

# # Neural net speech recognition and summarization for Videos

Victor Miguel Garcia Sanchez

1 de marzo de 2024

---

Loading libraries and global variables

```
import subprocess
import moviepy.editor as mp
from moviepy.video.io.ffmpeg_tools import ffmpeg_extract_subclip
import whisper
# https://openai.com/blog/whisper/
# pip install git+https://github.com/openai/whisper.git, not
whisper.py unrelated package
from pathlib import Path
from transformers import pipeline
import os
import re

summarizer = pipeline("summarization", model="facebook/bart-large-cnn")
# https://arxiv.org/pdf/1912.08777.pdf
```

Función para determinar la duración total del video

```
def get_length(filename):
    result = subprocess.run(["ffprobe", "-v", "error", "-show_entries",
                            "format=duration", "-of",
                            "default=noprint_wrappers=1:nokey=1",
                            filename],
                            stdout=subprocess.PIPE,
                            stderr=subprocess.STDOUT)
    return float(result.stdout)
```

Function to generate the video subtitles, separating them into blocks of chunk\_size seconds. The function can be limited so that it only generates the subtitles between the start and stop seconds, and can be used as:

-VideoToSRT("video.mp4")

-VideoToSRT("video.mp4", 10)

-VideoToSRT("video.mp4", 10, 52)

-VideoToSRT("video.mp4", 5, 34, 212)

```
def VideoToSRT(filename, chunk_size=5, start=0, stop=-1):
    num_seconds_video= int(get_length(filename))+1
    print("The video is {} seconds".format(num_seconds_video))
    l=list(range(0,num_seconds_video+1,chunk_size))
    rank=range(len(l)-1)
    if stop>num_seconds_video:
        print("Upper time limit exceeds the duration of the video, it
will be considered the end of the video")
    if start>0:
        emp=int(start/chunk_size)
    else:
        emp=0
    if stop>-1:
        term=int(stop/chunk_size)+1
    else:
        term=len(l)-1
    rank=range(emp,term)
    diz={}

    for i in rank:
        ffmpeg_extract_subclip(filename, l[i]-2*(l[i]!=0), l[i+1],
targetname="chunks/cut{}.mp4".format(i+1))
        clip = mp.VideoFileClip(r"chunks/cut{}.mp4".format(i+1))

clip.audio.write_audiofile(r"converted/converted{}.wav".format(i+1))
os.remove(r"chunks/cut{}.mp4".format(i+1))

        model = whisper.load_model("base")
        result =
model.transcribe("converted/converted{}.wav".format(i+1),
language="english")
os.remove(r"converted/converted{}.wav".format(i+1))
        print("Removing converted/converted{}.wav".format(i+1))

    print("-----")
    print(round((i-emp)*100/len(rank),2),"%:")
    diz['chunk{}'.format(i+1)]=result["text"]

    l_chunks=[str(i+1)+'\
n00:'+"{:02d}".format(int(chunk_size*i/60))
+':'+"{:02d}".format(chunk_size*i%60)+' ,000 -->
'+ '00:'+"{:02d}".format(int(chunk_size*(i+1)/60))
+':'+"{:02d}".format(chunk_size*(i+1)%60-1)+' ,999\
n'+diz['chunk{}'.format(i+1)] for i in range(min(rank),(len(diz)
+min(rank)-1))]
    text='\n\n'.join(l_chunks)

    with open(filename[:-3]+'txt',mode ='w') as file:
```

```

        file.write("\n")
        file.write(text)

    p = Path(filename[:-3]+'txt')
    p.rename(p.with_suffix('.srt'))
print("100%")
print("Finally ready!")

