

# 安裝說明 PyTorch /Cuda/ cuDnn install SOP

## 深度學習準備前置作業 工具如下

- Pytorch 套件
- CUDA 是 Nvidia 專用於平行化運算的框架。
- cuDNN (CUDA Deep Neural Network Library) 是CUDA 深度學習的函式庫。

### 小叮嚀:

- 30xx, 40xx 要安裝 CU11 以上的版本，否則無法執行
- 在一個獨立的python 環境中，能用 pip 安裝就用 pip，除非找不到適合的系統版本安裝包在使用 conda 指令安裝

CUDA 是否要安裝?

答:

如果電腦只有CPU，單純練習語法也是可以使用 pytorch，不一定要安裝 CUDA，但是如果部屬需求，例如導出TensorRT模型，則需要安裝CUDA

# 確認GPU驅動是否要更新

## 法一: 指令確認

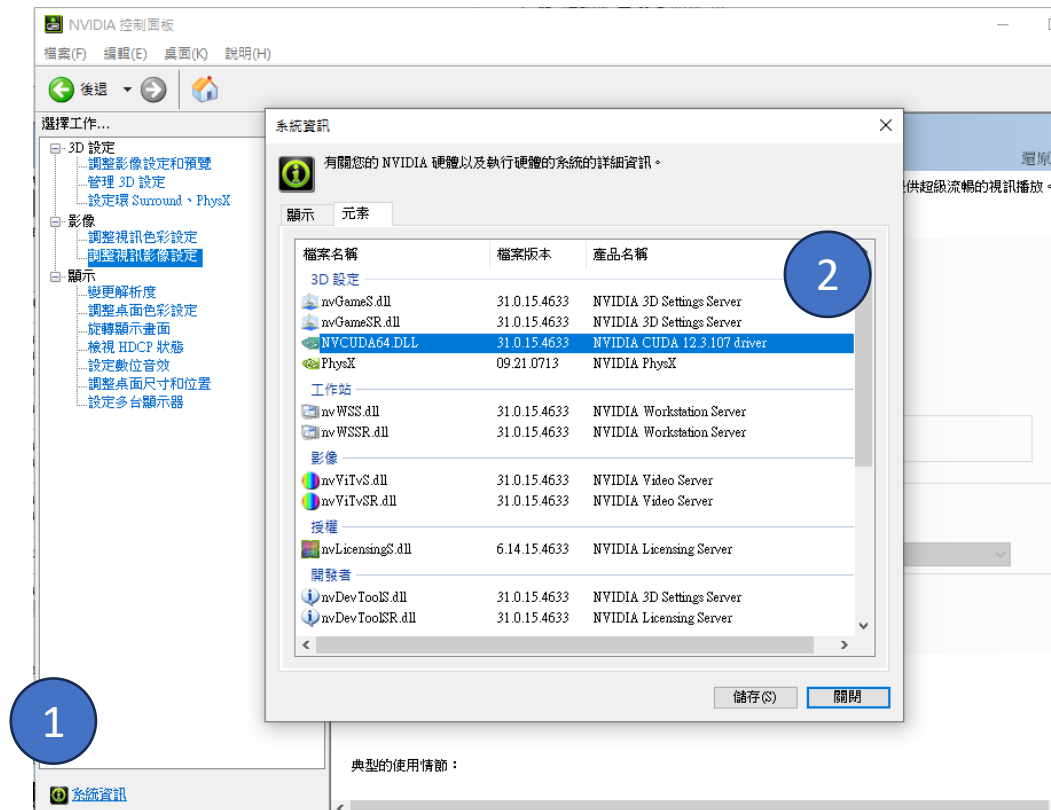
```
user@DESKTOP-QA3SSAC MINGW64 ~
$ nvidia-smi
Sun Jan 14 08:59:55 2024

+-----+
| NVIDIA-SMI 546.33                | Driver Version: 546.33      | CUDA Version: 12.3   |
+-----+-----+
| GPU Name                               | TCC/WDDM                    | Bus-Id                | Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf              Pwr:Usage/Cap |      Memory-Usage            | GPU-Util    Compute M. |
|                                       |                              |                    |    MIG M. |
+-----+-----+
| 0  NVIDIA GeForce RTX 3090          | WDDM                        | 00000000:01:00.0 On  |         | 0%                   |
| 0%   33C   P8              28W / 390W    | 624MiB / 24576MiB          |              Default   |         |
+-----+-----+

Processes:
+-----+
| GPU  GI  CI  PID  Type  Process name                        | GPU Memory |
| ID   ID   ID             |              | Usage        |
+-----+-----+
| 0    N/A N/A   984   C+G   ...n\120.0.2210.121\msedgewebview2.exe | N/A        |
| 0    N/A N/A  2624   C+G   ...5n1h2txyewy\ShellExperienceHost.exe | N/A        |
| 0    N/A N/A  3440   C+G   ...Search_cw5n1h2txyewy\SearchApp.exe  | N/A        |
| 0    N/A N/A  8708   C+G   ...yewy\Microsoft.AAD.BrokerPlugin.exe  | N/A        |
| 0    N/A N/A  9300   C+G   ...les\Microsoft.OneDrive\OneDrive.exe  | N/A        |
| 0    N/A N/A  10352  C+G   ...GeForce Experience\NVIDIA Share.exe   | N/A        |
| 0    N/A N/A  10756  C+G   C:\Windows\explorer.exe                  | N/A        |
| 0    N/A N/A  10968  C+G   ...CBS_cw5n1h2txyewy\TextInputHost.exe  | N/A        |
| 0    N/A N/A  13336  C+G   ...oogle\Chrome\Application\chrome.exe  | N/A        |
| 0    N/A N/A  14876  C+G   ...2txyewy\StartMenuExperienceHost.exe  | N/A        |
| 0    N/A N/A  23632  C+G   ...4__8wekyb3d8bbwe\bin\PBIDesktop.exe  | N/A        |
| 0    N/A N/A  27104  C+G   ...siveControlPanel\SystemSettings.exe   | N/A        |
| 0    N/A N/A  30444  C+G   ...ience\NVIDIA GeForce Experience.exe  | N/A        |
+-----+-----+
```

