Solutions for MeloTTS and Open-ai Whisper models, on web discussion for errors.

Run both models: Solution Using Method-3.

melotts: https://huggingface.co/myshell-ai/MeloTTS-Japanese

Whisper Turbo: https://huggingface.co/openai/whisper-large-v3-turbo

using Rocm as a framework on the device provided, you are able to run on Radeon GPU.

References : Rocm docs

https://rocm.docs.amd.com/projects/install-on-linux/en/latest/install/3rd-party/pytorch-install.html

🔧 Environment Setup

- 👤 **Author**: `lilly@PC-AMD` ⇒ To check my environment "**wsl -d Ubuntu-22.04**" in Windows powershell
- 🖥️ **Host OS**: Windows 11
- 🐧 **WSL2 Distro**: Ubuntu 22.04
- 🎮 **GPU**: AMD Radeon RX 9070 (Confirmed functional via ROCm & PyTorch)
- ⚙️ **ROCm Version**: 6.4.1
- 🧠 **PyTorch Version**: 2.6.0 (Nightly ROCm Build with HIP 6.4.x)
- 🐍 **Python Version**: 3.10.18
- 📦 **Frameworks & Tools**:
  - Hugging Face Transformers (`transformers`)
  - SoundFile (`soundfile`)

    - TorchAudio (`torchaudio`)
    - ROCm Toolkit (`rocminfo`, `hip-runtime`)
- 🧪 **Model Inference Tools**:
    - ✅ `myshell-ai/MeloTTS-Japanese` (Japanese TTS model — tested and exported `.wav` file)
    - ✅ `openai/whisper-large-v3-turbo` (ASR model — successful speech-to-text inference)
- 🧬 **Conda Environments**:
    - `rocmtest` – ROCm/PyTorch validation and setup
    - `melotts` – MeloTTS text-to-speech pipeline
    - `whisperturbo` – Whisper Turbo transcription

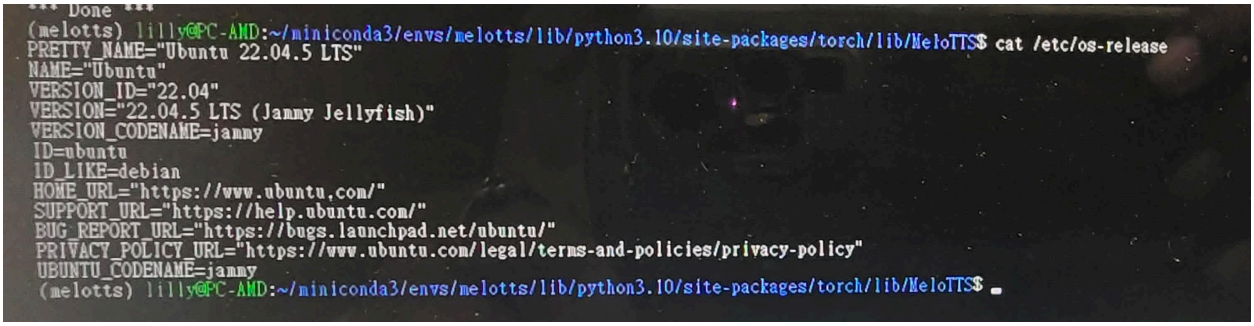
✅ Task Progress Summary

Step	Task	Status
1️⃣	Set up WSL2 (Ubuntu 22.04)	✅ Done
2️⃣	Verified AMD Radeon RX 9070 GPU	✅ Done
3️⃣	Installed ROCm 6.4.1 via `.deb` package	✅ Done
4️⃣	Installed PyTorch ROCm wheels in Conda (`rocmtest`)	✅ Done
5️⃣	Verified `rocminfo` detects CPU and GPU agents	✅ Done
6️⃣	Installed Python 3.10 + Conda env + ROCm PyTorch (2.6.0)	✅ Done
7️⃣	Verified `torch.cuda.is_available()` → `True`	✅ Done
8️⃣	Created `melotts` Conda environment, cloned MeloTTS from GitHub, and successfully ran inference on ROCm-enabled Radeon GPU (via `CUDA`)	✅ Done
9️⃣	Saved generated audio output (`melotts_output.wav`) successfully (some non-blocking warnings observed)	✅ Done
🔟	Created `whisper-turbo` Conda environment, installed dependencies, and successfully ran **Whisper Large v3**	✅ Done

