

**TRƯỜNG ĐẠI HỌC TRẦN ĐẠI NGHĨA**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO TUẦN**

**MÔN HỌC: MẠNG NEURAL**

**TP. HỒ CHÍ MINH, THÁNG 08 NĂM 2020**



**TRƯỜNG ĐẠI HỌC TRẦN ĐẠI NGHĨA**

**KHOA CÔNG NGHỆ THÔNG TIN**

**BÁO CÁO TUẦN**

**MÔN HỌC: MẠNG NEURAL**

**Nhóm báo cáo:**

**Nguyễn Tiểu Phụng**

**Huỳnh Đức Anh Tuấn**

**Giảng viên hướng dẫn:**

**Th.s Ngô Thanh Tú**

**TP. HỒ CHÍ MINH, THÁNG 06 NĂM 2020**



# **MỤC LỤC**

[**MỤC LỤC** 3](#_Toc47482489)

[CHƯƠNG 1: TÌM HIỂU NGÔN NGỮ LẬP TRÌNH PYTHON 1](#_Toc47482490)

[1.1. Giới thiệu ngôn ngữ lập trình Python 1](#_Toc47482491)

[1.1.1. Lịch sử phát triển 1](#_Toc47482492)

[1.1.2. Phiên bản 1](#_Toc47482493)

[1.1.3. Một số điểm khác nhau giữa phiên bản 3x và 2x 2](#_Toc47482494)

[1.1.4. Đặc điểm của Python 5](#_Toc47482495)

[1.2. Hướng dẫn cài đặt bằng Anaconda/Miniconda 6](#_Toc47482496)

[1.2.1. Giới thiệu Ananconda/Miniconda và cài đặt 6](#_Toc47482497)

[1.2.2. Download Anaconda/Miniconda và hướng dẫn cài đặt 6](#_Toc47482498)

[1.2.3. Hướng dẫn cài đặt CUDA và cuDNN 13](#_Toc47482499)

[1.2.4. Hướng dẫn cài thêm thư viện bằng conda 13](#_Toc47482500)

CHƯƠNG 1: TÌM HIỂU TỔNG QUAN

1. Giới thiệu về Machine Learning
2. Lịch sử phát triển

* 1950 - Nhà bác học Alan Turing đã tạo ra "Turing Test (phép thử Turing)" để xác định xem liệu một máy tính có trí thông minh thực sự hay không. Để vượt qua bài kiểm tra đó, một máy tính phải có khả năng đánh lừa một con người tin rằng nó cũng là con người.
* 1952 - Arthur Samuel đã viết ra chương trình học máy (computer learning) đầu tiên. Chương trình này là trò chơi cờ đam, và hãng máy tính IBM đã cải tiến trò chơi này để nó có thể tự học và tổ chức những nước đi trong chiến lược để giành chiến thắng.
* 1957 - Frank Rosenblatt đã thiết kế mạng nơron (neural network) đầu tiên cho máy tính, trong đó mô phỏng quá trình suy nghĩ của bộ não con người.
* 1967 - Thuật toán "nearest neighbor" đã được viết, cho phép các máy tính bắt đầu sử dụng những mẫu nhận dạng (pattern recognition) rất cơ bản. Nó được sử dụng để vẽ ra lộ trình cho một người bán hàng có thể bắt đầu đi từ một thành phố ngẫu nhiên nhưng đảm bảo anh ta sẽ đi qua tất cả các thành phố khác theo một quãng đường ngắn nhất.
* 1990s - Machine Learning đã dịch chuyển từ cách tiếp cận hướng kiến thức (knowledge-driven) sang cách tiếp cận hướng dữ liệu (data-driven). Các nhà khoa học bắt đầu tạo ra các chương trình cho máy tính để phân tích một lượng lớn dữ liệu và rút ra các kết luận - hay là "học" từ các kết quả đó.
* 006 - Geoffrey Hinton đã đưa ra một thuật ngữ "deep learning" để giải thích các thuật toán mới cho phép máy tính "nhìn thấy" và phân biệt các đối tượng và văn bản trong các hình ảnh và video.
* 2010 - Microsoft Kinect có thể theo dõi 20 hành vi của con người ở một tốc độ 30 lần mỗi giây, cho phép con người tương tác với máy tính thông qua các hành động và cử chỉ.
* 2011 - Google Brain đã được phát triển, và mạng deep nơron (deep neural network) của nó có thể học để phát hiện và phân loại nhiều đối tượng theo cách mà một con mèo thực hiện.
* 2014 - Facebook phát triển DeepFace, một phần mềm thuật toán có thể nhận dạng hoặc xác minh các cá nhân dựa vào hình ảnh ở mức độ giống như con người có thể.
* 2015 - Amazon ra mắt nền tảng machine learning riêng của mình.
* 2015 - Microsoft tạo ra Distributed Machine Learning Toolkit, trong đó cho phép phân phối hiệu quả các vấn đề machine learning trên nhiều máy tính.

