

Ollama

Ollama

- 1. Large Language Model (LLM)
- 2. Ollama Installation
 - Script Installation
- 3. Use Ollama
- 4. Ollama Uninstall
- References

Demo Environment

Development board: Jetson Orin series motherboard

SSD: 128G

Tutorial Scope

Motherboard Model	Supported
Jetson Orin NX 16GB	√
Jetson Orin NX 8GB	√
Jetson Orin Nano 8GB	√
Jetson Orin Nano 4GB	√

Ollama is an open source tool that aims to simplify the deployment and operation of large language models, allowing users to use high-quality language models in local environments.

1. Large Language Model (LLM)

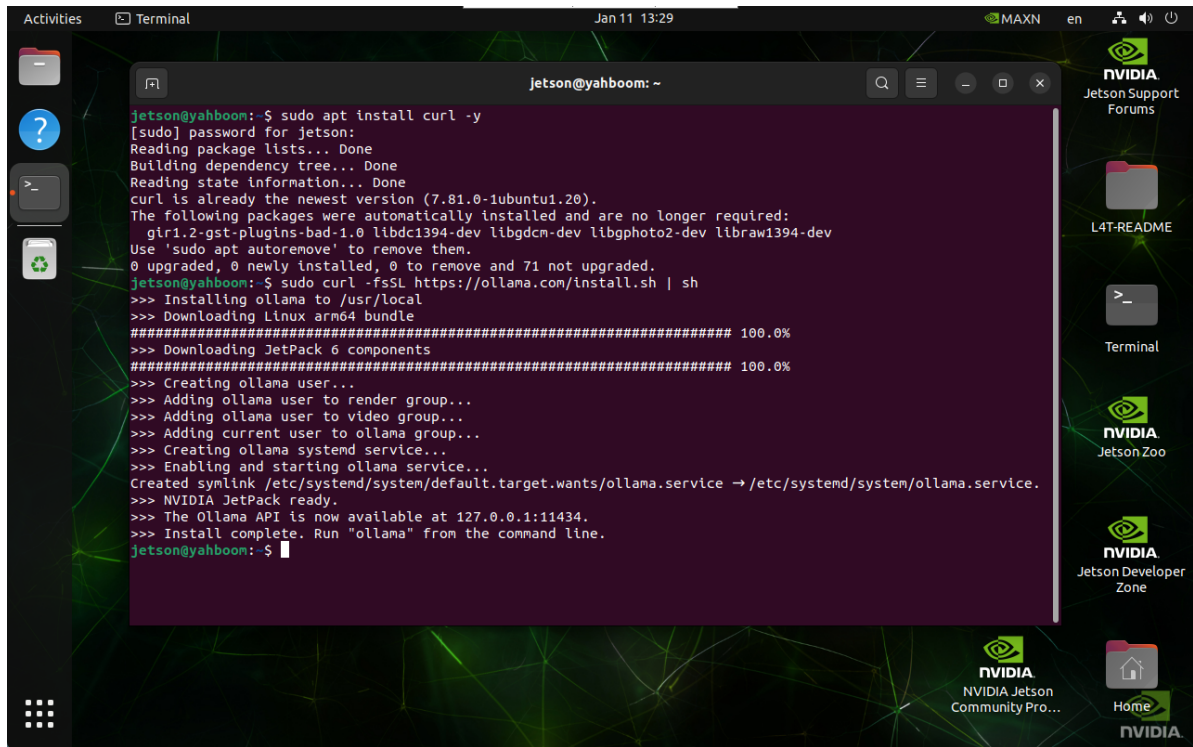
Large Language Models (LLM) are a type of advanced text generation system based on artificial intelligence technology. Its main feature is that it can learn and understand human language through large-scale training data and can generate natural and fluent text.

2. Ollama Installation

The tutorial demonstrates the use of scripts to install Ollama on the Jetson Orin series motherboard.

Script Installation

```
sudo apt install curl -y
sudo curl -fsSL https://ollama.com/install.sh | sh
```



