19/01/23

AI-BOT
AID IN WRITING AND UNDERSTANDING COMPOSITIONS

Abstract:

More and more students with visual impairments have attempted post-secondary education over the past few decades. Numerous new difficulties have been posed by this situation. One of them is the requirement for thorough, timely study of text and electronic documents. Students who are blind are unable to turn pages in a book, scan the text, or annotate with a highlighter.

In this article, we offer a solution in the form of an experimental prototype and demonstrate how NLP techniques can effectively help blind students achieve their academic goals. So, our title is AI-BOT for aiding writing and understanding composition which benefits visually impaired and other people by reducing their time for writing and understanding compositions. The methods used comprise automatic index generation, paraphrasing, summarizer, passage retrieval, and query rewriting. The paper demonstrates a technological application of a useful piece of software.

For index generation and passage retrieval, we will use GREP(Global-Regular-Expression-Print), further we will be seeing the scope of Document Term Matrix, Inverted index, Distributed Indexing and Dynamic Indexing. For Paraphrase génération we will be going through Multiple sequence alignment, Phrase-based Machine Translation, Long short-term memory and Transformers.

For Paraphrase recognition Recursive Auto-encoders, Skip-thought vectors and Transformers will be looked upon. For summarizer extractive and abstractive approaches can be explored. Query rewriting using WordNet we also be incorporated. Torch" provides deep learning algorithms while "Sentencepeice" is used to 'tokenize '(component breakdown) the text. Lastly, "Newspaper3k" is a web scraping library that is used to import articles from the internet. The 'Natural Language Toolkit 'is an NLP-based toolkit in Python that helps with text summarization.
Our proposed software aims to assist the blind community by creating a chatbot that utilizes advanced NLP techniques to facilitate information access on a specific topic. The chatbot employs automatic index creation, passage retrieval, and WordNet-based query rewriting to efficiently retrieve relevant information. The information is then read aloud to the user, allowing them to easily access useful information without the need for extensive navigation of electronic resources.

Team members:

20BCE7648 – Dhyan
20BCR7087 - Jeevantika Gurung
20BCR7106 - Shreyas M B

Keywords/Software Used:

Index generation, paraphrasing, summarizer, passage retrieval, and query rewriting, GREP (Global-Regular-Expression-Print, Long short-term memory, Transformers, Multiple sequence alignment, Phrase-based Machine Translation, Recursive Auto-encoders , Skip-thought vectors , Transformers, WordNet , Torch , Natural Language Toolkit , Sentencepeice , Newspaper3k.