1. Machine Learning là gì?

* Machine Learning là một tập con của AI. Theo định nghĩa của Wikipedia, Machine learning is the subfield of computer science that “gives computers the ability to learn without being explicitly programmed”.
* Nói đơn giản, Machine Learning là một lĩnh vực nhỏ của Khoa Học Máy Tính, nó có khả năng tự học hỏi dựa trên dữ liệu đưa vào mà không cần phải được lập trình cụ thể.
* Machine Learning thường được chia thành học có giám sát, trong đó máy tính học bằng ví dụ từ dữ liệu được gắn nhãn và học không giám sát, trong đó các máy tính nhóm các dữ liệu tương tự và xác định chính xác sự bất thường.

1. Deep-Learing?

* Deep Learing là nhóm thuật toán nhỏ của Machine Learning lấy ý tưởng dựa trên Neural Network (mạng neuron) của con người. Deep Learning thường yêu cầu lượng dữ liệu lớn và nguồn tài nguyên sử dụng nhiều hơn các phương pháp thông thường, tuy nhiên cho độ chính xác cao hơn.
* Những dấu mốc quan trọng trong lịch sử phát triển của DeepLearning.
* Perceptron (60s) - Perceptron là một thuật toán supervised learning giúp giải quyết bài toán phân lớp nhị phân, được khởi nguồn bởi Frank Rosenblatt năm 1957 trong một nghiên cứu được tài trợ bởi Văn phòng nghiên cứu hải quân Hoa Kỳ (U.S Office of Naval Research – từ một cơ quan liên quan đến quân sự).
* MLP và Backpropagation ra đời (80s)- Năm 1986, Geoffrey Hinton cùng với hai tác giả khác xuất bản một bài báo khoa học trên Nature với tựa đề “Learning representations by back-propagating errors”. Trong bài báo này, nhóm của ông chứng minh rằng neural nets với nhiều hidden layer (được gọi là multi-layer perceptron hoặc MLP) có thể được huấn luyện một cách hiệu quả dựa trên một quy trình đơn giản được gọi là backpropagation.
* Năm 2006 - neural networks với nhiều hidden layer được đổi tên thành deep learning.

|  |
| --- |
| Những dấu mốc quan trọng của Deep Learning |
|  |

1. Neural Network (Mạng Nơ-ron)

* Neural Network là một hệ thống các chương trình và cấu trúc dữ liệu mô phỏng cách vận hành của não người. Mỗi nơ-ron trong mạng nơ-ron là một hàm toán học lấy dữ liệu thông qua đầu vào, biến đổi dữ liệu đó thành dạng dễ điều chỉnh hơn và sau đó phun ra thông qua đầu ra.

|  |
| --- |
| The UIT Information Security Laboratory |
|  |

1. Hướng dẫn cài đặt bằng Anaconda/Miniconda
2. Giới thiệu Ananconda/Miniconda và cài đặt

* Anaconda là một Distribution miễn phí và mã nguồn mở của Python và R giúp đơn giản hóa việc cài đặt, quản lý và triển khai packages (numpy, scipy, tensorflow, ...).
* Anaconda phục vụ cho nhiều mục địch, đặc biệt trong Data Science (Khoa học dữ liệu), Machine learnig (Máy học), Big Data (Dữ liệu lớn), Image Processing (Xử lý ảnh), ...
* Anaconda hiện nay đã có hơn 20 triệu người dùng và hơn 7500 packages khoa học dữ liệu dành cho Windows, Linux và MacOS.
* Trong khi đó Spyder là 1 trong những IDE (môi trường tích hợp dùng để phát triển phần mềm) tốt nhất cho data science và quang trọng hơn là nó được cài đặt khi bạn cài đặt Anaconda.