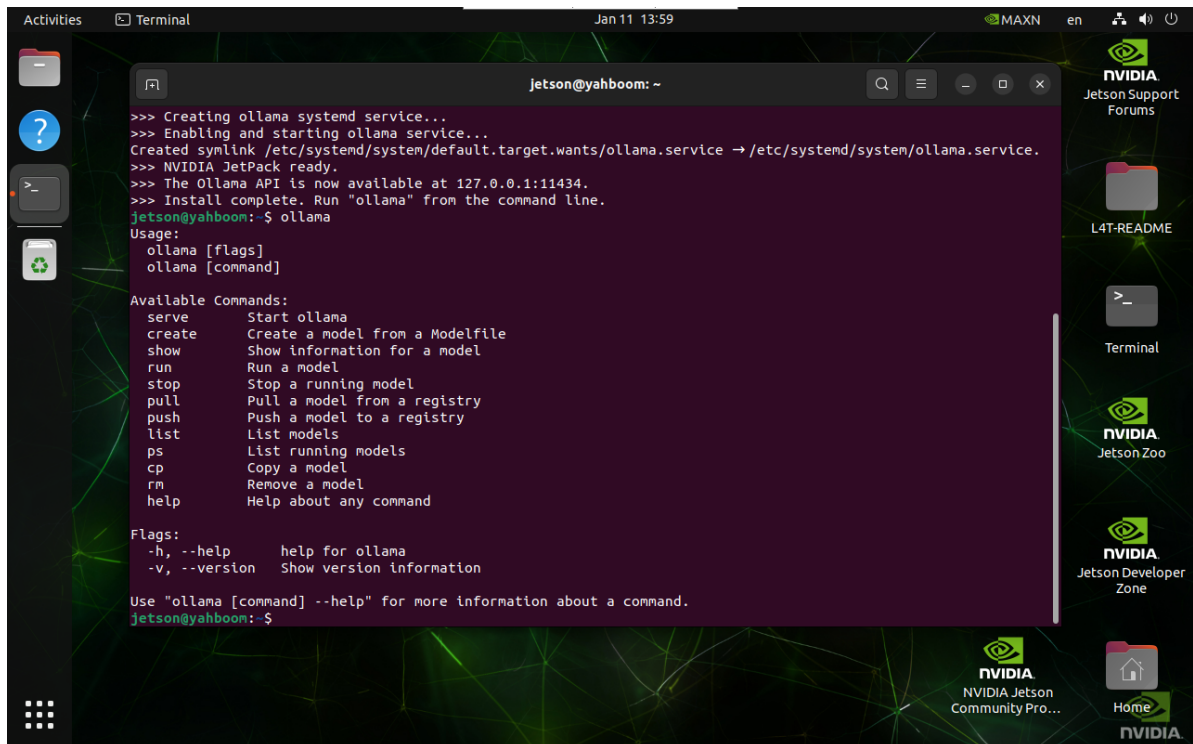
The image shows a terminal window on a Jetson device. The user is running the command `sudo apt install curl -y`. The terminal output shows that curl is already installed. Then, the user runs `sudo curl -fsSL https://ollama.com/install.sh | sh`. The terminal output shows the installation progress, including downloading the Linux arm64 bundle and JetPack 6 components. The installation is complete, and the Ollama API is now available at 127.0.0.1:11434.

```
jetson@yahboom:~$ sudo apt install curl -y
[sudo] password for jetson:
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
curl is already the newest version (7.81.0-1ubuntu1.20).
The following packages were automatically installed and are no longer required:
  gir1.2-gst-plugins-bad-1.0 libdc1394-dev libgdcn-dev libgphoto2-dev libraw1394-dev
Use 'sudo apt autoremove' to remove them.
0 upgraded, 0 newly installed, 0 to remove and 71 not upgraded.
jetson@yahboom:~$ sudo curl -fsSL https://ollama.com/install.sh | sh
>>> Installing ollama to /usr/local
>>> Downloading linux arm64 bundle
##### 100.0%
>>> Downloading JetPack 6 components
##### 100.0%
>>> Creating ollama user...
>>> Adding ollama user to render group...
>>> Adding ollama user to video group...
>>> Adding current user to ollama group...
>>> Creating ollama systemd service...
>>> Enabling and starting ollama service...
Created symlink /etc/systemd/system/default.target.wants/ollama.service → /etc/systemd/system/ollama.service.
>>> NVIDIA JetPack ready.
>>> The Ollama API is now available at 127.0.0.1:11434.
>>> Install complete. Run "ollama" from the command line.
jetson@yahboom:~$
```

The entire installation process takes a long time, please wait patiently!

3. Use Ollama

Type ollama in the terminal and you will see the prompt:



The image shows a terminal window on a Jetson device. The user has run the `ollama` command, which displays the usage information and available commands. The available commands are listed in a table format. The flags are also listed. The user is prompted to use `ollama [command] --help` for more information about a command.

```
jetson@yahboom:~$ ollama
Usage:
  ollama [flags]
  ollama [command]

Available Commands:
  serve    Start ollama
  create   Create a model from a Modelfile
  show     Show information for a model
  run      Run a model
  stop     Stop a running model
  pull     Pull a model from a registry
  push     Push a model to a registry
  list     List models
  ps       List running models
  cp       Copy a model
  rm       Remove a model
  help     Help about any command

Flags:
  -h, --help      help for ollama
  -v, --version    Show version information

Use "ollama [command] --help" for more information about a command.
jetson@yahboom:~$
```

Command	Purpose
ollama serve	Start ollama
ollama create	Create a model from a model file
ollama show	Show model information
ollama run	Run a model
ollama pull	Pull a model from a registry
ollama push	Push a model to a registry
ollama list	List models
ollama ps	List running models
ollama cp	Copy a model
ollama rm	Delete a model
ollama help	Get help information about any command

4. Ollama Uninstall

- Deleting a service

```
sudo systemctl stop ollama
sudo systemctl disable ollama
sudo rm /etc/systemd/system/ollama.service
```

- Deleting files

```
sudo rm $(which ollama)
```

- Deleting model and service users and groups

```
sudo rm -r /usr/share/ollama
sudo userdel ollama
sudo groupdel ollama
```

References

Ollama

Official website: <https://ollama.com/>

GitHub: <https://github.com/ollama/ollama>