Driver version: 546.33  
CUDA version: 12.3 (最高支援)

## 法二: GUI 確認



1. 以 RTX 3090 為例
2. 打開CMD, 輸入 **nvidia-smi** 確認 GPU 驅動程式以及 CuDA 版本

在終端機輸入 **nvidia-smi** 指令，會跳出目前使用的顯示卡以及Cuda Driver API 的版本

# 更新GPU驅動

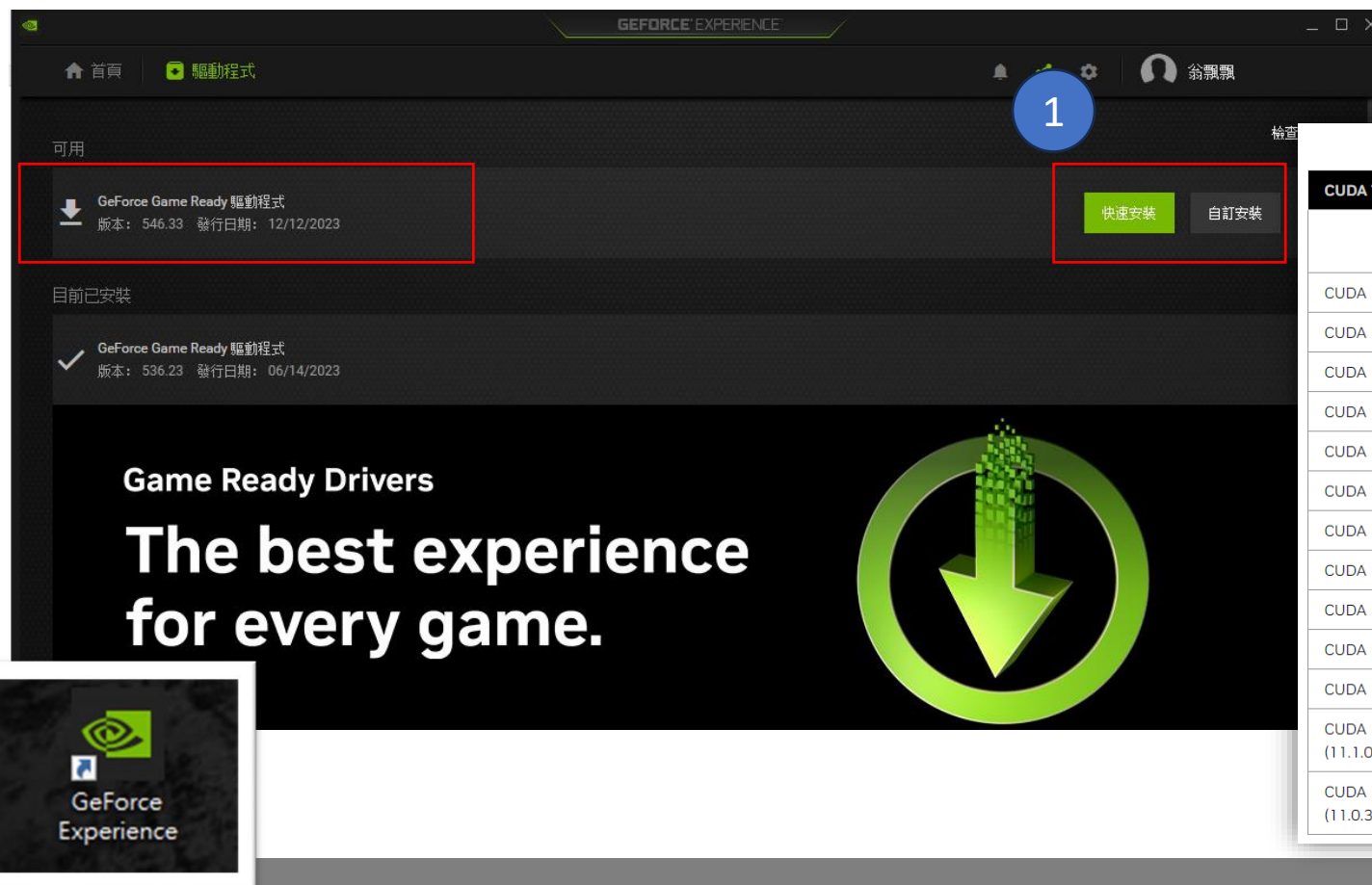
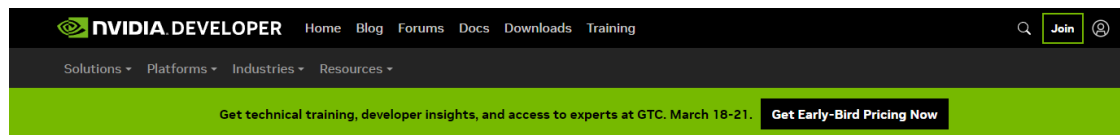


