Educacess - Software for Transcribing Video Lessons Into Compact PDFs Using Keyframe Extraction, Summarization and Transcription Algorithms

Carvalho, Ligia Keiko (School: Colegio Tecnico de Campinas)

To avoid suspending school activities during the Covid-19 pandemic, online video classes were used. Those classes required significant electronic equipment data storage space and internet packet data usage. This research aimed to develop a system that transcribes, summarizes and extracts the keyframes from a video class into a compact PDF. Compact PDFs can be used as an alternative material for students with limited internet connections. Therefore, three types of algorithms were developed: transcription, summarization and keyframe extraction. The algorithms evolved by going through coding and validation cycles, improving quality and processing time. To evaluate, metrics of Word Error Rate, Word Information Loss and Levenshtein Distance were used on transcription, ROUGE framework on summarization, and a qualitative and quantitative analysis on keyframe extraction. The algorithms showed positive results in technical tests and teachers validation, by creating compact PDFs with compact summaries, reducing data usage by 90% to 98%, and having acceptable transcription quality for classes with low amount of technical terms, according to the teachers. The research led to challenges and discoveries, such as: methods that allowed efficient ways to extract the main content from videos, an algorithm that splits the video in parts and merges the results to process long videos and the complexity behind implementing a proper keyframe extraction and summarization. The proposed technology assists students with a summary material from classes, without generating high costs as the algorithms and the simulated scenario of usage in schools were developed to make the summary compact. Also could be used when users need to summarize videos to turn their content into short texts.

Awards Won:

Oracle Academy: Award of $5,000 for outstanding project in the systems software category.

Fondazione Bruno Kessler: Award to participate in summer school "Web Valley"

Association for the Advancement of Artificial Intelligence: Honorable Mention

Association for the Advancement of Artificial Intelligence: AAAI Membership for the School Libraries of All 8 Winners (in-kind award / part of 1st-3rd prize and honorable mentions' prize)