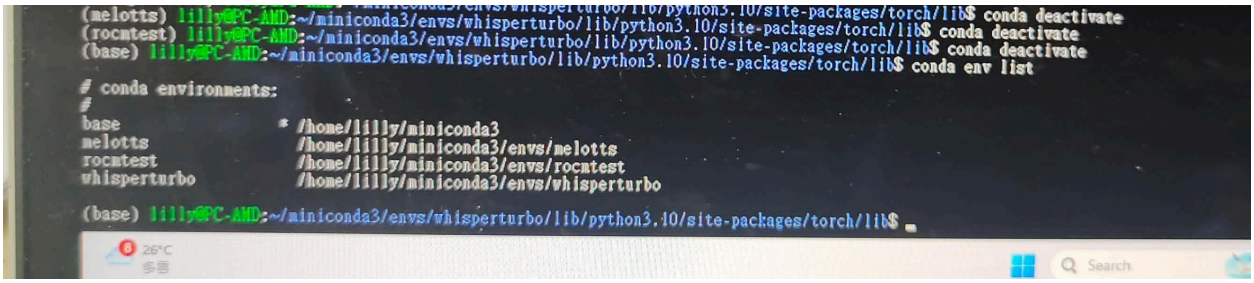
Step	Task	Status
	Turbo model on ROCm GPU with valid audio transcription	

Outputs

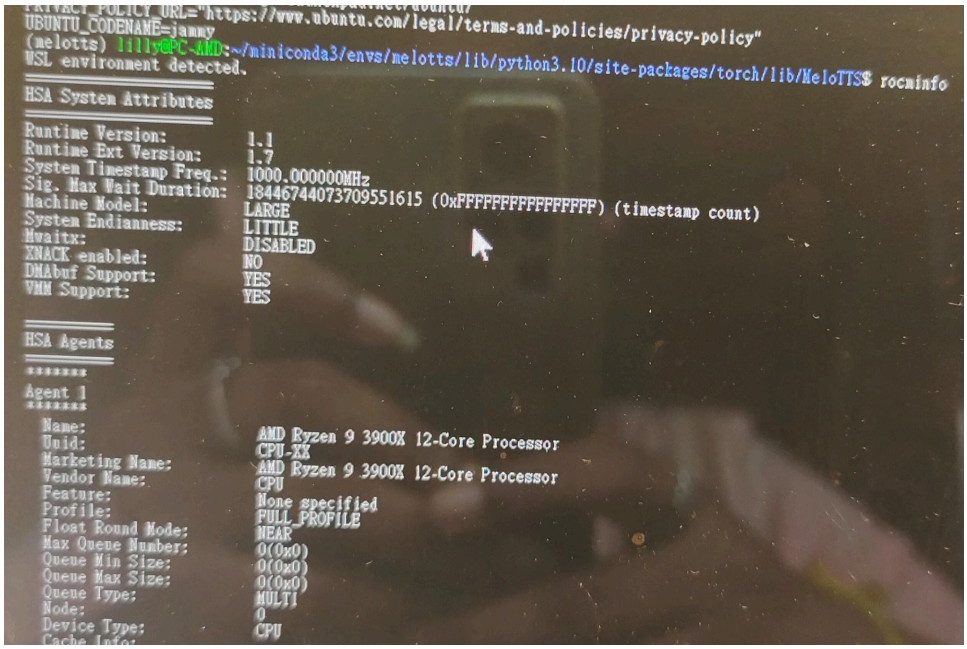
1. Set up WSL2 (Ubuntu 22.04)

