

Capstone Project

Problem Statement & Preliminary Findings



Joel

Ideas

1. Fear and Greed Index (For Investors)
2. Video and Text Summarizer (For Students)
3. Automated Essay Grader (For Teachers)



Problem Statement

Students struggle with a lack of time when it comes to reading long articles or watching videos. I wish to help reduce the time it takes to **summarise** said articles/videos.

Input

YouTube URL
Article URL
Readings PDF



Output

Summary in Text
Format

What is my data source?

Topical Web Scrape

basics of economics

Intro to Economics: Crash Course Econ #1
6.7M views • 7 years ago
CrashCourse
In which Jacob Clifford and Adriene Hill launch a brand new Crash Course on Economics! So, what is economics? Good question ...
CC
What Is Economics 5 moments

Basic Concepts of Economics - Needs, Wants, Demand, Supply, Market, Utility, Price, Value, GDP, GNP
631K views • 2 years ago
Academic Gain Tutorials
This video covers the detailed discussion on the Basic Concepts of Economics. After this class, we will have generated brief idea ...
Basic Concepts of Economics | Terms we have learnt under Demand & Supply | What is Market? | ... 13 moments

Intro to Economics
Khan Academy
Learning Objectives Learning Objectives 1. How the scarcity of economic resources leads to tradeoffs. 2. How to use the concept of comparative advantage to demonstrate the potential for gains from trade. 3. How to use the ...
Learning playlist
31 videos

Google introduction to economics

<https://www.encyclopedia.com > finance > introduction...>
Introduction: What Is Economics? | Encyclopedia.com
Economics is a social science devoted to the study of how people and societies get what they need and want. Or, in more formal language, **economics** is the study ...

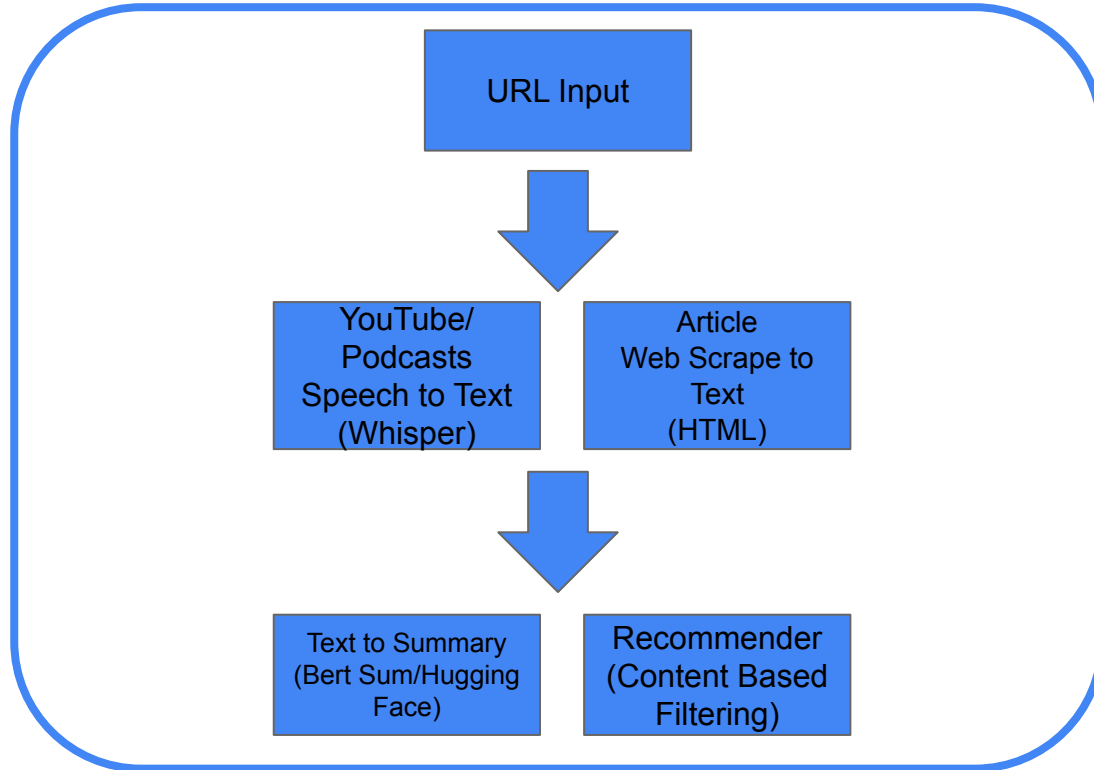
<https://openstax.org > principles-economics-2e > pages>
Ch. 1 Introduction - Principles of Economics 2e | OpenStax
Oct 11, 2560 BE — What Is **Economics**, and Why Is It Important? Microeconomics and Macroeconomics; How **Economists** Use Theories and Models to Understand **Economic** ...

<https://rlacollge.edu.in > Introduction-to-Economics | PDF>
INTRODUCTION TO ECONOMICS!!!! - the RLA College
INTRODUCTION TO. ECONOMICS. Choices, Choices, Choices, . . . **Economics** – the study of how individuals ... **ECONOMICS: 5 Economic.** Questions.
65 pages