```

## Get the video script

Function to get the complete video script at Script.txt file.

```

def VideoScript(filename):
    clip = mp.VideoFileClip(filename)
    clip.audio.write_audiofile(filename[:-3]+".wav")
    model = whisper.load_model("base")
    result = model.transcribe(filename[:-3]+".wav",
language="english")
    os.remove(filename[:-3]+".wav")
    Script=result["text"]
    with open('Script.txt',mode='w') as file:
        file.write(Script)
    print(Script)
    return(Script)

```

## Create a summary from the script

Defines the functions necessary to summarize the Script into a maximum of max\_len sentences

Running the functions

The following function only creates the video subtitles in an SRT format to test the Whisper library

```
VideoToSRT("How Clicking a Single Link Can Cost Millions.mp4",5,0,180)
```

```
The video is 861 seconds
```

```
Moviepy - Running:
```

```
>>> "+ " ".join(cmd)
```

```
Moviepy - Command successful
```

```
MoviePy - Writing audio in converted/converted1.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted1.wav
```

```
-----
```

```
0.0 %:  
Moviepy - Running:  
>>> "+" ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted2.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted2.wav  
-----
```

```
2.7 %:  
Moviepy - Running:  
>>> "+" ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted3.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted3.wav  
-----
```

```
5.41 %:  
Moviepy - Running:  
>>> "+" ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted4.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted4.wav  
-----
```

```
8.11 %:  
Moviepy - Running:  
>>> "+" ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted5.wav
```

```
MoviePy - Done.
```

Removing converted/converted5.wav

-----  
10.81 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted6.wav

MoviePy - Done.

Removing converted/converted6.wav

-----  
13.51 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted7.wav

MoviePy - Done.

Removing converted/converted7.wav

-----  
16.22 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted8.wav

MoviePy - Done.

Removing converted/converted8.wav

-----  
18.92 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted9.wav

MoviePy - Done.

Removing converted/converted9.wav

-----  
21.62 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted10.wav

MoviePy - Done.

Removing converted/converted10.wav

-----  
24.32 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted11.wav

MoviePy - Done.

Removing converted/converted11.wav

-----  
27.03 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted12.wav

MoviePy - Done.

Removing converted/converted12.wav

-----  
29.73 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted13.wav

MoviePy - Done.

Removing converted/converted13.wav

-----  
32.43 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted14.wav

MoviePy - Done.

Removing converted/converted14.wav

-----  
35.14 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted15.wav

MoviePy - Done.

Removing converted/converted15.wav

-----  
37.84 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted16.wav

MoviePy - Done.

Removing converted/converted16.wav

-----  
40.54 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted17.wav

MoviePy - Done.

Removing converted/converted17.wav

-----  
43.24 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted18.wav

MoviePy - Done.

Removing converted/converted18.wav

-----  
45.95 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted19.wav

MoviePy - Done.

Removing converted/converted19.wav

-----  
48.65 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted20.wav

MoviePy - Done.

Removing converted/converted20.wav

-----  
51.35 %:

Moviepy - Running:

>>> "+ " ".join(cmd)



```
Moviepy - Command successful
MoviePy - Writing audio in converted/converted21.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted21.wav
```

```
-----
54.05 %:
```

```
Moviepy - Running:
>>> "+ " ".join(cmd)
Moviepy - Command successful
MoviePy - Writing audio in converted/converted22.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted22.wav
```

```
-----
56.76 %:
```

```
Moviepy - Running:
>>> "+ " ".join(cmd)
Moviepy - Command successful
MoviePy - Writing audio in converted/converted23.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted23.wav
```

```
-----
59.46 %:
```

```
Moviepy - Running:
>>> "+ " ".join(cmd)
Moviepy - Command successful
MoviePy - Writing audio in converted/converted24.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted24.wav
```

```
-----
62.16 %:
```

```
Moviepy - Running:  
>>> "+ " ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted25.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted25.wav
```

```
-----  
64.86 %:  
Moviepy - Running:  
>>> "+ " ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted26.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted26.wav
```

```
-----  
67.57 %:  
Moviepy - Running:  
>>> "+ " ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted27.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted27.wav
```

```
-----  
70.27 %:  
Moviepy - Running:  
>>> "+ " ".join(cmd)  
Moviepy - Command successful  
MoviePy - Writing audio in converted/converted28.wav
```

```
MoviePy - Done.
```

Removing converted/converted28.wav

-----  
72.97 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted29.wav

MoviePy - Done.

Removing converted/converted29.wav

-----  
75.68 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted30.wav

MoviePy - Done.

Removing converted/converted30.wav

-----  
78.38 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted31.wav

MoviePy - Done.

Removing converted/converted31.wav

-----  
81.08 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted32.wav

MoviePy - Done.