Table 2: CUDA Toolkit and Minimum Required Driver Version for CUDA Minor Version Compatibility

CUDA Toolkit	Minimum Required Driver Version for CUDA Minor Version Compatibility*	
	Linux x86_64 Driver Version	Windows x86_64 Driver Version
CUDA 12.3.x	>=525.60.13	>=527.41
CUDA 12.2.x	>=525.60.13	>=527.41
CUDA 12.1.x	>=525.60.13	>=527.41
CUDA 12.0.x	>=525.60.13	>=527.41
CUDA 11.8.x	>=450.80.02	>=452.39
CUDA 11.7.x	>=450.80.02	>=452.39
CUDA 11.6.x	>=450.80.02	>=452.39
CUDA 11.5.x	>=450.80.02	>=452.39
CUDA 11.4.x	>=450.80.02	>=452.39
CUDA 11.3.x	>=450.80.02	>=452.39
CUDA 11.2.x	>=450.80.02	>=452.39
CUDA 11.1 (11.1.0)	>=450.80.02	>=452.39
CUDA 11.0 (11.0.3)	>=450.36.06**	>=451.22**

1. NVIDIA CUDA Toolkit Release
2. 到 <https://docs.nvidia.com/cuda/cuda-toolkit-release-notes/index.html> 確認 GPU 驅動對應的 CUDA 版本
3. 以 RTX 3090 為例，點選 nvidia Geforce experience 自動更新 (version: 536.23 -> 546.33)

# 確認 CUDA 版本相依性 -1



## Your GPU Compute Capability

1

Are you looking for the compute capability for your GPU, then check the tables below. You can learn more about Compute Capability here.

NVIDIA GPUs power millions of desktops, notebooks, workstations and supercomputers around the world, accelerating computationally-intensive tasks for consumers, professionals, scientists, and researchers.

