```
!pip install transformers torch datasets soundfile librosa

Requirement already satisfied: transformers in
/usr/local/lib/python3.11/dist-packages (4.48.3)
Requirement already satisfied: torch in
/usr/local/lib/python3.11/dist-packages (2.5.1+cu124)
Collecting datasets
  Downloading datasets-3.3.2-py3-none-any.whl.metadata (19 kB)
Requirement already satisfied: soundfile in
/usr/local/lib/python3.11/dist-packages (0.13.1)
Requirement already satisfied: librosa in
/usr/local/lib/python3.11/dist-packages (0.10.2.post1)
Requirement already satisfied: filelock in
/usr/local/lib/python3.11/dist-packages (from transformers) (3.17.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.24.0 in
/usr/local/lib/python3.11/dist-packages (from transformers) (0.28.1)
Requirement already satisfied: numpy>=1.17 in
/usr/local/lib/python3.11/dist-packages (from transformers) (1.26.4)
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) (0.21.0)
Requirement already satisfied: safetensors>=0.4.1 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: typing-extensions>=4.8.0 in
/usr/local/lib/python3.11/dist-packages (from torch) (4.12.2)
Requirement already satisfied: networkx in
/usr/local/lib/python3.11/dist-packages (from torch) (3.4.2)
Requirement already satisfied: jinja2 in
/usr/local/lib/python3.11/dist-packages (from torch) (3.1.5)
Requirement already satisfied: fsspec in
/usr/local/lib/python3.11/dist-packages (from torch) (2024.10.0)
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch)
  Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch)
  Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch)
  Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
```

```
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch)
  Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch)
  Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cufft-cu12==11.2.1.3 (from torch)
  Downloading nvidia_cufft_cu12-11.2.1.3-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-curand-cu12==10.3.5.147 (from torch)
  Downloading nvidia_curand_cu12-10.3.5.147-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cusolver-cu12==11.6.1.9 (from torch)
  Downloading nvidia_cusolver_cu12-11.6.1.9-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cusparse-cu12==12.3.1.170 (from torch)
  Downloading nvidia_cusparse_cu12-12.3.1.170-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in
/usr/local/lib/python3.11/dist-packages (from torch) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch) (12.4.127)
Collecting nvidia-nvjitlink-cu12==12.4.127 (from torch)
  Downloading nvidia_nvjitlink_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Requirement already satisfied: triton==3.1.0 in
/usr/local/lib/python3.11/dist-packages (from torch) (3.1.0)
Requirement already satisfied: sympy==1.13.1 in
/usr/local/lib/python3.11/dist-packages (from torch) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in
/usr/local/lib/python3.11/dist-packages (from sympy==1.13.1->torch)
(1.3.0)
Requirement already satisfied: pyarrow>=15.0.0 in
/usr/local/lib/python3.11/dist-packages (from datasets) (18.1.0)
Collecting dill<0.3.9,>=0.3.0 (from datasets)
  Downloading dill-0.3.8-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: pandas in
/usr/local/lib/python3.11/dist-packages (from datasets) (2.2.2)
Collecting xxhash (from datasets)
  Downloading xxhash-3.5.0-cp311-cp311-
manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (12 kB)
Collecting multiprocess<0.70.17 (from datasets)
  Downloading multiprocess-0.70.16-py311-none-any.whl.metadata (7.2
kB)
Requirement already satisfied: aiohttp in
/usr/local/lib/python3.11/dist-packages (from datasets) (3.11.13)
Requirement already satisfied: cffi>=1.0 in
/usr/local/lib/python3.11/dist-packages (from soundfile) (1.17.1)
Requirement already satisfied: audioread>=2.1.9 in
```

```
