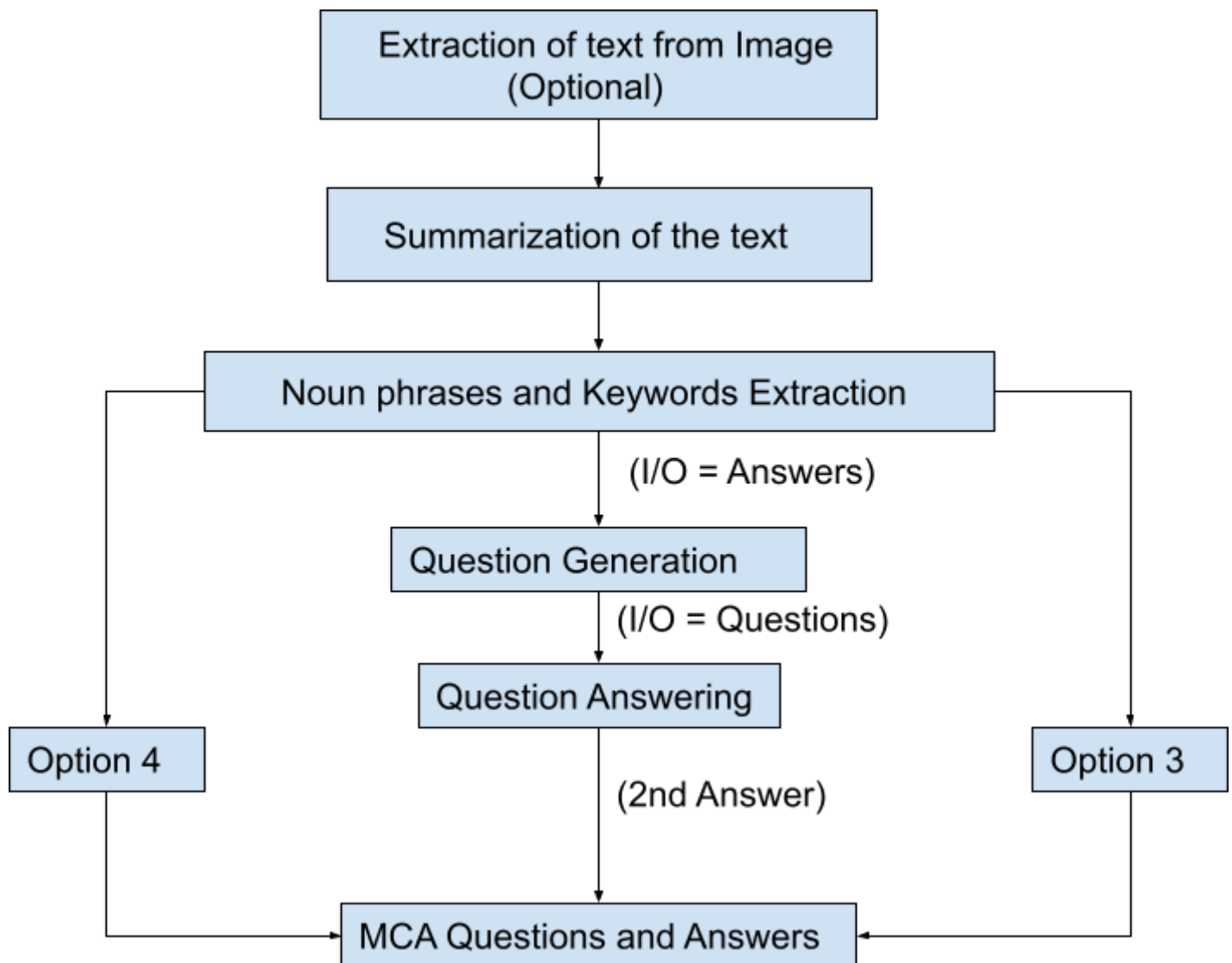


NLP ASSIGNMENT - DOCUMENTATION

The model produces MCA questions from the paragraph given. An overview of the model is depicted in the flowchart.

OVERVIEW



EXPLANATION:

- Extraction of words from the image using Pytesseract as in the problem statement image was asked to give as input.
- Summarization of the text using T5 transformers.
- Noun phrase extraction by NLTK.
- Keyword extraction by KeyBERT.
- Using Noun phrases and Keywords as answers to generate questions by Transformers library.
- Again generating answers for the question generated using Transformers to make a second answer.
- Generating random options from Noun phrases and keywords.
- MCA questions as a list of strings inside a dictionary.

CREDITS/CITATIONS:

*) Inspired by Questgen.ai.

*) Thanks and Credits to
Tutorial by Ramsrigoutham for MCQ questions generation from text.

*) @misc{mromero2021t5-base-finetuned-question-generation-ap,
title={T5 (base) fine-tuned on SQUAD for QG via AP},
author={Romero, Manuel},
publisher={Hugging Face},
journal={Hugging Face Hub},

howpublished={\url{<https://huggingface.co/mrm8488/t5-base-finetuned-question-generation-ap>}},
year={2021}
}

*) <https://github.com/MaartenGr/keyBERT>