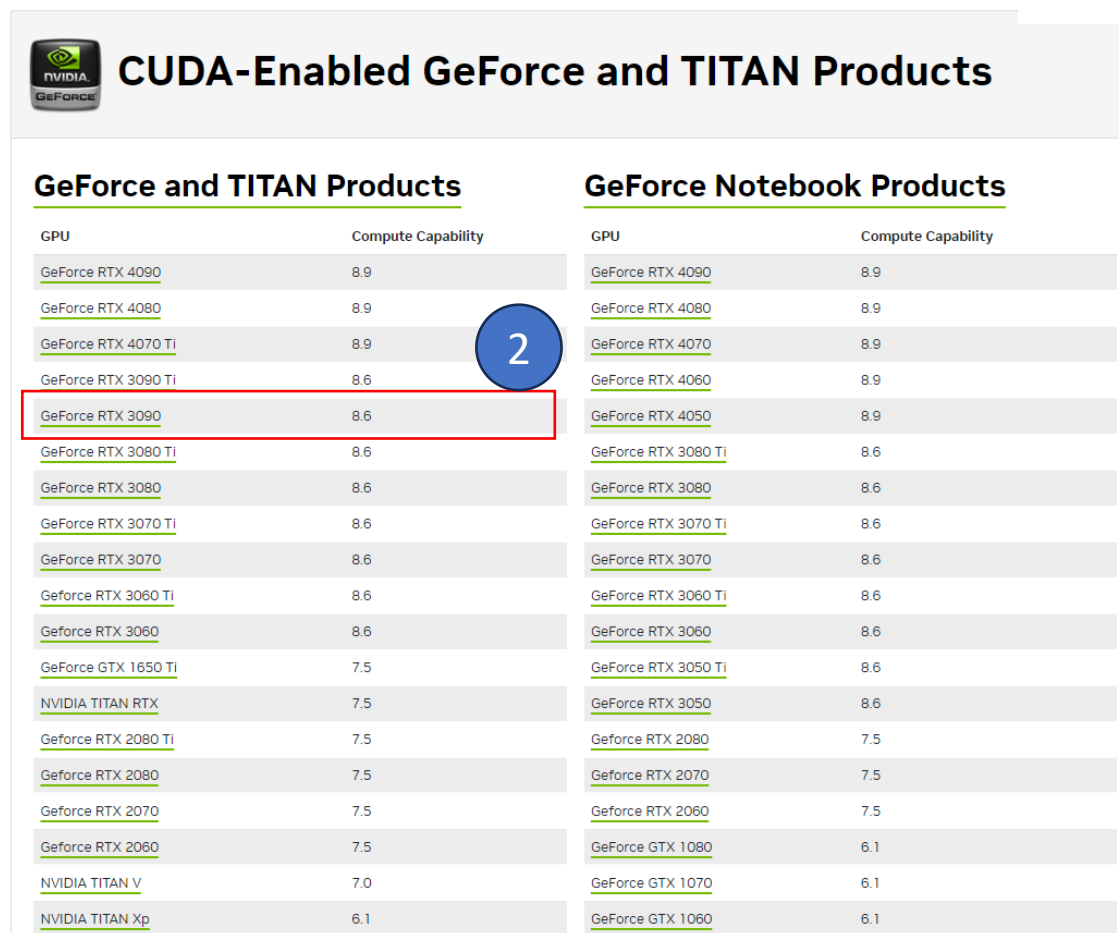
Get started with CUDA and GPU Computing by joining our free-to-join NVIDIA Developer Program.

- Learn about the CUDA Toolkit
- Learn about Data center for technical and scientific computing
- Learn about RTX for professional visualization
- Learn about Jetson for AI autonomous machines

If you have an older NVIDIA GPU you may find it listed on our legacy CUDA GPUs page

Click the sections below to expand

<https://developer.nvidia.com/cuda-gpus>



### CUDA-Enabled GeForce and TITAN Products

GPU	Compute Capability
<a href="#">GeForce RTX 4090</a>	8.9
<a href="#">GeForce RTX 4080</a>	8.9
<a href="#">GeForce RTX 4070 Ti</a>	8.9
<a href="#">GeForce RTX 3090 Ti</a>	8.6
<a href="#">GeForce RTX 3090</a>	8.6
<a href="#">GeForce RTX 3080 Ti</a>	8.6
<a href="#">GeForce RTX 3080</a>	8.6
<a href="#">GeForce RTX 3070 Ti</a>	8.6
<a href="#">GeForce RTX 3070</a>	8.6
<a href="#">GeForce RTX 3060 Ti</a>	8.6
<a href="#">GeForce RTX 3060</a>	8.6
<a href="#">GeForce GTX 1650 Ti</a>	7.5
<a href="#">NVIDIA TITAN RTX</a>	7.5
<a href="#">GeForce RTX 2080 Ti</a>	7.5
<a href="#">GeForce RTX 2080</a>	7.5
<a href="#">GeForce RTX 2070</a>	7.5
<a href="#">GeForce RTX 2060</a>	7.5
<a href="#">NVIDIA TITAN V</a>	7.0
<a href="#">NVIDIA TITAN Xp</a>	6.1

GPU	Compute Capability
<a href="#">GeForce RTX 4090</a>	8.9
<a href="#">GeForce RTX 4080</a>	8.9
<a href="#">GeForce RTX 4070</a>	8.9
<a href="#">GeForce RTX 4060</a>	8.9
<a href="#">GeForce RTX 4050</a>	8.9
<a href="#">GeForce RTX 3080 Ti</a>	8.6
<a href="#">GeForce RTX 3080</a>	8.6
<a href="#">GeForce RTX 3070 Ti</a>	8.6
<a href="#">GeForce RTX 3070</a>	8.6
<a href="#">GeForce RTX 3060 Ti</a>	8.6
<a href="#">GeForce RTX 3060</a>	8.6
<a href="#">GeForce RTX 3050 Ti</a>	8.6
<a href="#">GeForce RTX 3050</a>	8.6
<a href="#">GeForce RTX 2080</a>	7.5
<a href="#">GeForce RTX 2070</a>	7.5
<a href="#">GeForce RTX 2060</a>	7.5
<a href="#">GeForce GTX 1080</a>	6.1
<a href="#">GeForce GTX 1070</a>	6.1
<a href="#">GeForce GTX 1060</a>	6.1

