GPU Usage Manual – NVIDIA RTX with CUDA

Guide for checking, monitoring, and using the NVIDIA RTX GPU with CUDA acceleration on your Debian 12 workstation.

1. Verifying CUDA installation

- Open MATE Terminal.

- Check CUDA toolkit version: `nvcc --version`

- Verify driver installation: `nvidia-smi`

2. Testing GPU access in Python

- Activate your Python venv: `source ~/Desktop/user-venv/bin/activate`

- Run: `python -c "import torch; print(torch.cuda.is_available())"`

- Expected output: `True` indicates the GPU is available to PyTorch.

3. Monitoring GPU usage

- Run: `watch -n 1 nvidia-smi`

- Displays GPU utilization, memory usage, running processes, and temperature in real time.

4. Running GPU-accelerated workloads

- Use ML/DL frameworks such as PyTorch or TensorFlow inside the venv.

- Ensure that `device='cuda'` (PyTorch) or `/GPU:0` (TensorFlow) is set in your scripts.

5. Best practices

- Avoid running multiple heavy GPU workloads at once unless intended.

- Monitor GPU temperature — keep below $85\,°\mathrm{C}$ for stability.

- If CUDA is not detected, contact support before reinstalling drivers.

6. Troubleshooting

- If `nvidia-smi` reports *No devices were found*, restart the system.

- For persistent issues, verify correct driver version and CUDA compatibility.

- Do not update the NVIDIA driver or kernel without approval.

7. Support

- Provide `nvidia-smi` output, CUDA version, and error messages when contacting support.