted from the video is bound to have several sentences that contain prepositions that refer to nouns in different sentences which are resolved with the use of coreference resolution. Several processing steps such as stop- words removal, lemmatization, tf-IDF, and bigram-trigram formation are used to process the transcript for topic modeling. The purpose of topic modeling is to extract topic words which would be used to filter out important sentences from the transcript. The LDA model is used to find the optimal number of topics within the transcript provided based on the coherence score of the model. Then the words from the words probability distribution with probability higher than the set threshold are selected as topic words that are used for filtration.

Algorithm 1: REFERENCE SET CREATION

  Input: MCQ Dataset
  Output: Reference set for sentence selection
1 Import MCQ question dataset
2 Extract questions and choices from json file
3 **for** *each question* **do**
4   Extract the parse tree for each MCQ question
5   Using the parse tree and answer, generate assertive sentences
6   Extract parts of speech string from each assertive sentence
7 **end**
8 **for** *each extracted parts of speech string* **do**
9   Compare it with other extracted parts of speech strings to find similarities
10   **if** *the strings are similar* **then**
11   Keep deleting the other parts of speech strings
12   **end**
13 **end**
14 **return** A set of unique parts of speech strings is obtained as a result

---

3.3   BERT Extractive Summarizer

The distil-BERTSUM model is used to generate an extractive summary from the transcript. This summary extracted does not change the structure of the sentence and it is assumed to contain the informative sentences which are to be used for the generation of MCQs.

3.4   Reference Set Creation and Sentence Selection

Sentence selection as depicted in Algorithm 2 is used to extract factual sentences from the sentences which were filtered based on topic modeling. These factual sentences are extracted by comparing them with a reference set made of existing MCQs from the SciQ dataset as depicted in Algorithm 1. The existing MCQs are converted into assertive sentences by processing and replacing any option within the question.

Algorithm 2: SENTENCE SELECTION

   Input: Processed text and Reference set
   Output: Informative Sentences
1 Filter the sentences containing the topic words extracted from LDA model
2 **for** *each filtered sentence* **do**
3   Generate Parts of speech string for each filtered sentence
4   **for** *each parts of speech string in reference set* **do**
5   **if** *POS of reference set is present as substring in POS of filtered sentence* **then**
6   we consider the filtered sentence to be valid
7   **end**
8   Extract and save the filtered sentence
9   **end**
10 **end**

These assertive sentences are then converted into POS strings which is a concatenation of parts of speech of each word within the sentence. String comparison is used to compare the POS strings of the reference set and the POS string of the filtered sentence to find a match in patterns.

3.5 Question Formation

The selected sentence and keyword are sent to the T5 transformer as input where the questions will be formed by rephrasing the sentence according to the keyword. The same sentence can be transformed into multiple questions based on the keywords identified by the keyword extraction phase.

3.6 Distractor Generation

The purpose of distractor generation is to provide false options to confuse the student. The options need to be as similar to the keyword and should be different from each other for optimal effect. AGeES uses the Sense2vec to generate a list of distractors out of which three are selected using Maximum Marginal Relevance (MMR) to make the options as diverse as possible.

4 Discussion and Results

The approach taken by the AGeES system in generating multiple-choice questions from video lectures is a promising and effective method. Unlike the traditional parse tree method used in existing MCQ generation systems, AGeES adopts a more flexible approach by creating Parts of Speech strings of the reference set and the input transcript. This enables the system to select sentences that contain factual information even in cases where the syntax and grammar are not perfect, and the speech is informal. With an increase in the number of keywords, the process of generating desired multiple-choice questions becomes simpler and more efficient. Additionally, the system's use of text preprocessing techniques like coreference resolution and topic modeling further enhances the effectiveness of the overall MCQ generation process.

A few examples of generated MCQs are given below:

1. *What is the capital of France?*
   (a) *Berlin*
   (b) *Madrid*
   (c) *Paris*
   (d) *Rome*
2. *What type of field will form when atoms gain a positive or negative charge?*
   (a) *Electric*
   (b) *Hydro*
   (c) *Pneumatic*
   (d) *A/C*
3. *What is responsible for the exchange of gases?*
   (a) *Pores*
   (b) *Sebum*
   (c) *Oiliness*
   (d) *Scalp*

The proposed MCQ generation system as in [2] addresses some of the limitations faced by previous systems in generating MCQs from video lectures. Previous systems relied on syntactic and semantic parsing of the input text, which proved to be inadequate in handling informal speech and imperfect syntax. In contrast, AGeES utilizes a flexible approach of creating Parts of Speech (POS) strings of the reference set and input transcript and checks if any POS string of the reference set is contained within the sentence from the transcript, which is then selected for further processing. This method ensures that factual information is accurately identified from video lectures, thereby improving the quality of the generated MCQs.

The preprocessing techniques in AGeES, including Co reference resolution and topic modeling, contribute to the effectiveness of the MCQ generation system. Coreference resolution, as described in [5], helps in obtaining unambiguous sentences that are easily understood by computers and in converting complex sentences into simpler ones. Topic modeling facilitates the discarding of unimportant sentences by detecting whether a sentence comes under any specific topic. These techniques improve the accuracy of identifying factual information and generating MCQs from video lectures.