1. 以 RTX 3090 為例
2. 到 NVIDIA DEVELOPER 網站，查看GPU型號的算力版本 (compute capability: 8.6)

# 確認 CUDA 版本相依性 -2

Compute Capability (CUDA SDK support vs. Microarchitecture)

CUDA SDK version(s)	Tesla	Fermi	Kepler (early)	Kepler (late)	Maxwell	Pascal	Volta	Turing	Ampere	Ada Lovelace	Hopper
1.0 <sup>[34]</sup>	1.0 – 1.1										
1.1	1.0 – 1.1+x										
2.0	1.0 – 1.1+x										
2.1 - 2.3.1 <sup>[35][36][37][38]</sup>	1.0 – 1.3										
3.0 - 3.1 <sup>[39][40]</sup>	1.0 –	2.0									
3.2 <sup>[41]</sup>	1.0 –	2.1									
4.0 - 4.2	1.0 –	2.1+x									
5.0 - 5.5	1.0 –			3.5							
6.0	1.0 –			3.5							
6.5	1.1 –				5.x						
7.0 - 7.5		2.0 –			5.x						
8.0		2.0 –				6.x					
9.0 - 9.2			3.0 –				7.0				
10.0 - 10.2			3.0 –					7.5			
11.0 <sup>[42]</sup>				3.5 –					8.0		
11.1 - 11.4 <sup>[43]</sup>				3.5 –					8.6		
11.5 - 11.7.1 <sup>[44]</sup>				3.5 –					8.7		
11.8 <sup>[45]</sup>				3.5 –							9.0
12.0 - 12.2					5.0 –						9.0


GPU 硬體架構

因為 RTX 3090 的 Compute Capability = 8.6  
所以CUDA 最低要求為 11.1 版

1

1. 以 RTX 3090 為例
2. 搭配 CUDA & CC 版本對照表確認版本號最低需求為 11.1 版
3. 以 CC = 8.6 的算力版本而言，CUDA 版本 11.1 – 12.2 都是可以的

# 安裝 PyTorch

PyTorch

Get StartedEcosystem ▾Edge ▾BlogTutorialsDocs ▾Resources ▾GitHub🔍

## INSTALL PYTORCH

Select your preferences and run the install command. Stable represents the most currently tested and supported version of PyTorch. This should be suitable for many users. Preview is available if you want the latest, not fully tested and supported, builds that are generated nightly. Please ensure that you have **met the prerequisites below (e.g., numpy)**, depending on your package manager. Anaconda is our recommended package manager since it installs all dependencies. You can also [install previous versions of PyTorch](#). Note that LibTorch is only available for C++.

**NOTE:** Latest PyTorch requires Python 3.8 or later. For more details, see Python section below.

PyTorch Build	Stable (2.1.2)		Preview (Nightly)	
Your OS	Linux	Mac	Windows	
Package	Conda	Pip	LibTorch	Source
Language	Python		C++ / Java	
Compute Platform	CUDA 11.8	CUDA 12.1	ROCm 5.6	CPU


Run this Command:


```
pip3 install torch torchvision torchaudio --index-url https://download.pytorch.org/whl/cu118
```


[Previous versions of PyTorch >](#)

## QUICK START WITH CLOUD PARTNERS

Get up and running with PyTorch quickly through popular cloud platforms and machine learning services.

