**Project Title:**

**MCQ Generation using Abstractive Summarization achieved by PEGASUS model**

by Ayushi Mathur

**Abstract**

A lot of sectors are suffering because of COVID, including the education industry. To continue to provide quality education, institutions needed to establish an online presence that covered all aspects of their offline operation. A student’s development needs to be assessed. A Multiple-Choice Question format provides a deep and concise check of knowledge, making it one of the most popular formats. However, in the current scenario, the questions for practice are easily available, which results in the lack of unique questions to test the students’ understanding. The majority of question papers are simply extracted from a text. This requires humans to create unique questions out of chapters, requiring a great deal of time and effort.

This research overcomes the above by automating it using natural language processing supported by abstractive summarization. The text or chapter imported by the user is transformed with the help of Pre-training with Extracted Gap-Sentences for Abstractive Summarization developed by Google AI in 2020 (PEGASUS) enabling us to get important sentences in a paraphrased manner. These sentences will be used as questions. The key-value or word which will be the answer and will be removed from the sentence is determined with the help of KeyBert which uses Bidirectional Encoder Representations from Transformers (BERT) embeddings and was developed by Maarten Grootendorst. KeyBERT considers the semantics of a word while performing keyword extraction. To create the other options of the questions similar to the answer, we use sense2vec which was trained on Reddit comments and returns distractors (word similar to our keyword/answer).

We will generate questions along with the options based on the text entered by the user by the end of the process.