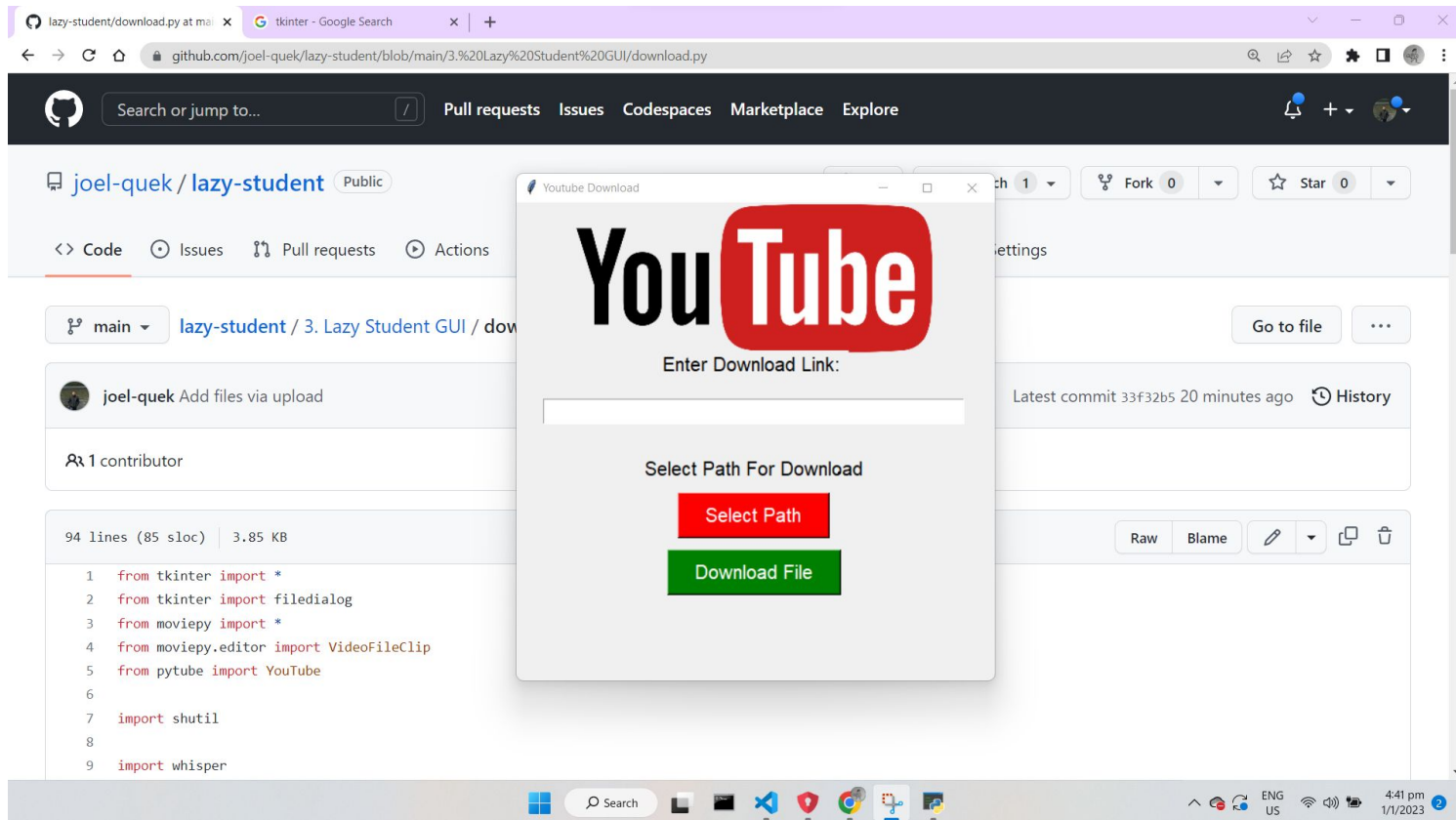
<https://www.london.ac.uk > Course Finder>
Introduction to economics EC1002 | University of London
This course will **introduce** you to the fundamentals of **economic** analysis and reasoning. It is the course upon which subsequent, more specialised **economics** ...

Project Architecture

GUI



GUI (tkinter)



YouTube Downloader (pytube)

```
1  from pytube import YouTube
2
3  def Download():
4      print("Put your YouTube link here:")
5      link = input()
6      youtubeObject=YouTube(link)
7      youtubeObject=youtubeObject.streams.get_highest_resolution()
8
9      try:
10         youtubeObject.download(filename = "downloaded")
11     except:
12         print("There has been an error in downloading this video")
13
14     print("Download Successful!")
```

Speech to Text (whisper)

```
model = whisper.load_model('base')
model_text = model.transcribe('3. Lazy Student GUI\downloaded_video', fp16=False)['text']
text_file = open(user_path + r'\transcribed_text.txt', "w")
text_file.write(model_text)
text_file.close()
```


Text to Summary (huggingface, transformers)

```
summarizer = pipeline('summarization')
with open(user_path + r'\transcribed_text.txt') as file:
    lines = file.readlines()
summary_text=summarizer(lines[0])
text_file2 = open(user_path + r"\summary.txt", "wt")
text_file2.write(summary_text[0]['summary_text'])
text_file2.close()
```

What are potential challenges?

- Too ambitious?
 - Progress has been made
- Not technically a “data science” project (one degree removed)
 - I can focus on tuning the hyperparameters of the huggingface summariser
 - The model is the summariser. I am assessing an already **pre-trained** model
- A lot of new things to learn
 - GUI especially
 - Tuning the huggingface parameters

Next Steps

1. Summariser for Non-Video Websites (text-based)
- 2. Output summarised article in point form**
3. Summariser for PDFs (KIV)

Is this a reasonable project given the time constraints?

Should be can

Thank you :)

Improvements

Make app usable from desktop (instead of python)

Summarised output to include the **max number of words** determined by the user in the GUI.

Train test data for summariser

Show and explain how huggingface summariser works (help end user to understand)

Rephrase problem statement into a business use case (reframe project as an implementation for end-user)

Overcome memory problem on my end (huggingface cant summarise past 4000 word without running into error). Define a threshold.

Improvements

- GPT summarizer
- Desktop app
- Pdf scrape