 Amazon Web Services >

 Google Cloud Platform >

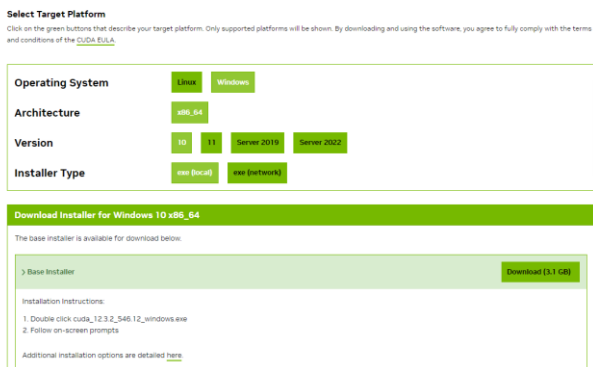
 Microsoft Azure >

Copy to cmd

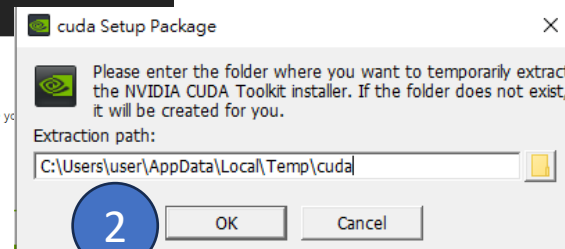
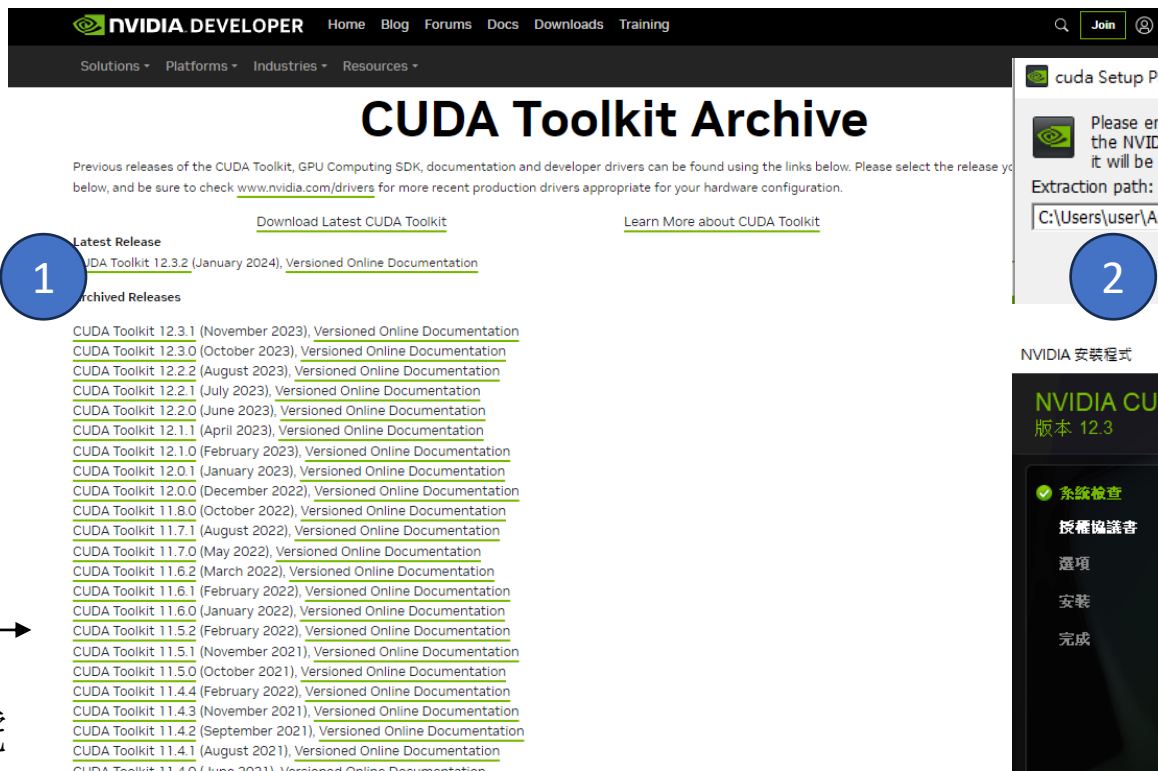
1. 以 RTX 3090 為例
2. 到 pytorch 官網根據硬體型號選擇安裝指令(自動帶出)
3. 須注意最新版本의 python 可能還沒有 pytorch 支援，python 需要降版才可以安裝

# 安裝 CUDA

## CUDA Toolkit 12.3 Update 2 Downloads



不想安裝最新版  
可以到封存區找版本號



<https://developer.nvidia.com/cuda-toolkit-archive>

1. 以 RTX 3090 為例
2. 根據前面統計的版本區間，安裝對應的 CUDA (12.3)
3. 下載完點 OK
4. 用 `nvcc --version` 確認是否安裝成功 (linux 指令, cmd 不會吃)