1. Download Anaconda/Miniconda và hướng dẫn cài đặt
2. Các bước cài đặt

* Yêu cầu phần cứng và phần mềm:
* Hệ điều hành: Win 7, Win 8/8.1, Win 10, Red Hat Enterprise Linux/CentOS 6.7, 7.3, 7.4, and 7.5, and Ubuntu 12.04+.
* Ram tối thiểu 4GB.
* Ổ cứng trống tối thiểu 3GB để tải và cài đặt.

**Cài đặt:**

* Bước 1: truy cập vào trang web <https://www.anaconda.com/>

|  |
| --- |
|  |
| Hình 1.1 Trang chủ ananconda.com |

* Chọn xuống mục download: và chọn tải phiên bản thích hợp, ở đây em chọn hệ điều hành windows bản python 3.7 và 64-bit Graphical.

|  |
| --- |
|  |
| Hình 1.2 Download bản cài đặt |

* Bước 2: Chạy file cài đặt với quyền admin

|  |
| --- |
|  |
| Hình 1.3 tiến hình cài đặt |

* Bước 3: Chấp nhập các yêu cầu thiết lập và tiến hình cài đặt

|  |  |
| --- | --- |
|  |  |
| Hình 1.4 Giao diện cài đặt | |

* Bước 4: Chọn tài khoản và vị trí để cài đặt. Ở đây em chọn recommended và ổ C:/

|  |  |
| --- | --- |
|  |  |
| Hình 1.5 Chọn tài khoản và vị trí để cài đặt | |

* Bước 5: Giao diện Anaconda khi cài đặt xong

|  |
| --- |
|  |
| Hình 1.6 Giao diện của phần mền Anaconda |

* Bước 6: Kiểm tra lại phiên bản phần mềm Anaconda

|  |
| --- |
| Hình 1.7 Giao diện home của Anaconda |

* Chọn Launch ở Environments base(root) để mở CMD ở môi trường conda để kiểm tra lại phiên bản cài đặt bằng lệnh: ***conda -V***

|  |
| --- |
|  |
| Hình 1.8 kiểm tra phiên bản conda |
|  |
| Hình 1.9 kiểm tra phiên bản python |

1. Quản lý môi trường

* Với Anaconda có nhiều packages khoa học phụ thuộc vào các phiên bản cụ thể của các packages khác. Các nhà khoa học dữ liệu thường sử dụng nhiều phiên bản của nhiều package và sử dụng nhiều môi trường để phân tách các phiên bản khác nhau này.
* Chương trình dòng lệnh (command-line program conda) vừa là trình quản lý các package vừa là trình quản lý môi trường (environment manager). Điều này giúp các nhà khoa học dữ liệu đảm bảo rằng mỗi phiên bản của mỗi package có tất cả các phụ thuộc mà nó yêu cầu và hoạt động chính xác.
* Anaconda Navigator cung cấp cho người dùng một giao diện đồ họa để quản lý các environment (môi trường) và package. Ta sẽ có environment mặc định là base (root) chứa các package cơ bản.
* Ở ngăn giao diện Home là nơi quản lý các Application (ứng dụng) tại một environment (trong vòng đỏ).

