Indic ASR Setup and Execution Guide (Ubuntu with GPU)

This guide details the steps to set up the environment and run inference using AI4Bharat's IndicWhisper and IndicConformer models on an Ubuntu system with an NVIDIA GPU.

Prerequisites

- Ubuntu OS
- NVIDIA GPU with drivers installed (`nvidia-smi` should work)
- Python 3.8 or higher
- `git`, `ffmpeg`, `venv`

Jetson Orin Setup (Special Instructions)

Important: Jetson devices (Orin, Xavier, Nano) use the ARM64 architecture and require a specific version of PyTorch provided by NVIDIA JetPack. You **cannot** install PyTorch using the standard `pip install torch` command from PyPI.

1. Verify JetPack Version

Check your JetPack version:

```
sudo apt-cache show nvidia-jetpack
```

2. Install PyTorch for Jetson

Follow the official NVIDIA guide to install the PyTorch wheel compatible with your JetPack version:
NVIDIA Jetson PyTorch Installation Guide

Example for JetPack 6.0 (CUDA 12.6):

```
# Install system dependencies
sudo apt-get install -y python3-pip libopenblas-base libopenmpi-dev libomp-dev

# Install PyTorch (check the link above for the exact URL for your JetPack
version)
# Example URL (verify this!):
wget https://developer.download.nvidia.com/compute/redist/jp/v60/pytorch/torch-
2.4.0a0+6dd6c25.nv24.07-cp310-cp310-linux_aarch64.whl
pip install torch-2.4.0a0+6dd6c25.nv24.07-cp310-cp310-linux_aarch64.whl

# Install torchvision (must be compiled from source)
sudo apt-get install -y libjpeg-dev zlib1g-dev libpython3-dev libavcodec-dev
libavformat-dev libswscale-dev
```

```
git clone --branch v0.19.0 https://github.com/pytorch/vision torchvision
cd torchvision
export BUILD_VERSION=0.19.0
python3 setup.py install --user
cd ..
```

3. Install Other Dependencies

After installing PyTorch and torchvision manually:

```
pip install -r requirements.txt
pip install Cython packaging
```

4. Install NeMo

```
git clone https://github.com/AI4Bharat/NeMo.git
cd NeMo
bash reinstall.sh
cd ..
```

Standard Ubuntu Setup (x86_64 only)

Use this section ONLY if you are on a standard desktop/server Ubuntu (NOT Jetson).

1. **Make the setup script executable:**

   ```
   chmod +x setup_ubuntu.
   ```

2. **Run the setup script:**
   This script will create a virtual environment, install dependencies, install NeMo, and download the IndicWhisper model.

   ```
   ./setup_ubuntu.sh
   ```

3. **Activate the environment:**

   ```
   source venv/bin/activate
   ```

Manual Setup Steps

If you prefer to run steps manually:

1. **Install System Dependencies:**

```
sudo apt-get update && sudo apt-get install -y ffmpeg libsndfile1 git
python3-venv
```

2. **Create and Activate Virtual Environment:**

```
python3 -m venv venv
source venv/bin/activate
```

3. **Install PyTorch (with CUDA support):**

```
pip install torch torchvision torchaudio --index-url
https://download.pytorch.org/whl/cu118
# Note: Adjust 'cu118' based on your CUDA version (check with nvidia-smi)
```

4. **Install Python Dependencies:**

```
pip install -r requirements.txt
pip install Cython packaging
```

5. **Install NeMo (Required for IndicConformer):**

```
git clone https://github.com/AI4Bharat/NeMo.git
cd NeMo
bash reinstall.sh
cd ..
```

6. **Download IndicWhisper Model:**

```
python download_indic_whisper_model.py
```

Running Inference

Ensure your virtual environment is active (`source venv/bin/activate`) and you have an audio file (e.g., `audio.wav`) in the directory.

1. Test IndicWhisper

```
python test_indic_whisper.py <path_to_audio_file>
```

Example: `python test_indic_whisper.py "hindi podcast.wav"`

2. Test IndicConformer

```
python test_indic_conformer.py <path_to_audio_file>
```

Example: `python test_indic_conformer.py "hindi podcast.wav"`

Troubleshooting

- **CUDA/GPU Issues:** Run `python -c "import torch; print(torch.cuda.is_available())"` to verify GPU access. If `False`, reinstall PyTorch with the correct CUDA version.
- **NeMo Import Errors:** Ensure you ran `bash reinstall.sh` inside the `NeMo` directory and that the installation completed successfully.
- **Audio Format:** `ffmpeg` is required to handle various audio formats. Ensure it is installed.