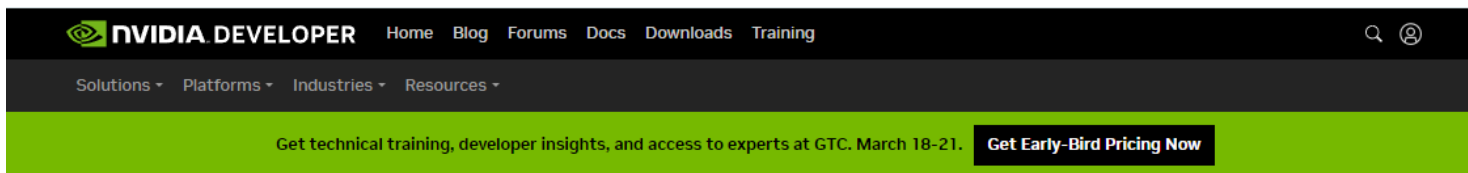
MINGW64:/c:/Users/user

```
user@sunny MINGW64 ~  
$ nvcc --version  
nvcc: NVIDIA (R) Cuda compiler driver  
Copyright (c) 2005-2023 NVIDIA Corporation  
Built on Wed_Nov_22_10:30:42_Pacific_Standard_Time_2023  
Cuda compilation tools, release 12.3, V12.3.107  
Build cuda_12.3.r12.3/compiler.33567101_0
```

4



# 安裝 cuDNN



## cuDNN Download

NVIDIA cuDNN is a GPU-accelerated library of primitives for deep neural networks.

☒ I Agree To the Terms of the [cuDNN Software License Agreement](#)

Note: Please refer to the [Installation Guide](#) for release prerequisites, including supported GPU architectures and compute capabilities, before downloading.

For more information, refer to the cuDNN Developer Guide, Installation Guide and Release Notes on the [Deep Learning SDK Documentation](#) web page.

Download cuDNN v8.9.7 (December 5th, 2023), for CUDA 12.x

Download cuDNN v8.9.7 (December 5th, 2023), for CUDA 11.x

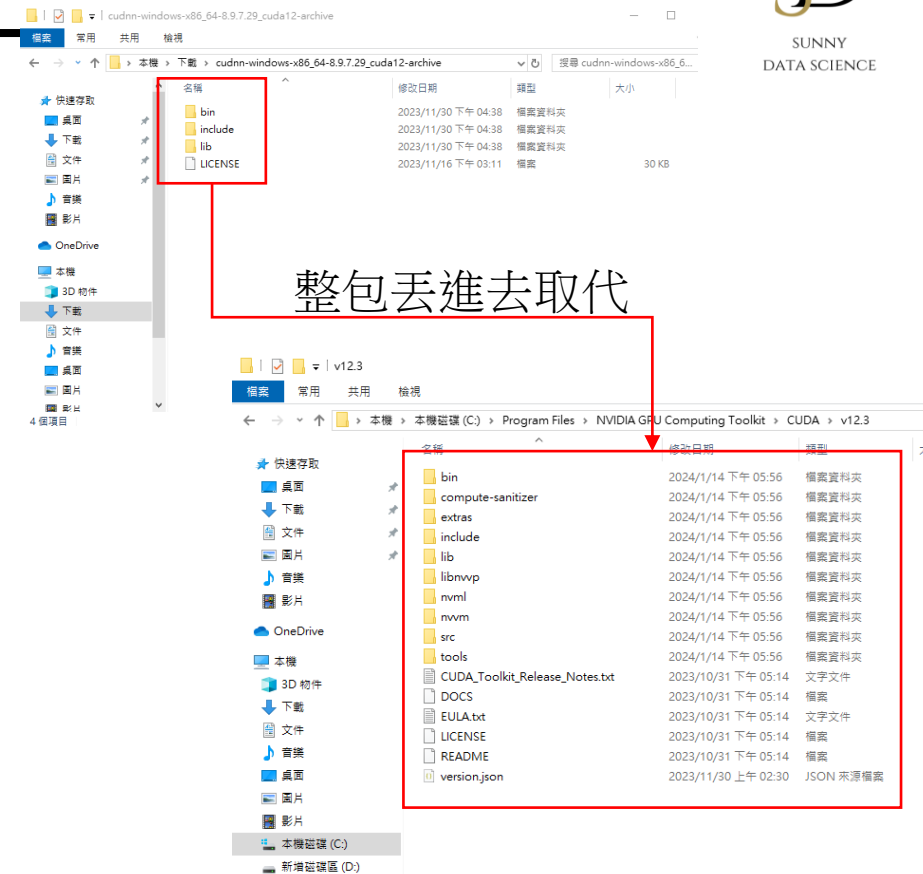
[Archived cuDNN Releases](#)

### Ethical AI

NVIDIA's platforms and application frameworks enable developers to build a wide array of AI applications. Consider potential algorithmic bias when choosing or creating the models being deployed. Work with the model's developer to ensure that it meets the requirements for the relevant industry and use case; that the necessary instruction and documentation are provided to understand error rates, confidence intervals, and results; and that the model is being used under the conditions and in the manner intended.