/usr/local/lib/python3.11/dist-packages (from librosa) (3.0.1)
Requirement already satisfied: scipy>=1.2.0 in
/usr/local/lib/python3.11/dist-packages (from librosa) (1.13.1)
Requirement already satisfied: scikit-learn>=0.20.0 in
/usr/local/lib/python3.11/dist-packages (from librosa) (1.6.1)
Requirement already satisfied: joblib>=0.14 in
/usr/local/lib/python3.11/dist-packages (from librosa) (1.4.2)
Requirement already satisfied: decorator>=4.3.0 in
/usr/local/lib/python3.11/dist-packages (from librosa) (4.4.2)
Requirement already satisfied: numba>=0.51.0 in
/usr/local/lib/python3.11/dist-packages (from librosa) (0.60.0)
Requirement already satisfied: pooch>=1.1 in
/usr/local/lib/python3.11/dist-packages (from librosa) (1.8.2)
Requirement already satisfied: soxr>=0.3.2 in
/usr/local/lib/python3.11/dist-packages (from librosa) (0.5.0.post1)
Requirement already satisfied: lazy-loader>=0.1 in
/usr/local/lib/python3.11/dist-packages (from librosa) (0.4)
Requirement already satisfied: msgpack>=1.0 in
/usr/local/lib/python3.11/dist-packages (from librosa) (1.1.0)
Requirement already satisfied: pycparser in
/usr/local/lib/python3.11/dist-packages (from cffi>=1.0->soundfile)
(2.22)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(2.4.6)
Requirement already satisfied: aiosignal>=1.1.2 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(1.3.2)
Requirement already satisfied: attrs>=17.3.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(25.1.0)
Requirement already satisfied: frozenlist>=1.1.1 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(1.5.0)
Requirement already satisfied: multidict<7.0,>=4.5 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(6.1.0)
Requirement already satisfied: propcache>=0.2.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(0.3.0)
Requirement already satisfied: yarl<2.0,>=1.17.0 in
/usr/local/lib/python3.11/dist-packages (from aiohttp->datasets)
(1.18.3)
Requirement already satisfied: llvmlite<0.44,>=0.43.0dev0 in
/usr/local/lib/python3.11/dist-packages (from numba>=0.51.0->librosa)
(0.43.0)
Requirement already satisfied: platformdirs>=2.5.0 in
/usr/local/lib/python3.11/dist-packages (from pooch>=1.1->librosa)
(4.3.6)
```

```
Requirement already satisfied: charset-normalizer<4,>=2 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers)
(3.4.1)
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)
(2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.11/dist-packages (from requests->transformers)
(2025.1.31)
Requirement already satisfied: threadpoolctl>=3.1.0 in
/usr/local/lib/python3.11/dist-packages (from scikit-learn>=0.20.0-
>librosa) (3.5.0)
Requirement already satisfied: MarkupSafe>=2.0 in
/usr/local/lib/python3.11/dist-packages (from jinja2->torch) (3.0.2)
Requirement already satisfied: python-dateutil>=2.8.2 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets)
(2.8.2)
Requirement already satisfied: pytz>=2020.1 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets)
(2025.1)
Requirement already satisfied: tzdata>=2022.7 in
/usr/local/lib/python3.11/dist-packages (from pandas->datasets)
(2025.1)
Requirement already satisfied: six>=1.5 in
/usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2-
>pandas->datasets) (1.17.0)
Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-
manylinux2014_x86_64.whl (363.4 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 363.4/363.4 MB 3.6 MB/s eta
0:00:00
anylinux2014_x86_64.whl (13.8 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 13.8/13.8 MB 25.1 MB/s eta
0:00:00
anylinux2014_x86_64.whl (24.6 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 24.6/24.6 MB 36.3 MB/s eta
0:00:00
e_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (883 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 883.7/883.7 kB 27.8 MB/s eta
0:00:00
anylinux2014_x86_64.whl (664.8 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 664.8/664.8 MB 2.5 MB/s eta