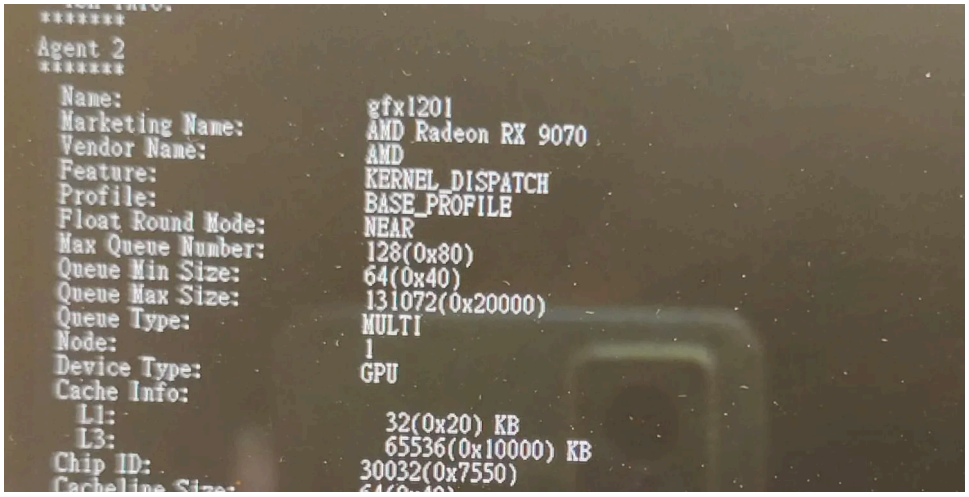


conda environments :