The keyword extraction approach used in AGeES outperforms previous methods, as described in [2]. The proposed method uses a total of 11 keyword extraction techniques, ensuring that a higher number of keywords are extracted without compromising on their quality. This approach addresses the limitations of previous methods, such as NER's inability to identify domain-specific words and RAKE's tendency to give long phrases as keywords on odd occasions.

In summary, AGeES is an effective approach for generating MCQs from video lectures. The combination of flexible POS string comparison, multiple keyword extraction techniques, and advanced text preprocessing techniques, such as coreference resolution and topic modeling, contribute to the accuracy and effectiveness of the system. The proposed system works efficiently and represents a significant step forward in the field of MCQ generation from video lectures.

5 Conclusion and Future Works

In conclusion, AGeES can be a potent tool for not just improving the quality of education but also enhancing the processes of learning and assessment. Moreover, the system saves a large amount of time and energy for the teaching community in framing questions. Such a system has the capability to efficiently extract important material from video lectures or tutorials, produce relevant questions, and assess the responses of students in real-time. However, there are still certain issues that need to be resolved, such as assuring the accuracy and reliability of the questions, reducing bias and mistakes, and accommodating various learning preferences and styles. The question generation model can be upgraded to form multiple kinds of questions other than just 'wh' questions. The current distractor generation model faces setbacks when it comes to generating subject-specific words, which can be improved with the addition of input context. The current model is restricted to creating MCQs for science-based videos. This can be improved by expanding the reference set.

References

1. Srivastava, A., Shinde, S., Patel, N., Despande, S., Dalvi, A., Tripathi, S.: Questionator - automated question generation using deep learning. In: International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (2020)
2. Kalpakchi, D., Boye, J.: BERT-based distractor generation for Swedish reading comprehension questions using a small-scale dataset. In: 14th International Conference on Natural Language Generation (2021)
3. CH, D.R., Saha, S.K.: Generation of multiple-choice questions from textbook contents of school-level subjects. IEEE Trans. Learn. Technol. **16**, 40–52 (2022)
4. Liang, C., Yang, X., Dave, N., Wham, D., Pursel, B., Giles, C.L.: Distractor generation for multiple choice questions using learning to rank. In Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications, pp. 284–290 (2018)
5. Devi, M.K., Fathima, S., Baskaran, R.: CBCS-comic book cover synopsis: generating synopsis of a comic book with unsupervised abstractive dialogue. Procedia Comput. Sc. **172**, 701–708 (2020)
6. Dhanya, N.M., Balaji, R.K., Akash, S.: Aixam-ai assisted online MCQ generation platform using Google t5 and sense2vec. In: 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), pp. 38–44. IEEE (2022)
7. Lu, J., Ye, X., Ren, Y., Yang, Y.: Good, better, best: Textual distractors generation for multiple-choice visual question answering via reinforcement learning. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2022)
8. Karthika Devi, M.S., Fathima, S., Baskaran, R.: SYNC—Short, yet novel concise natural language description: generating a short story sequence of album images using multimodal network. In: Fong, S., Dey, N., Joshi, A. (eds.) ICT Analysis and Applications. LNNS, vol. 93, pp. 235–245. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-0630-7_23
9. Faizan, A., Lohmann, S.: Automatic generation of multiple choice questions from slide content using linked data. In: Proceedings of the 8th International Conference on Web Intelligence, Mining, and Semantics (2018)
10. Ma, T., Pan, Q., Rong, H., Qian, Y., Tian, Y., Al-Nabhan, N.: T-Bertsum: topic-aware text summarization based on BERT. IEEE Trans. Comput. Soc. Syst. **9**(3), 879–890 (2021)

11. Majumder, M., Saha, S.K.: A system for generating multiple choice questions: With a novel approach for sentence selection. In: Conference: Proceedings of the 2nd Workshop on Natural Language Processing Techniques for Educational Applications (2015)
12. Maniar, S., Parmani, J., Bodke, M., Saxena, K.: Generation and grading of arduous MCQs using NLP and OCR detection using OpenCV. In: 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–7. IEEE (2021)
13. Mehta, P.K., Jain, P., Makwana, C., Raut, C.M.: Automated MCQ generator using natural language processing. In: Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications pp. 284–290 (2021)
14. Liu, M., Rus, V., Liu, L.: Automatic Chinese multiple choice question generation using mixed similarity strategy. EEE Trans. Learn. Technol. **11**, 193–202 (2018)
15. Nwafor, C.A., Onyenwe, I.E.: An automated multiple-choice question generation using natural language processing techniques. arXiv preprint: arXiv:2103.14757 (2021)
16. Kusuma, S.F., Siahaan, D.O., Fatichah, C.: Automatic question generation in education domain based on ontology. In: International Conference on Computer Engineering, Network, and Intelligent Multimedia (2020)
17. Qiu, Z., Wu, X., Fan, W.: Automatic distractor generation for multiple choice questions in standard tests. In: Conference: Proceedings of the 28th International Conference on Computational Linguistics (2020)