0:00:00
anylinux2014_x86_64.whl (211.5 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 211.5/211.5 MB 6.2 MB/s eta
0:00:00
anylinux2014_x86_64.whl (56.3 MB)
```

```
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 56.3/56.3 MB 12.6 MB/s eta
0:00:00
anylinux2014_x86_64.whl (127.9 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 127.9/127.9 MB 8.3 MB/s eta
0:00:00
anylinux2014_x86_64.whl (207.5 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 207.5/207.5 MB 4.9 MB/s eta
0:00:00
anylinux2014_x86_64.whl (21.1 MB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 21.1/21.1 MB 48.8 MB/s eta
0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 485.4/485.4 kB 26.8 MB/s eta
0:00:00
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 116.3/116.3 kB 9.3 MB/s eta
0:00:00
ultiprocess-0.70.16-py311-none-any.whl (143 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 143.5/143.5 kB 12.2 MB/s eta
0:00:00
anylinux_2_17_x86_64.manylinux2014_x86_64.whl (194 kB)
━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 194.8/194.8 kB 13.4 MB/s eta
0:00:00
e-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupti-cu12, nvidia-cublas-
cu12, dill, nvidia-cusparse-cu12, nvidia-cudnn-cu12, multiprocess,
nvidia-cusolver-cu12, datasets
  Attempting uninstall: nvidia-nvjitlink-cu12
    Found existing installation: nvidia-nvjitlink-cu12 12.5.82
    Uninstalling nvidia-nvjitlink-cu12-12.5.82:
      Successfully uninstalled nvidia-nvjitlink-cu12-12.5.82
  Attempting uninstall: nvidia-curand-cu12
    Found existing installation: nvidia-curand-cu12 10.3.6.82
    Uninstalling nvidia-curand-cu12-10.3.6.82:
      Successfully uninstalled nvidia-curand-cu12-10.3.6.82
  Attempting uninstall: nvidia-cufft-cu12
    Found existing installation: nvidia-cufft-cu12 11.2.3.61
    Uninstalling nvidia-cufft-cu12-11.2.3.61:
      Successfully uninstalled nvidia-cufft-cu12-11.2.3.61
  Attempting uninstall: nvidia-cuda-runtime-cu12
    Found existing installation: nvidia-cuda-runtime-cu12 12.5.82
    Uninstalling nvidia-cuda-runtime-cu12-12.5.82:
      Successfully uninstalled nvidia-cuda-runtime-cu12-12.5.82
  Attempting uninstall: nvidia-cuda-nvrtc-cu12
    Found existing installation: nvidia-cuda-nvrtc-cu12 12.5.82
    Uninstalling nvidia-cuda-nvrtc-cu12-12.5.82:
      Successfully uninstalled nvidia-cuda-nvrtc-cu12-12.5.82
  Attempting uninstall: nvidia-cuda-cupti-cu12
    Found existing installation: nvidia-cuda-cupti-cu12 12.5.82
    Uninstalling nvidia-cuda-cupti-cu12-12.5.82:
      Successfully uninstalled nvidia-cuda-cupti-cu12-12.5.82
  Attempting uninstall: nvidia-cublas-cu12
```

```
    Found existing installation: nvidia-cublas-cu12 12.5.3.2
    Uninstalling nvidia-cublas-cu12-12.5.3.2:
      Successfully uninstalled nvidia-cublas-cu12-12.5.3.2
  Attempting uninstall: nvidia-cusparse-cu12
    Found existing installation: nvidia-cusparse-cu12 12.5.1.3
    Uninstalling nvidia-cusparse-cu12-12.5.1.3:
      Successfully uninstalled nvidia-cusparse-cu12-12.5.1.3
  Attempting uninstall: nvidia-cudnn-cu12
    Found existing installation: nvidia-cudnn-cu12 9.3.0.75
    Uninstalling nvidia-cudnn-cu12-9.3.0.75:
      Successfully uninstalled nvidia-cudnn-cu12-9.3.0.75
  Attempting uninstall: nvidia-cusolver-cu12
    Found existing installation: nvidia-cusolver-cu12 11.6.3.83
    Uninstalling nvidia-cusolver-cu12-11.6.3.83:
      Successfully uninstalled nvidia-cusolver-cu12-11.6.3.83
Successfully installed datasets-3.3.2 dill-0.3.8 multiprocess-0.70.16
nvidia-cublas-cu12-12.4.5.8 nvidia-cuda-cupti-cu12-12.4.127 nvidia-
cuda-nvrtc-cu12-12.4.127 nvidia-cuda-runtime-cu12-12.4.127 nvidia-
cudnn-cu12-9.1.0.70 nvidia-cufft-cu12-11.2.1.3 nvidia-curand-cu12-
10.3.5.147 nvidia-cusolver-cu12-11.6.1.9 nvidia-cusparse-cu12-
12.3.1.170 nvidia-nvjitlink-cu12-12.4.127 xxhash-3.5.0
```