|  |
| --- |
|  |
| Hình 1.10 Giao diện environment |

|  |
| --- |
| 1  2  3 |
| Hình 1.11 các ngăn của giao diện Environments |

* Vùng số 1 là danh sách các environment ta đã tạo.
* Vùng số 2 là nút để tạo environment mới, sau nhấp chuột vào ta sẽ có giao diện như hình 1.12. Chúng ta chọn version của Python và đặt tên cho môi trường.
* Tương tự Clone là để sao chép một bản environment với các package giống một environment đã tạo. Import dùng để tạo environment bằng file có sẵn. Remove để xóa environment.
* Vùng thứ 3 dùng để tìm kiếm và cài đặt các package trong environment bạn đã chọn trong vùng thứ nhất.
* Ví dụ tạo mới một environment imgprocess.

|  |
| --- |
|  |
| Hình 1.12 tạo mới một environment imgprocess |
|  |
| Hình 1.13 environment imgprocess đã được tạo thành công |

1. Hướng dẫn cài đặt CUDA và cuDNN
2. Cài đặt CUDA

* Vào trang: <https://developer.nvidia.com/cuda-toolkit> chọn DOWNLOADS

|  |
| --- |
|  |
|  |

* Chọn tiếp CUDA Toolkit 🡪 DOWNLOAD

|  |
| --- |
|  |
|  |

* Tiếp tục nhấp vào các nút màu xanh mô tả nền tảng phù hợp (nên chọn Type exe local). Bằng cách tải xuống và sử dụng phần mềm, bạn cần đồng ý tuân thủ đầy đủ các điều khoản và điều kiện của CUDA EULA.

|  |
| --- |
|  |
|  |

* Sau khi download thành công chuyển sang bước cài đặt.

|  |
| --- |
|  |
|  |

* Quá trình cài đặt trải qua 5 giai đoạn.

|  |
| --- |
|  |
|  |

1. Tải cuDNN cho bản Pytorch 10.1

* Truy cập đường dẫn: <https://developer.nvidia.com/rdp/cudnn-archive> và tiến hành đăng nhập (nếu có) hoặc đăng ký tài khoản (nếu chưa có).

|  |
| --- |
|  |
|  |

* Chọn phiên bản phù hợp nhất (10.1) và mới nhất và chọn hệ điều hành phù hợp sau đó tiến hành tải bộ thư viện về.

|  |
| --- |
|  |
|  |

1. Hướng dẫn cài thêm thư viện bằng conda
2. Các thư viện sử dụng

* Tạo một environment neural để cài các thư viện hỗ trợ.

|  |
| --- |
|  |
| Hình 1.14 create environment neural |

* Các thư viện cần cài đặt thêm như:
* Pytorch
* Numpy
* Matloplit

1. Cài đặt pytorch bằng dòng lệnh

* Truy cập vào trang chủ <https://pytorch.org/>
* Lựa chọn các gói hỗ trợ phù hợp (Sử dụng bản 10.1).

|  |
| --- |
|  |
| Hình 1.15 Các điều kiện cài đặt |

* Mở CMD.exe Prompt với environment neural
* Thư viện openCV: dùng lệnh ***conda install pytorch torchvision cudatoolkit=10.1 -c pytorch***

|  |
| --- |
|  |
| Hình 1.16 cài thư viện Pytorch trên env neural bằng cmd |

1. Bài toán mô phỏng Training a classifier

Chuẩn bị dữ liệu

* Bộ dữ liệu được sử dụng gồm 10 lớp từ CIFAR-10: ‘airplane’, ‘automobile’, ‘bird’, ‘cat’, ‘deer’, ‘dog’, ‘frog’, ‘horse’, ‘ship’, ‘truck’.
* Các hình ảnh trong CIFAR-10 có kích thước 3x32x32, tức là hình ảnh màu 3 kênh có kích thước 32x32 pixel.

Đào tạo lớp phân loại hình ảnh gồm 5 bước:

* Tải và chuẩn hóa bộ dữ liệu kiểm tra và huấn luyện CIFAR10 bằng cách sử dụng torchvision.
* Xác định một mạng nơ ron kết hợp.
* Xác định chức năng Loss và tối ưu hóa.
* Đào tạo cho mạng dữ liệu.
* Kiểm tra mạng với bộ dữ liệu thử nghiệm.