Youtube2Speech ⬛ GitHub

This noteboo uses an [open source project](#) to transcribe a given Youtube video 😄

▼ Install DeepSpeech

```
import os
from os.path import exists, join, basename, splitext

if not exists('deepspeech-0.6.1-models'):
  !apt-get install -qq sox
  !pip install -q deepspeech-gpu==0.6.1 youtube-dl
  !wget https://github.com/mozilla/DeepSpeech/releases/download/v0.6.1/deepspeech-0.6.1-model

  !tar xvfz deepspeech-0.6.1-models.tar.gz

from IPython.display import YouTubeVideo
```
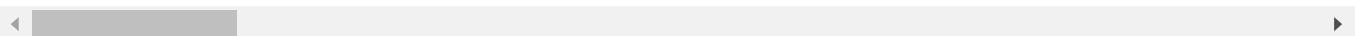
```
--2021-03-06 14:52:52--  https://github.com/mozilla/DeepSpeech/releases/download/v0.6.1/
Resolving github.com (github.com)... 140.82.121.3
Connecting to github.com (github.com)|140.82.121.3|:443... connected.
HTTP request sent, awaiting response... 302 Found
Location: https://github-releases.githubusercontent.com/60273704/f29e6300-33cd-11ea-8523
--2021-03-06 14:52:53--  https://github-releases.githubusercontent.com/60273704/f29e6300
Resolving github-releases.githubusercontent.com (github-releases.githubusercontent.com)
Connecting to github-releases.githubusercontent.com (github-releases.githubusercontent.c
HTTP request sent, awaiting response... 200 OK
Length: 1229020343 (1.1G) [application/octet-stream]
Saving to: 'deepspeech-0.6.1-models.tar.gz'

deepspeech-0.6.1-mo 100%[===================>]   1.14G  79.8MB/s    in 19s

2021-03-06 14:53:11 (62.9 MB/s) - 'deepspeech-0.6.1-models.tar.gz' saved [1229020343/122

._deepspeech-0.6.1-models
deepspeech-0.6.1-models/
deepspeech-0.6.1-models/._lm.binary
deepspeech-0.6.1-models/lm.binary
deepspeech-0.6.1-models/._output_graph.pbmm
deepspeech-0.6.1-models/output_graph.pbmm
deepspeech-0.6.1-models/._output_graph.pb
deepspeech-0.6.1-models/output_graph.pb
deepspeech-0.6.1-models/._trie
deepspeech-0.6.1-models/trie
deepspeech-0.6.1-models/output_graph.tflite
```
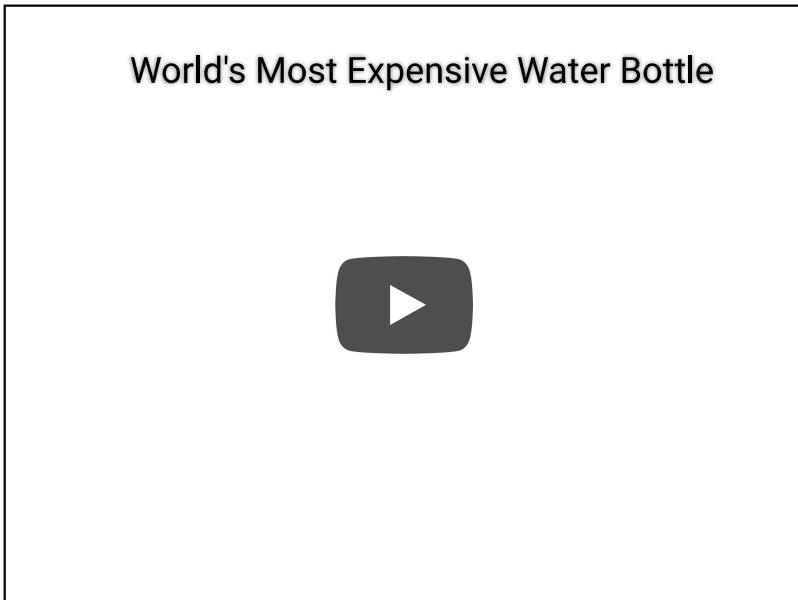
Transcribe Youtube Video

We are going to make speech recognition on the following Youtube video:

```
YOUTUBE_ID = 'e5nlD2LfGd0'

YouTubeVideo(YOUTUBE_ID)
```


World's Most Expensive Water Bottle

Download the above video, convert to a WAV file and do speech recognition 😊😒

```
!rm -rf *.wav
!youtube-dl --extract-audio --audio-format wav --output "test.%(ext)s" https://www.youtube.co
!deepspeech --model deepspeech-0.6.1-models/output_graph.pbmm --lm deepspeech-0.6.1-models/lm
```

```
    [youtube] e5nlD2LfGd0: Downloading webpage
    [download] Destination: test.webm
    [download] 100% of 459.05KiB in 00:00
    [ffmpeg] Destination: test.wav
    Deleting original file test.webm (pass -k to keep)
    Loading model from file deepspeech-0.6.1-models/output_graph.pbmm
    TensorFlow: v1.14.0-21-ge77504a
    DeepSpeech: v0.6.1-0-g3df20fe
    2021-03-06 14:57:04.400500: I tensorflow/core/platform/cpu_feature_guard.cc:142] Your CF
    2021-03-06 14:57:04.454574: I tensorflow/stream_executor/platform/default/dso_loader.cc
    2021-03-06 14:57:04.544182: E tensorflow/stream_executor/cuda/cuda_driver.cc:318] failec
    2021-03-06 14:57:04.544251: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:156] k
    Loaded model in 0.215s.
    Loading language model from files deepspeech-0.6.1-models/lm.binary deepspeech-0.6.1-moc
    Loaded language model in 0.00526s.
    Warning: original sample rate (48000) is different than 16000hz. Resampling might produc
    Running inference.
    he is a shunella bottle wire to letting the crow i have no idea was consolatory but the
    Inference took 13.864s for 84.755s audio file.
```

Follow Me

Repl.it

GitHub

CodePen

LinkedIn

LeetCode

HackerRank