Removing converted/converted32.wav

-----  
83.78 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted33.wav

MoviePy - Done.

Removing converted/converted33.wav

-----  
86.49 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted34.wav

MoviePy - Done.

Removing converted/converted34.wav

-----  
89.19 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted35.wav

MoviePy - Done.

Removing converted/converted35.wav

-----  
91.89 %:

Moviepy - Running:

>>> "+ " ".join(cmd)

Moviepy - Command successful

MoviePy - Writing audio in converted/converted36.wav

```
MoviePy - Done.
```

```
Removing converted/converted36.wav
```

```
-----  
94.59 %:
```

```
Moviepy - Running:
```

```
>>> "+ " ".join(cmd)
```

```
Moviepy - Command successful
```

```
MoviePy - Writing audio in converted/converted37.wav
```

```
MoviePy - Done.
```

```
Removing converted/converted37.wav
```

```
-----  
97.3 %:
```

```
100%
```

```
Finally ready!
```

The VideoScript function generates and prints the entire video script

```
ClickingMillions=VideoScript("How Clicking a Single Link Can Cost  
Millions.mp4")
```

```
MoviePy - Writing audio in How Clicking a Single Link Can Cost  
Millions..wav
```

```
MoviePy - Done.
```

I received a phone call from somebody who needed my help and they explained to me that this organisation suffered a cyber attack, more specifically a ransomware attack which is designed to both steal your data and make it unusable. It replicates itself throughout the business and can drive you down to paper-based controls. And this was an opportunity that I saw where I could influence something positively. And it was my job to investigate what had happened, how it happened and why. And I saw something that I hadn't experienced before firsthand. In 2017, the NHS suffered something similar and it cost nearly £100m to recover. This incident cost around £5m to recover and took 14 months. Yet what I saw was the human impact. How this happened, a single individual clicked a link and a single individual enabled this, unknowingly, to happen to an organisation. Multiple people were signed off sick due to stress and multiple people weren't able to go to work the next day and carry out their job. Now for me, cyber security is a very technological focused term. And yet IBM did a

study in 2021 and 95% of cyber attacks use the human element. Now that's all well and good but what does that actually mean? It means people can be exploited too. There's no lines of code and there's no fancy software. Cyber security is the, as far as the media is concerned, maybe teenagers and their bedrooms causing trouble stealing things and learning how to use them. Yet what people don't see is the impact and how it's day to day life. And this incident for me made me think slightly differently around cyber security. And recently I had an opportunity which presented this thought process. I was commissioned to evade security controls for a very well-known building in London. That's a snazzy way of saying break in. And effectively it was my job to see if I could get past the security controls and get into the building. And so for me, thinking kind of outside of the box, this building has floor to ceiling doors, 24, 7 security team, endless budget for this kind of thing based on where they are. And so thinking slightly outside, I needed to come up with a different plan. And what I did was I tried to go down the social engineering route which is the art of kind of deception and making people believe something without the full information. And what I did was I walked in the front door, dressed quite similarly to this. And I was greeted by eight people. And I thought, oh, it's very over the top. And it's because every single person should have the right information and should know where they're going. It's very rare for them to be visitors. And this person asked me, well, why are you here? Who are you here to see? And I explained I didn't have an appointment but I was here to see a specific person. And they said, yeah, there's no chance you're getting in. And I thought, I'll try to go this way. And yeah, what I know is people are empathetic and people want to help each other. And so I made up a story. And I said I was here for a legal matter. And I was only able to achieve what I needed to achieve on these premises. And they said, yeah, sorry, we're still. And I explained the urgency. And I made them feel sorry for me. And when I was thinking about giving this talk, I was going to pause and I was going to pretend that I was struggling. And that motion that you would have felt where you wanted to help me or you wanted me to continue is exactly how this person felt. They felt they were stopping me from doing my job, which they were. But not for how they expected it. And so I know I pretended to be on the phone in the 4A, pacing up and down, pretending to be aggravated. And then the manager came across with a QR code for me and said, so sorry, so sorry for the issues, no problem. And they showed me around a side passage away from the two rounds of security. So I had my laptop bag with me with the evidence. And it wasn't checked. And I was able to go in. And I was able to go to the floor that I needed to. And I was paid as a cybersecurity expert to evade the controls of this building. And all I did was ask for access and make someone feel sorry for me. And so that's two very different perspectives. One, the 5 million pound job and took 14 months to recover where I was helping people. But the second I was the aggressor where the person trying to get in. Now this is all enabled through the