CUDA 版本要略大於 cuDNN (向下兼容)

舉例 CUDA == 12.1, cuDNN == 12.0



整包丟進去取代

<https://developer.nvidia.com/rdp/cudnn-download>

1. 以 RTX 3090 為例
2. CuDnn 安裝 下載前要用 NVIDIA 帳號進行登入
3. GPU 加速套件
4. 安裝完解壓縮放入對應資料夾

- 將解壓縮後所有內容複製到CUDA安裝路徑
- \*C:\Program Files\NVIDIA GPU Computing Toolkit\CUDA\v12.3\***。(每個人位置可能不一樣)



# Python torch 確認 GPU 是否可讀取成功

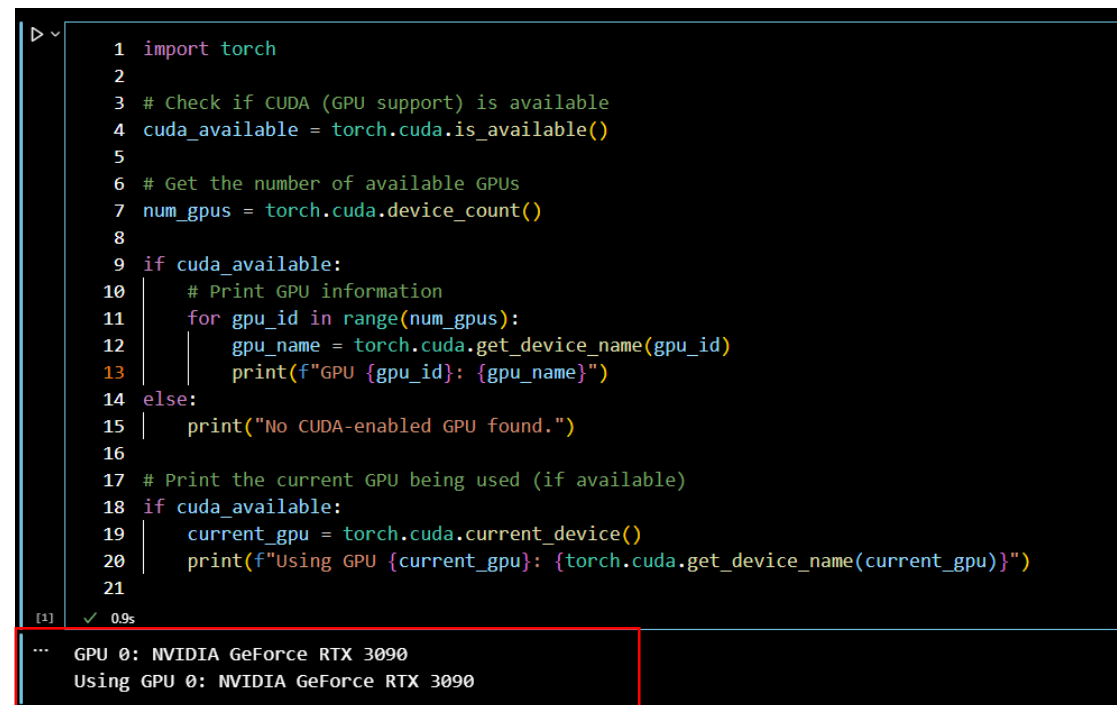
```
import torch

# Check if CUDA (GPU support) is available
cuda_available = torch.cuda.is_available()

# Get the number of available GPUs
num_gpus = torch.cuda.device_count()

if cuda_available:
    # Print GPU information
    for gpu_id in range(num_gpus):
        gpu_name = torch.cuda.get_device_name(gpu_id)
        print(f"GPU {gpu_id}: {gpu_name}")
else:
    print("No CUDA-enabled GPU found.")

# Print the current GPU being used (if available)
if cuda_available:
    current_gpu = torch.cuda.current_device()
    print(f"Using GPU {current_gpu}: {torch.cuda.get_device_name(current_gpu)}")
```



```
1 import torch
2
3 # Check if CUDA (GPU support) is available
4 cuda_available = torch.cuda.is_available()
5
6 # Get the number of available GPUs
7 num_gpus = torch.cuda.device_count()
8
9 if cuda_available:
10     # Print GPU information
11     for gpu_id in range(num_gpus):
12         gpu_name = torch.cuda.get_device_name(gpu_id)
13         print(f"GPU {gpu_id}: {gpu_name}")
14 else:
15     print("No CUDA-enabled GPU found.")
16
17 # Print the current GPU being used (if available)
18 if cuda_available:
19     current_gpu = torch.cuda.current_device()
20     print(f"Using GPU {current_gpu}: {torch.cuda.get_device_name(current_gpu)}")
21
```

[1] ✓ 0.9s

GPU 0: NVIDIA GeForce RTX 3090  
Using GPU 0: NVIDIA GeForce RTX 3090

## 1. 程式碼貼到 jupyter notebook 確認結果