

Automated Question Generation for Enhanced Learning

Siddharth Kadam, Asma Narmawala, Ananya Adarsh and Tarun Mati

Abstract

Generating meaningful questions from large text data is a crucial yet labor-intensive task in education. The ParaQG system [1] addresses this by leveraging advanced sequence-to-sequence models with dynamic dictionaries, copy mechanisms, and global sparse-max attention to generate syntactically and semantically valid questions from paragraphs. The system features an interactive interface, customizable answer selection, and robust filtering to ensure relevance and quality.

Keywords

Question Generation, Seq2Seq Model

1. Problem Statement

Manual question generation from paragraphs is time consuming and limits scalability. Most existing automated models primarily focus on single-sentence input, which fails to address real-world use cases requiring contextual understanding of entire paragraphs. To bridge this gap, a solution must support paragraph-level input, user interactivity, and efficient filtering of unanswerable questions.

2. Dataset

The ParaQG system leverages the dev-v2.0 subset of the SQuAD 2.0 dataset. This dataset comprises:

- 1,204 unique contexts.
- 11,873 questions, including both answerable and adversarially constructed unanswerable questions.

3. Evaluation

The Seq2Seq model for generating questions was evaluated using the METEOR metric due to its comprehensive approach in measuring precision, recall, and alignment between generated and reference questions. METEOR provides a more nuanced evaluation by accounting for word order, making it particularly suitable for assessing the semantic and syntactic quality of generated questions.

4. Results

The ParaQG system achieves an average METEOR score of 0.483, demonstrating its ability to produce relevant and grammatically correct questions. Additionally:

- A fine-tuned BERT model was used for filtering unanswerable questions, significantly enhancing output quality.
- The system's customizable interface improves user experience by offering precise answer selection and control over generated questions.

5. Key Challenges and Learnings

5.1. Challenges

- **Data Preprocessing:** Maintaining clean and uniform input through tokenization, punctuation handling, and BIO annotation was essential for robust model training.
- **Model Complexity:** Integrating dynamic dictionaries and sparse-max attention improved contextual understanding but increased computational overhead.

5.2. Learnings

- Dynamic dictionaries enhance the model's adaptability to out-of-vocabulary terms.
- Sparse-max attention ensures the system focuses on critical context segments for question generation.
- Robust filtering mechanisms like BERT improve the relevance and quality of the generated questions.

6. Future Works

To extend the impact and applicability of ParaQG, future efforts will focus on:

- Expanding dataset diversity to include non-English and domain-specific texts.
- Exploring multi-lingual capabilities to support a wider range of use cases.

References

- [1] V. Kumar, S. Muneeswaran, G. Ramakrishnan, Y.-F. Li, Paraqg: A system for generating questions and answers from paragraphs, arXiv preprint arXiv:1909.01642 (2019).