

Daily Log

Monday November 18

Continued to fetch transcripts from both the code and the demo site, updating the spreadsheet dataset.

Wednesday November 20

DeepAI Summarization Model - <https://deepai.org/machine-learning-model/summarization>: Successfully imported relevant packages locally in Python and resolved some errors with API queries and JSON calls. Continued to work on implementation of algorithm for eventual text summarization. Also continued to fetch transcripts from both the code and the demo site, updating the spreadsheet dataset.

Friday November 22

MeaningCloud Summarization API - <https://www.meaningcloud.com/products/automatic-summarization>: Successfully configured parameters for API calls and tested with one example. Partial summary returned, so worked to resolve errors in calls. Continued to work on method for eventual text summarization. Also continued to fetch transcripts from both the code and the demo site, updating the spreadsheet dataset.

Timeline

Date	Goal	Met
November 8th	Discuss with Mr. White whether need to go about implementing Mellin and Discrete Cosine Transform or any other audio processing algorithms; checkpoint: have a successful summarization working and determine what next steps are if any issues arise; continue to scale dataset in parallel	No, haven't discussed with Mr. White yet, will actually do so later on when we get to comparing results of our models' preprocessing once our initial proof-of-concept is working for text summarization as well.
November 15th	Continue to research text summarization APIs and algorithms, continue to scale transcripts dataset in parallel	Yes, continued research API options, started to experiment with two, resolving errors, collected more transcripts for the dataset
November 22th	Start implementation of algorithm utilizing decided upon APIs' various methods and comparing results, continue to scale transcripts dataset in parallel	Yes, started implementation of local algorithm/general method for each of the two APIs currently under consideration, though still working to resolve lingering errors.
November 29th	Continue and approach finalization of implementation of algorithm utilizing decided upon APIs' various methods and comparing results, continue to scale transcripts dataset in parallel	
December 6th	Finalize or almost finalize implementation of algorithm utilizing decided upon APIs' various methods and comparing results, continue to scale transcripts dataset in parallel	
Winter Goal	Implement local, interactive Python code (maybe with GUI) that receives TED talk URL as input, queries IBM Watson's Speech-to-Text API, returns a transcript, summarizes it, and displays it to the user for a quick read. The app may potentially be able to work on a few limited cases of videos other than TED talks, but that may be more in the future.	

Reflection

I've run into a couple small challenges with API return calls for both the DeepAI and Meaning-Cloud APIs, but that was to be expected. Hopefully I can get sample/initial text summary results soon so that I can evaluate how good these summaries are and how the APIs compare/stack up

to one another. Then, I need to finish developing scalable algorithm(s) for the dataset for each of the API(s), in support of the core implementation for our Winter Goal.