The "base" model is one of the intermediate sizes offered by OpenAI, falling between the smaller "tiny" and "small" models and the larger "medium" and "large" models.

```
from transformers import
WhisperProcessor,WhisperForConditionalGeneration
import torch
import soundfile as sf
import time

model_size="base" #it could be riny,small,
medium,large,largev2,largev3,
```

WhisperProcessor-its a class,preprocesses audio into a format that whisper model can understand (feature extraction,tokenization)

WhisperForConditionalGeneration-its a class,loads pre-trained whisper model, provides method for audio to text generation

sf-reading and writing audio files in various formats (WAV)

```
processor=WhisperProcessor.from_pretrained(f"openai/whisper-
{model_size}")
#processor is an instance of WhisperProcessor class
#from_pretrained(): This is a static method that downloads and loads a
pre-trained processor from the Hugging Face Model Hub.
```

```
'''
the above line makes a request to the hugging face model hub
it downloads config and vocab files associated with the base model's
processor
these files initiales the object of WhisperProcessor class and knows
how to preprocess audio files for the "base" model
'''

model=WhisperForConditionalGeneration.from_pretrained(f"openai/whisper
-{model_size}")
'''
another request is made to the transformers model hub
it downloads config files(network architecture) and pre-trained
weights that is the learned parameters
pytorch uses these files to build the nn in memory
'''
```

```
/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
secret in your Google Colab and restart your session.
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(
```


```
{"model_id":"777be9760db941a3821c53e84e65dc55","version_major":2,"vers
ion_minor":0}

{"model_id":"259be75a442d4fb093be3315b4aeaa0d","version_major":2,"vers
ion_minor":0}

{"model_id":"396f577979ec4883963f55bbd9031beb","version_major":2,"vers
ion_minor":0}

{"model_id":"75ef4038a6d240a09e4b79aacaf1e902","version_major":2,"vers
ion_minor":0}

{"model_id":"758714bb102b48ce92c98459f5f4add4","version_major":2,"vers
ion_minor":0}

{"model_id":"55e3b9072bfa487aab8c36599b7a6635","version_major":2,"vers
ion_minor":0}

{"model_id":"775e3c512ab5446880d955e3d90306c5","version_major":2,"vers
ion_minor":0}

{"model_id":"f6752a8680ba4992908db7d45c936083","version_major":2,"vers
ion_minor":0}
```

{"model_id":"fd568de0d00b4195b07b26b3214c512f","version_major":2,"version_minor":0}

{"model_id":"fe07400d596c44168642e487757d63a7","version_major":2,"version_minor":0}

{"model_id":"4fd7d43049a24b089fee902c65b99ba4","version_major":2,"version_minor":0}

{"type":"string"}

```
device="cuda" if torch.cuda.is_available() else "cpu"
model.to(device)#loads whisper model into GPU or CPU
#o/p is the detailed view of NN architecture

WhisperForConditionalGeneration(
  (model): WhisperModel(
    (encoder): WhisperEncoder(
      (conv1): Conv1d(80, 512, kernel_size=(3,), stride=(1,),
padding=(1,))
      (conv2): Conv1d(512, 512, kernel_size=(3,), stride=(2,),
padding=(1,))
      (embed_positions): Embedding(1500, 512)
      (layers): ModuleList(
        (0-5): 6 x WhisperEncoderLayer(
          (self_attn): WhisperSdpaAttention(
            (k_proj): Linear(in_features=512, out_features=512,
bias=False)
            (v_proj): Linear(in_features=512, out_features=512,
bias=True)
            (q_proj): Linear(in_features=512, out_features=512,
bias=True)
            (out_proj): Linear(in_features=512, out_features=512,
bias=True)
          )
          (self_attn_layer_norm): LayerNorm((512,), eps=1e-05,
elementwise_affine=True)
          (activation_fn): GELUActivation()
          (fc1): Linear(in_features=512, out_features=2048, bias=True)
          (fc2): Linear(in_features=2048, out_features=512, bias=True)
          (final_layer_norm): LayerNorm((512,), eps=1e-05,
elementwise_affine=True)
        )
      )
      (layer_norm): LayerNorm((512,), eps=1e-05,
elementwise_affine=True)
    )
    (decoder): WhisperDecoder(
      (embed_tokens): Embedding(51865, 512, padding_idx=50257)
      (embed_positions): WhisperPositionalEmbedding(448, 512)
```

```
      (layers): ModuleList(
        (0-5): 6 x WhisperDecoderLayer(
          (self_attn): WhisperSdpaAttention(
            (k_proj): Linear(in_features=512, out_features=512,
bias=False)
            (v_proj): Linear(in_features=512, out_features=512,
bias=True)
            (q_proj): Linear(in_features=512, out_features=512,
bias=True)
            (out_proj): Linear(in_features=512, out_features=512,
bias=True)
          )
          (activation_fn): GELUActivation()
          (self_attn_layer_norm): LayerNorm((512,), eps=1e-05,
elementwise_affine=True)
          (encoder_attn): WhisperSdpaAttention(
            (k_proj): Linear(in_features=512, out_features=512,
bias=False)
            (v_proj): Linear(in_features=512, out_features=512,
bias=True)
            (q_proj): Linear(in_features=512, out_features=512,
bias=True)
            (out_proj): Linear(in_features=512, out_features=512,
bias=True)
          )
          (encoder_attn_layer_norm): LayerNorm((512,), eps=1e-05,
elementwise_affine=True)
          (fc1): Linear(in_features=512, out_features=2048, bias=True)
          (fc2): Linear(in_features=2048, out_features=512, bias=True)
          (final_layer_norm): LayerNorm((512,), eps=1e-05,
elementwise_affine=True)
        )
      )
      (layer_norm): LayerNorm((512,), eps=1e-05,
elementwise_affine=True)
    )
  )
  (proj_out): Linear(in_features=512, out_features=51865, bias=False)
)
```

