



VAA_NC_REPORTWEEK2_THUONG



LÊ THỊ THƯƠNG



23/4/2021

1. Tìm hiểu về Google Colab

Google Colab (Colaboratory) là một dịch vụ đám mây miễn phí, hiện nay có hỗ trợ GPU (Tesla K80) và TPU (TPUv2). Do được phát triển dựa trên Jupiter Notebook nên việc sử dụng Google Colab cũng tương tự như việc sử dụng Jupyter Notebook. Google Colab là một công cụ lý tưởng để chúng ta rèn luyện kỹ năng lập trình với ngôn ngữ Python thông qua các thư viện của deep learning. Google Colab cài đặt sẵn cho chúng ta những thư viện rất phổ biến trong nghiên cứu Deep Learning như PyTorch, TensorFlow, Keras và OpenCV.

Google Colab cho phép chạy các dòng code python thông qua trình duyệt, đặc biệt phù hợp với Data analysis, machine learning và giáo dục. Colab không cần yêu cầu cài đặt hay cấu hình máy tính, mọi thứ có thể chạy thông qua trình duyệt, bạn có thể sử dụng tài nguyên máy tính từ CPU tốc độ cao và cả GPUs và cả TPUs đều được cung cấp cho bạn.

2. Tìm hiểu về GPU và các thông số của nó

2.1 GPU

GPU là từ viết tắt của Graphics Processing Unit – gọi là Đơn vị xử lý đồ họa chuyên dụng nhận nhiệm vụ tăng tốc, xử lý đồ họa cho bộ xử lý trung tâm CPU. GPU có các tính năng vượt trội, vượt xa so với các trình điều khiển đồ họa cơ bản (ví dụ như GPU của Intel).

Các GPU hiện đại có năng suất rất cao trong xử lý đồ họa máy tính. Với cấu trúc mang tính xử lý song song mạnh mẽ của mình, tốc độ xử lý dữ liệu GPU có thể tiếp nhận hàng ngàn luồng dữ liệu cùng một lúc vì thế có thể tăng tốc một số phần mềm tới hơn 100 lần so với một CPU.

GPU được sử dụng trong các hệ thống nhúng, điện thoại di động, máy tính cá nhân, máy trạm, máy chơi game v...v.. Trong máy tính cá nhân, một GPU có thể xuất hiện ở card đồ họa, hoặc nó cũng có thể được gắn trên mainboard.

GPU cũng là con chip điện tử nhưng có chức năng như một bộ vi xử lý riêng của card đồ họa, tự nó có đủ khả năng và sức mạnh xử lý tất cả mọi vấn đề có liên quan tới hình ảnh của máy tính. Hai dòng chip đồ họa phổ biến nhất là: NVIDIA và AMD

2.2 GPU Google Colab sử dụng

▼ GPU (Tesla K80)

- tổng bộ nhớ 2 x 12 GB GDDR5
- băng thông bộ nhớ 480 GB/s
- hiệu năng tính toán đạt 8,74 TFLOPS (single-precision) và 2,91 TFLOPS (double-precision)
- GPU Boost cho phép đẩy hiệu năng xử lý đối với một số ứng dụng đặc biệt
- 4992 lõi xử lý CUDA
- cần nguồn điện tối thiểu là 300 W

3. Python version trên Google Colab

Google Colab hỗ trợ trên Python :2.7 và 3.6

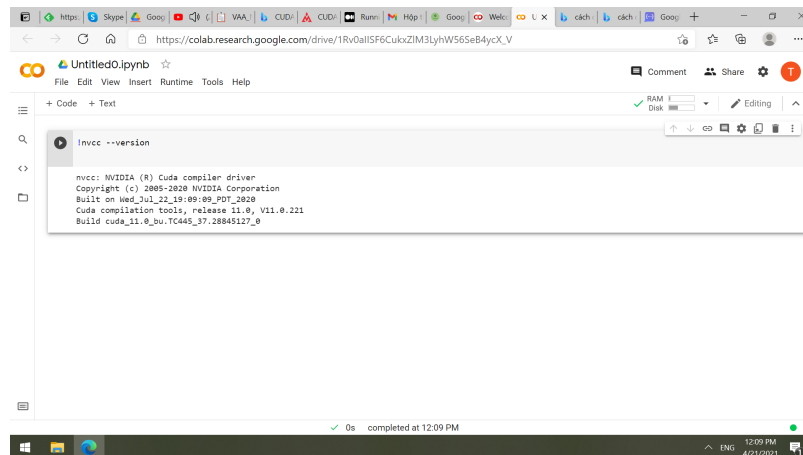
<https://trituenhantao.io/lap-trinh/lam-quen-voi-google-colab/> -Nguồn tham khảo

4. Tìm hiểu về CUDA và Cudnn

4.1 CUDA

CUDA (*Compute Unified Device Architecture*) - Kiến trúc hợp nhất tính toán của các thiết bị điện tử được phát triển độc quyền bởi hãng công nghệ NVIDIA. Còn về CUDA core hay nhân CUDA thì chúng ta có thể hiểu đây là một nhân xử lý trong GPU của card đồ họa - đơn vị chịu trách nhiệm tính toán các thông tin, dữ liệu đồ họa cần được kết xuất. Nhân CUDA tích hợp trong GPU của card đồ họa rời, càng nhiều nhân CUDA thì khả năng tính toán đồng thời nhiều thông tin càng nhanh và chính xác

CUDA version trên Google Colab :11.0



```
nvcc --version

nvcc: NVIDIA (R) Cuda compiler driver
Copyright (c) 2005-2020 NVIDIA Corporation
Built on Wed Jul 22 19:09:09 PDT 2020
Cuda compilation tools, release 11.0, V11.0.221
Build cuda_11_0_nv.TC445_37.28845127_0
```

Kiểm tra version của CUDA

4.2 Cudnn

CUDA Deep Neural Network library (cuDNN) là một thư viện giúp tăng tốc GPU khi huấn luyện các model deep learning. Thư viện này cung cấp quá trình tối ưu huấn luyện feed forward và backpropagation trên các layers: convolution, pooling, normalization, activation.