way that humans exist and human behavior. And cybersecurity as a whole doesn't really represent that in a way that is sufficient, I don't think. And so I have one more narrative and different perspective to share. And it's when I was a victim. This happened only a few weeks ago. And what happened was I received a phone call. It was around 8 p.m. I received a phone call from a phone number. And they said, hello is this Mr. Pullin? I said yes. And they said, we've seen your bank cards be used in a different part of the country. And I thought, oh, goodness. And what they explained was they explained there's been three different transactions. And would they like to block them for me? I said, yes, please. That would be really helpful. And I googled the number out of instinct and it was the phone number from the fraud line in the bank. And something didn't add up. And I'm a bit of a pessimist. I don't really trust people. And so I was instantly on the back foot, and they're saying all of these things. And they were confirming my identity. They told me where I lived. They told my mother's maiden name. And they told me a few other bits of information that bank would know. And all of this is to build a perception of credibility. Why shouldn't I trust you? And why shouldn't you be phoning me to help me? And we go back and forth for around an hour and a half. And there was a few things that didn't sit right with me. And so when I was a hold when they were blocking my transactions, I phoned the actual fraud line and I said, is there a way that I can verify their identity? The person on the phone said, they sounded very professional on legitimate. And they were. I asked for their name and they had a fake LinkedIn profile. They had a fake crime reference number for me. And me experiencing this first hand, having investigated things like this on a regular basis for mortgages and transactions ending up in the wrong place, I knew something wasn't sitting quite right. And this, the true person put a note on my account and I explained to the person, can you tell me what the note says, please? And the first time they got a little bit flustered. It took them five minutes and they said, yeah, we'll go and check with the accounts team. But in the meantime, can you tell me the code that it says in your mobile app, at which point I can hang up, got my cards replaced, and I was okay. But these three narratives of cybercrime or scams or criminal behavior are all technology focused with the end goal, but are human led. And you may ask, how is this possible? Why can this be so easy? I've literally just walked into a building and asked someone to let me in with a fake story. And someone's phoned me up with small piece of information and built this incredible picture around, okay, yes, I should trust you. And it's because data has a value in different pockets. And with small bits of information, you can build quite a narrative, as you can see. And so today, what you would be able to do on the kind of criminal underground, if you like, would be by a thousand email addresses and passwords for around \$6 US. A cup of coffee in some places, right? That's a thousand people's account details that you may be able to log into or have tangible information to create a case. And that might be, potentially, to be

Amazon for a passable reset. It might be what location you went on holiday and we're going to do a bit more of a targeted attack that way. And this information is available because of vulnerabilities from a technical standpoint. Yeah, this is to exploit the human behaviors. Take my parents, for example, I think I'm inside of a security because my parents give me a balance. My mum is 100%, 110% optimist, nothing's going to go wrong. Everything's again, no one's going to hurt my little boy and all of this sort of stuff. And my dad's much more on the pessimist again, where, why'd you want to know me? Why'd you want this information? And so that balance for me brings kind of both sides of the story. And my mum is the sort of person that would have shared the traditional WhatsApp messages, 250 pounds of Christmas, all how lovely that would be, pay for your Christmas lunch and all of those sorts of things. And that then becomes a whole different attack vector because it's coming from someone you trust and they're sharing your link and they're sharing something you might want to click and you begin to trust even more. And so my talk is around really focusing on the ways in which human behavior is exploited and how we can benefit and protect each other. And it's okay to call these things out. And so there's some basic things you can do, such as resetting passwords and making sure you're not using the same password for your accounts. Because if one of your passwords did get leaked, you would like to know, okay, it's just this one account and I understand that's the one I need to look after. When many people use the same profile Facebook, their bank, their online banking, sorry. And sites that you can purchase things. So you might be able to go on Amazon and buy an iPhone when someone's using the password, right? Bank account details are stored. And that creates a whole different perspective of risk and cybercrime. And so for me, I don't believe any generation can avoid this anymore. Children are being raised with iPads and older generations are online shopping because of convenience and accessibility to services they may not have had before. And so I believe that understanding how these things may happen and putting some light on them can really impact the way in which people conduct themselves and challenge when things may not feel quite right. And so for me, going through this journey and those three different perspectives, the one where I was the person helping, five million pounds and seeing people really suffer. The second one where I was putting people potentially in that position, however, fully ethically and I was there meant to be there for my job. And the third where I was the victim, it shows that it can take many different shapes based on information and information can come from social media. And so if you're going on holidays in Mexico, say for your honeymoon, you've saved up all of this money. Wonderful. Have a lovely time. Yet someone you know or an acquaintance or you have public visibility of your arrangements. If someone knows that information and they know the bank you may work with, they could phone you whilst you land and say, we've seen your card be used in this location. Now, how are you going to feel if someone saying your card's being used and it's you? You're