preparing the audio data

```
from datasets import load_dataset
df=load_dataset("hf-internal-testing/librispeech_asr_dummy","clean",sp
lit="validation")#clean config is free of noise
sample_audio=df[0]["audio"]["array"]#audio data from the first item in
the dataset
sample_rate=df[0]["audio"]["sampling_rate"]
```

{"model_id":"0361530dff1a4b38877ae815967628d4","version_major":2,"version_minor":0}

{"model_id":"7adff4d236564df19ebf33f1a8806033","version_major":2,"version_minor":0}

{"model_id":"98c5f60d1e0e41c79085c3c19624a663","version_major":2,"version_minor":0}

The audio data in the Hugging Face datasets library is initially represented as raw numerical samples in a NumPy array.

We convert it to WAV format to make it compatible with external tools and for easy playback and inspection.

```
sf.write("test.wav", sample_audio, sample_rate)#saving the audio as wav file
```

test.wav - name of the audio file

```
!pip install psutil
```

```
Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (5.9.5)
```

```
import psutil
import os
import resource

def transcribe (audio,processor,model,device):

input_features=processor(audio,sampling_rate=sample_rate,return_tensors="pt").input_features.to(device)#returning it in pytroch since whisper is a pytorch model , to(device) moves the input feature to cpu or gpu
  start_time=time.time()
  predicted_ids=model.generate(input_features)#generates predicted token IDs using whisper model

transcription=processor.batch_decode(predicted_ids,skip_special_tokens=True)#decodes predicted tokens to text , ignores the padding tokens
  end_time=time.time()
  latency=end_time-start_time
  return transcription,latency
def get_cpu_mem_use():
  process=psutil.Process(os.getpid())
  mem_use_byte=process.memory_info().rss
  mem_use_mb=mem_use_byte/(1024*1024)
  return mem_use_mb
loop=10
```

```
latencies=[]

if device =="cuda":
  torch.cuda.reset_peak_memory_stats(device=device)
  start_memory=torch.cuda.memory_allocated(device=device)

  for _ in range(loop):

transcription,latency=transcribe(sample_audio,processor,model,device)
    latencies.append(latency)

  end_memory=torch.cuda.memory_allocated(device=device)
  peak_memory=torch.cuda.max_memory_allocated(device=device)
  memory_usage=end_memory-start_memory
  avg_latency=sum(latencies)/loop

  print(f"GPU Transcription:{transcription[0]}")
  print(f"GPU Average Latency:{avg_latency:.4f} seconds")
  print(f"GPU Memory Usage:{memory_usage/(1024*1024):.2f}MB")
  print(f"GPU Peak Memory Usage:{peak_memory/(1024*1024):.2f}MB")

else:
  start_memory=get_cpu_mem_use()

  for _ in range(loop):

transcription,latency=transcribe(sample_audio,processor,model,device)
    latencies.append(latency)
  end_memory=get_cpu_mem_use()
  memory_usage=end_memory-start_memory

  avg_latency=sum(latencies)/loop
  print(f"CPU Transcription:{transcription[0]}")
  print(f"CPU Average Latency:{avg_latency:.4f} seconds")
  print(f"CPU Memory Usage: {memory_usage:.2f} MB")



CPU Transcription: Mr. Quilter is the apostle of the middle classes,
and we are glad to welcome his gospel.
CPU Average Latency:5.9084 seconds
CPU Memory Usage: 0.46 MB

import torch

if torch.cuda.is_available():
    print("GPU is available!")
    print(f"Device name: {torch.cuda.get_device_name(0)}")
else:
    print("GPU is NOT available.")
```

```
GPU is NOT available.
```