

Course Project

Jair Vieira Morais Filho – jairv2@illinois.edu

MCS – CS410 – Fall 2021

Final Report

Theme 2: Intelligent Learning Platform

Installation

pip install spacy

python -m spacy download en_core_web_sm

python -m pip install --upgrade pymupdf

This will be used for the similarity between two documents.

Execution

After Installation - Run the command line

C:\python project.py [Video Content text files folder] [pdf file] > output.txt

Example

python project.py C:\cs410\FinalProject\Videos C:\cs410\FinalProject\Book\MITEI-The-Future-of-the-Electric-Grid.pdf > output.txt

Result

The result will be on the text file : output.txt

Using Bat File

run the bat file: c:\.project C:\cs410\FinalProject\Videos
C:\cs410\FinalProject\Book\MITEI-The-Future-of-the-Electric-Grid.pdf

Video Lecture Content and Text Book Document

Each video content is a text document and it will be calculate the similarity with each page from the Textbook The Result is the closed the Video content are from the page they are related, using that we can relate the age with each page.

The Result

The Final result is the list of videos content and the pages that content are related Video 1 - related to page 1,2, 5,10 Video 2 - related to page 3,4,11

Result is on

Workload

It page of pdf was converted to a documento using pymupdf and put in a array of documents - 4hr

It is necessary to get each video content to was transcript to a text file -6hr

It was implemented the similarity function using spacy (Package that calculate the similarity of documents) - 12hr

In the code there are a cut on 0.7 on the similarity.