

HS-MVideoSumm

Desenvolvida em Python 3 e macOS High Sierra - versão 10.13.6

Recursos necessários:

- **OpenCV**

`pip3 install opencv-contrib-python==3.4.2.16`

- **numpy**

`pip3 install numpy`

- **scipy**

`pip3 install scipy`

- **scikit-learn**

`pip3 install -U scikit-learn`

- **parse**

`pip3 install parse`

- **NLTK (Natural Language Toolkit)**

`pip3 install --user -U nltk`

Mais informações: <https://www.nltk.org/install.html>

Abra o interpretador python e digite os seguintes comandos:

```
import nltk
```

```
nltk.download()
```

Uma nova janela deve abrir, mostrando o NLTK Downloader. Altere “Download Directory”. Para instalação central, defina C:\nltk_data (Windows), /usr/local/share/nltk_data (Mac), or /usr/share/nltk_data (Unix). Então, selecione os pacotes ou coleções que deseja fazer download (Collections – all).

Mais informações: <https://www.nltk.org/data.html>

- **moviepy**

`pip3 install moviepy`

- **Acesso a Google Cloud Natural Language API**

<https://console.cloud.google.com>

Menu de navegação → APIs e serviços → Credenciais → Criar projeto → Nome do projeto → Criar

Criar credenciais → Conta de serviço → Nome da conta (o ID é preenchido automaticamente) → Criar → Concluir

Clique sobre o e-mail da conta de serviço criada → aba Chaves → Adicionar chave → Criar nova chave → selecione Tipo de chave = JSON → Criar

Um arquivo JSON é salvo em seu computador

Configure a variável de ambiente, para Linux ou macOS:

```
export GOOGLE_APPLICATION_CREDENTIALS="KEY_PATH"
```

Substitua KEY_PATH pelo caminho do arquivo JSON que contém a chave da conta de serviço.

Mais informações: <https://cloud.google.com/docs/authentication/getting-started>

- **Ativar API no projeto criado na Google Cloud Platform**

<https://console.cloud.google.com>

Com o projeto selecionado:

Menu de navegação → APIs e serviços → Painel → Ativar APIs e serviços → Cloud Natural Language API → Ativar

Obs.: Para ativar é necessário criar uma conta de faturamento

- **Bibliotecas de cliente do Cloud**

```
pip3 install google-cloud-language==1.3.0
```

Para executar:

```
python3 main.py <video_set> <video1> <video2> [<video3> ...]
```

<video_set>: nome do conjunto

<video1> <video2> [<video3> ...]: nomes dos vídeos/textos a serem processados (sem extensão)

Obs.:

- Textos de notícias relacionadas separados em subtópicos e organizados em arquivo JSON com campos:

- content: texto do subtópico;
 - begin, end: tempos de início e fim do subtópico no formato HH:MM:SS.
- Vídeos em formato MP4.

Os limiares utilizados para detectar opinião a partir dos valores magnitude e score da Google Cloud Natural Language API foram definidos em janeiro/2020 considerando a base de vídeos <https://github.com/tamires/base-de-videos>