```
!pip install --upgrade transformers
```

```
Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-packages (4.52.4)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from transformers) (3.18.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.30.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.33.0)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (24.2)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from transformers) (6.0.2)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2024.11.6)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from transformers) (2.32.3)
Requirement already satisfied: tokenizers<0.22,>=0.21 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.21.1)
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.5.3)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.11/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub<1.0,>=0.30.0->trans
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub<1.0,>=0.3
Requirement already satisfied: hf-xet<2.0.0,>=1.1.2 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub<1.0,>=0.30.0->t
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2025.6.1
```

```
import transformers
print(transformers.__version__)
```

```
4.52.4
```

```
import librosa
import torch
import IPython.display as display
from transformers import Wav2Vec2ForCTC, Wav2Vec2Tokenizer
import numpy as np
tokenizer = Wav2Vec2Tokenizer.from_pretrained("facebook/wav2vec2-base-960h")
model = Wav2Vec2ForCTC.from_pretrained("facebook/wav2vec2-base-960h")
```

```
/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as s
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access public models or datasets.
  warnings.warn(
tokenizer_config.json: 100%  163/163 [00:00<00:00, 9.41kB/s]
vocab.json: 100%  291/291 [00:00<00:00, 22.2kB/s]
special_tokens_map.json: 100%  85.0/85.0 [00:00<00:00, 7.62kB/s]
config.json: 100%  1.60k/1.60k [00:00<00:00, 130kB/s]
The tokenizer class you load from this checkpoint is not the same type as the class this function is called from. It may result in u
The tokenizer class you load from this checkpoint is 'Wav2Vec2CTCTokenizer'.
The class this function is called from is 'Wav2Vec2Tokenizer'.
/usr/local/lib/python3.11/dist-packages/transformers/models/wav2vec2/tokenization_wav2vec2.py:720: FutureWarning: The class `Wav2Vec
  warnings.warn(
model.safetensors: 100%  378M/378M [00:06<00:00, 66.6MB/s]
Some weights of Wav2Vec2ForCTC were not initialized from the model checkpoint at facebook/wav2vec2-base-960h and are newly initializ
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
```

```
from google.colab import files

uploaded = files.upload()
```

```
Choose Files No file chosen  Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving WhatsApp Audio 2025-06-22 at 15.53.00_0b91cce5.dat.unknown to WhatsApp Audio 2025-06-22 at 15.53.00_0b91cce5.dat.unknown
```

```
import os

os.rename("WhatsApp Audio 2025-06-22 at 15.53.00_0b91cce5.dat.unknown", "myspeech.m4a")

Audio("myspeech.m4a", autoplay=True)
```

```
0:00 / 0:07
```

```
import librosa
import soundfile as sf
```

```
audio, sr = librosa.load("myspeech.m4a", sr=16000)


sf.write("myspeech.wav", audio, sr)
```

```
/tmp/ipython-input-13-1215277700.py:5: UserWarning: PySoundFile failed. Trying audioread instead.
  audio, sr = librosa.load("myspeech.m4a", sr=16000)
```

```
import torch
import torchaudio
from transformers import Wav2Vec2ForCTC, Wav2Vec2Processor


processor = Wav2Vec2Processor.from_pretrained("facebook/wav2vec2-base-960h")
model = Wav2Vec2ForCTC.from_pretrained("facebook/wav2vec2-base-960h")


waveform, sample_rate = torchaudio.load("myspeech.wav")


if sample_rate != 16000:
    resampler = torchaudio.transforms.Resample(orig_freq=sample_rate, new_freq=16000)
    waveform = resampler(waveform)


input_values = processor(waveform.squeeze().numpy(), return_tensors="pt", sampling_rate=16000).input_values


with torch.no_grad():
    logits = model(input_values).logits


predicted_ids = torch.argmax(logits, dim=-1)
transcription = processor.decode(predicted_ids[0])


print("🗣️ Transcribed Text:\n", transcription)
```

```
Some weights of Wav2Vec2ForCTC were not initialized from the model checkpoint at facebook/wav2vec2-base-960h and are newly initiali
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
🗣️ Transcribed Text:
 HALLO THIS EASER TUSTIFIEL OF CONVERTING SPEECH TO TEXT
```

Start coding or generate with AI.