Step 1) Install all the dependencies based on the requirements.txt file

Step 2) The function summarizer returns the summarized string for the videoclip passed as parameter. The sub steps in the function include:

* Extracting audio from the video and storing it in a .wav file (This is important) , with help of "moviepy.editor"
* Breakdown the audio into smaller chunks
* Use speech\_recognition library to convert speech to text and append the output to the resulting string each time
* Now that we have the text spoken in the video, we divide them into sentences and then use gensim's summarize module (another deep learning NLP library) to summarize the content of video. The gensim's summarize module internally uses :
  + Text Ranking Algorithm to find the summary (while utilizing the similarity matrix, and similarity score)

Step 3) Call the function and store the output for the 2 video's in a .json format file