```python
1 from google.colab import drive
2 drive.mount('/content/drive')
3
```

⤓ Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

```python
1 import torch
2 print(torch.__version__)
3 print(torch.cuda.is_available())
4
```

⤓ 2.7.1+cu118
True

```python
1 !pip install --upgrade torch torchvision torchaudio
2 !pip install --upgrade diffusers transformers accelerate scipy
3 !pip install speechrecognition pydub
4
```

⤓ Requirement already satisfied: nvidia-cusolver-cu11==11.4.1.48 in /usr/local/lib/python3.11/dist-packages (from torch) (11.4.1.48
Requirement already satisfied: nvidia-cusparse-cu11==11.7.5.86 in /usr/local/lib/python3.11/dist-packages (from torch) (11.7.5.86
Requirement already satisfied: nvidia-nccl-cu11==2.21.5 in /usr/local/lib/python3.11/dist-packages (from torch) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu11==11.8.86 in /usr/local/lib/python3.11/dist-packages (from torch) (11.8.86)
Requirement already satisfied: triton==3.3.1 in /usr/local/lib/python3.11/dist-packages (from torch) (3.3.1)
Requirement already satisfied: setuptools>=40.8.0 in /usr/local/lib/python3.11/dist-packages (from triton==3.3.1->torch) (75.2.0)
Requirement already satisfied: numpy in /usr/local/lib/python3.11/dist-packages (from torchvision) (2.0.2)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in /usr/local/lib/python3.11/dist-packages (from torchvision) (11.2.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from sympy>=1.13.3->torch) (1.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->torch) (3.0.2)
Requirement already satisfied: diffusers in /usr/local/lib/python3.11/dist-packages (0.33.1)
Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-packages (4.52.4)
Requirement already satisfied: accelerate in /usr/local/lib/python3.11/dist-packages (1.7.0)
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (1.15.3)
Requirement already satisfied: importlib-metadata in /usr/local/lib/python3.11/dist-packages (from diffusers) (8.7.0)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from diffusers) (3.18.0)
Requirement already satisfied: huggingface-hub>=0.27.0 in /usr/local/lib/python3.11/dist-packages (from diffusers) (0.32.4)
Requirement already satisfied: numpy in /usr/local/lib/python3.11/dist-packages (from diffusers) (2.0.2)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-packages (from diffusers) (2024.11.6)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from diffusers) (2.32.3)
Requirement already satisfied: safetensors>=0.3.1 in /usr/local/lib/python3.11/dist-packages (from diffusers) (0.5.3)
Requirement already satisfied: Pillow in /usr/local/lib/python3.11/dist-packages (from diffusers) (11.2.1)
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: tokenizers<0.22,>=0.21 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.21.1)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.11/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (from accelerate) (5.9.5)
Requirement already satisfied: torch>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from accelerate) (2.7.1+cu118)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.27.0->diffuse
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.27.
Requirement already satisfied: hf-xet<2.0.0,>=1.1.2 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.27.0->dif
Requirement already satisfied: sympy>=1.13.3 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (1.13.3)
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.5)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.1.6)
Requirement already satisfied: nvidia-cuda-nvrtc-cu11==11.8.89 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->acc
Requirement already satisfied: nvidia-cuda-runtime-cu11==11.8.89 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->a
Requirement already satisfied: nvidia-cuda-cupti-cu11==11.8.87 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->acc
Requirement already satisfied: nvidia-cudnn-cu11==9.1.0.70 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->acceler
Requirement already satisfied: nvidia-cublas-cu11==11.11.3.6 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accel
Requirement already satisfied: nvidia-cufft-cu11==10.9.0.58 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accele
Requirement already satisfied: nvidia-curand-cu11==10.3.0.86 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accel
Requirement already satisfied: nvidia-cusolver-cu11==11.4.1.48 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->acc
Requirement already satisfied: nvidia-cusparse-cu11==11.7.5.86 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->acc
Requirement already satisfied: nvidia-nccl-cu11==2.21.5 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate
Requirement already satisfied: nvidia-nvtx-cu11==11.8.86 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerat
Requirement already satisfied: triton==3.3.1 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.3.1)
Requirement already satisfied: setuptools>=40.8.0 in /usr/local/lib/python3.11/dist-packages (from triton==3.3.1->torch>=2.0.0->a
Requirement already satisfied: zipp>=3.20 in /usr/local/lib/python3.11/dist-packages (from importlib-metadata->diffusers) (3.22.0
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->diffusers) (3.
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->diffusers) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->diffusers) (2.4.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->diffusers) (2025.4.2
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from sympy>=1.13.3->torch>=2.0.0->a
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->torch>=2.0.0->accelerate)
Requirement already satisfied: speechrecognition in /usr/local/lib/python3.11/dist-packages (3.14.3)
Requirement already satisfied: pydub in /usr/local/lib/python3.11/dist-packages (0.25.1)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.11/dist-packages (from speechrecognition) (4.14.0)

```python
1 !apt install ffmpeg
2
```

⤓ Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
ffmpeg is already the newest version (7:4.4.2-0ubuntu0.22.04.1).

```
  0 upgraded, 0 newly installed, 0 to remove and 35 not upgraded.


 1 from google.colab import files
 2 uploaded = files.upload()

 3

 4 from pydub import AudioSegment

 5

 6 # This works if FFmpeg is correctly installed
 7 audio = AudioSegment.from_file("1.opus")  # removed format="opus"
 8 audio.export("converted.wav", format="wav")

 9

10 import speech_recognition as sr

11

12 recognizer = sr.Recognizer()
13 with sr.AudioFile("converted.wav") as source:
14     audio_data = recognizer.record(source)

15

16 try:
17     text = recognizer.recognize_google(audio_data)
18     print("🎤 Recognized Text:", text)
19 except sr.UnknownValueError:
20     print("Could not understand the audio.")
21     text = None

22

23 from diffusers import StableDiffusionPipeline
24 import torch

25

26 pipe = StableDiffusionPipeline.from_pretrained(
27     "runwayml/stable-diffusion-v1-5", torch_dtype=torch.float16
28 ).to("cuda")

29

30 if text:
31     image = pipe(text).images[0]
32     image.save("output.png")
33     display(image)

34
```

Choose Files 1.opus
  • **1.opus**(audio/ogg) - 3977 bytes, last modified: 10/6/2025 - 100% done
  Saving 1.opus to 1 (3).opus
  🎤 Recognized Text: rose flowers

Loading pipeline components...: 100%   7/7 [00:19<00:00, 3.10s/it]

100%   50/50 [00:07<00:00, 6.86it/s]