

MET CS622 Final Project Description (20 points)

You need to build a web crawler that download videos from a given link on video platform (youtube, Instagram and Tiktok) and given keywords. Then, trim the video and keep the important content of the video and discard the rest of the video.

- 1) Choose a tool that can download videos based on give words, such as <https://youtube-dl.org>, <https://github.com/Litarvan/vget> or <https://github.com/drawrowfly/tiktok-scraper> , etc.
- 2) Write a script to use the video downloading tool along with the given keyword and search the video site (including Youtube, Instagram, and Tiktok). You can use <https://www.selenium.dev> to automate the download, <https://jsoup.org/>, <https://beautiful-soup-4.readthedocs.io/> or etc.

You should be able to download ten videos for a given keyword. For example, we give “bench press” as a keyword and your code should download, at least 10 videos that include “bench press” from tiktok/Instagram/youtube. Physical activities are good keywords to search, e.g. “dead lift”, “bench press”, “leg squad”, “pull up”, etc.

- 3) Write a code that automatically, splits each downloaded video based on its sense changes into several clips. This results in having more than one video from each downloaded video. You can use tools like this and use the command line, <https://scenedetect.com>
- 4) Build a GUI for the entire project that an end user can use your work.

You are free to choose your language and you might not be able to write all codes in one language, such as Java or Python. For example, you need a command line and call them from java and build a UI with React. Or you might decide to write everything as a shell script. However, the structure of your code, along with the functionality of the code, is very important for your final project grade.

NOTES:

* For the project, you do not need to make yourself familiar with Python, but you can do it in Python as well.

* The project should be done alone and not in a group.

* The deadline to deliver your final project is on the last day of the class and not the exam date.