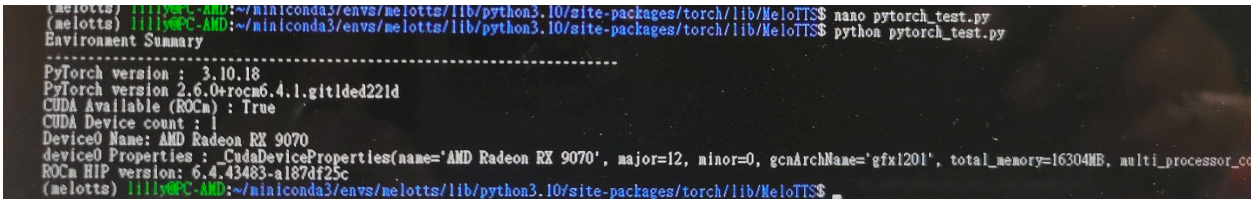


2. Verfied rocminfo detects CPU and GPU agents

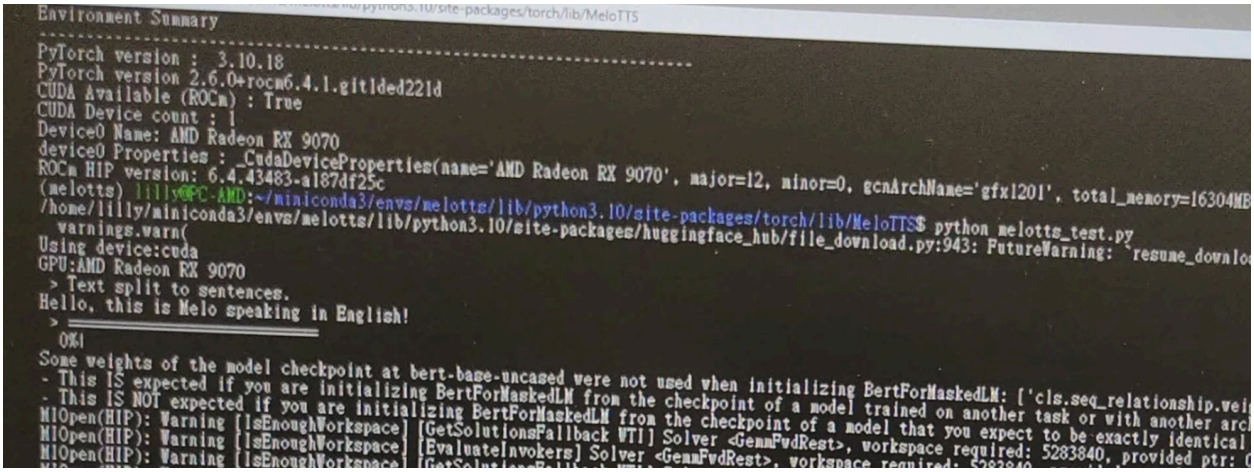


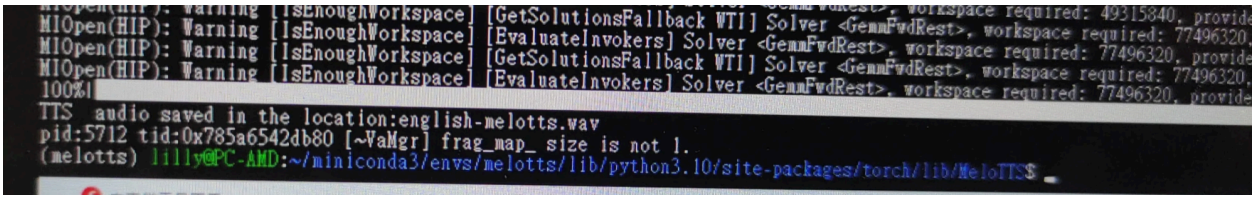


3. Verified `torch.cuda.is_available()` → `True`

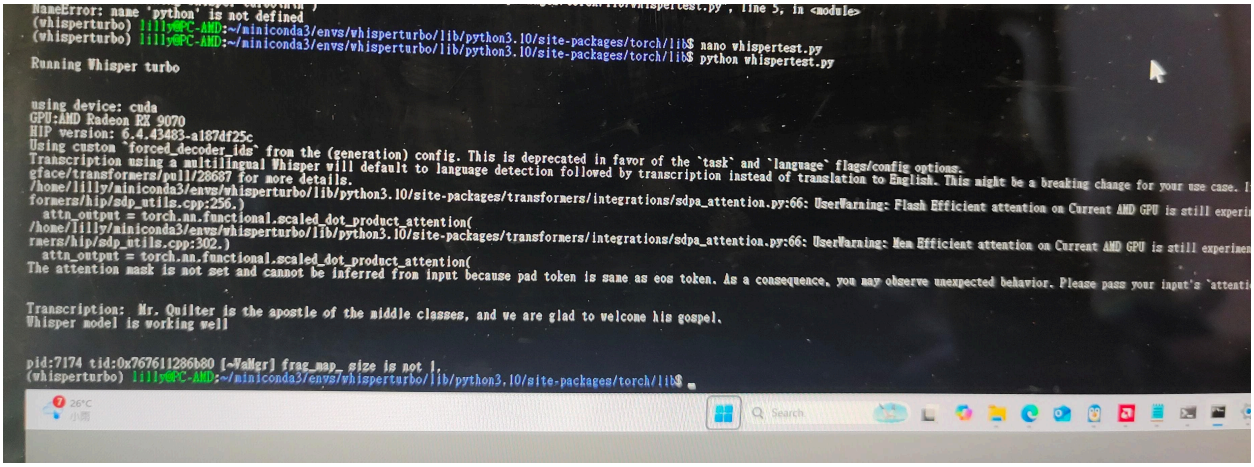


4. Successfully ran MeloTTS model on CUDA (Radeon GPU).





5. successfully ran **Whisper Large v3 Turbo** model on ROCm GPU with valid audio transcription



Wonders.ai company test