Daily Log

Monday December 8

Started Implementing a Method to Split the Transcript into Several Sections to See if Results were Improved if Multiple Summaries were Combined

Tuesday December 9

Finished Implementing the Method for Splitting Transcripts into Sections. Ran into an Error that said Summaries cannot be Generated from Text with Less than 2 Sentences. Had to Add a Fix for this Possibility. Ran Script and Looked Through Final Summary and Saw an Overall Improvement. Summary seemed more Lengthy that Expected.

Thursday December 11

Decided to Experiment with Number of Sections to See if it Would Improve Summary Quality and Length of Summary. Best Number of Sections Ranged Between Four and Six Sections. Decided to use Five Sections for the Time Being and Tried Researching any other Possible Optimizations.

Timeline

Date	Goal	Met
Nov 22	Begin Implementing a Text Summarization Algorithm	Yes, Started Working with APIs (Text Teaser and Gensim)
Dec 6	Finish Implementing a Text Summarization Algorithm for Piece of Text	Yes, Summary Returned for Short Paragraph
Dec 13	Continue Implementing a Text Summarization Algorithm for a TED Talk	Yes, Got Better Results with New Implementation
Winter Goal	Summarize a TED Talk by Implementing a Text-Summarization API	Yes, Was Able to Summarize a TED Talk to an Extent

Reflection

I was happy to see that my idea to split the transcript into sections improved my results. As of now, my winter goal is met, but I am still looking for ways to improve my results for next week. Some ways I was thinking to do this was to keep experimenting with the number of sections or looking for other signs for the end of sentences. I might do work over the weekend to experiment further with these small changes to see if a better summary is produced. After break, I will research ways to code text summarization from scratch using different machine learning and natural language processing techniques.