going to feel, okay, cool. Yeah, this is me. No problem. And they say, okay, can you just confirm your identity because I want to make sure this is you. Can you just tell me your card number? So you do. And then you ask, why are you there? I'm on my honeymoon, I have a lovely time. All of these social engineering empathetic side of behaviors. And then you get down into the more conversational elements. Okay, can you just confirm your card isn't going to expire? When does it expire please? There's many different ways you can post questions to make people feel acceptance. And then lastly, can you just check the security pins or I know which card I'm going to disable? And by that time what you've done is, you've told someone you've got money in your bank because you've been saving for this wonderful occasion. And also you're not going to be in the country to do anything about it. And so from a cybersecurity perspective, exploitation can happen in many different ways and I don't think it's publicized around the human elements enough. And so if you take one thing from today, I ask that you see this as your opportunity to make sure that you protect your own information and your loved ones and your identity online. There's no problem with using social media. All I ask is you consider who you're sharing that information with. The reason being that information is valuable even if it's not to you. It could build a picture and it could cause you some trouble. Consider who you share your information with. Thank you.

## Summary

```
def AbstractiveSummarization(Script, max_len=60):
    Sentences=Script.split(". ")
    textsize=len(Sentences)
    if(textsize<=50):
        Summary=summarizer(Script, max_length=max_len+10,
min_length=max_len-10, do_sample=False)[0]["summary_text"]
    else:
        chunks=int(textsize/50)+1
        k,m=divmod(textsize,chunks)
        Summary=""
        for i in range(chunks):
            subscript=[]
            subscript.append('. '.join(Sentences[i*k+min(i,m):
(i+1)*k+min(i,m)]))
            if i>0:
                Summary+='.'
            try:
                Summary+=summarizer(subscript,
max_length=int((max_len+10)/chunks), min_length=int((max_len-
10)/chunks), do_sample=False)[0]["summary_text"]
            except:
                #Too long paragraph, splitting again
```

```

        Subsentences=subscript[0].split(". ")
        Subk,Subm=divmod(len(Subsentences),2)
        for j in range(2):
            subsubscript=[]
            subsubscript.append('.')
            '.join(Subsentences[j*Subk+min(j,Subm):(j+1)*Subk+min(j,Subm)])
            Summary+=summarizer(subsubscript,
max_length=int((max_len+10)/2*chunks), min_length=int((max_len-
10)/2*chunks), do_sample=False)[0]["summary_text"]
            Summary=Summary.replace("\'", "'")
        return(Summary)

```

AbstractiveSummarization(ClickingMillions, 60)

Token indices sequence length is longer than the specified maximum sequence length for this model (1036 > 1024). Running this sequence through the model will result in indexing errors

'In 2017, the NHS suffered something similar and it cost nearly £100m to recover. This incident. Mr Pullin was paid as a cybersecurity expert to evade the controls of this building. He says he. "This is to exploit the human behaviors. And so there's some basic things you can do, such as resetting passwords and making sure you're not using the same password for your accounts" "I don't believe any generation can avoid this anymore. Children are being raised with iPads and older generations are online shopping because of convenience and accessibility to services they may not have had before"If you're going on holidays in Mexico, say for your honeymoon, you've saved up all of this money. Have a lovely time. Yet someone you know or an acquaintance or you have public visibility of your arrangements. If someone knows that information and they know the bank you may work with, they could phone you whilst you land and say, we've seen your card be used in this location.'