**Hướng dẫn sử dụng**

**Dự án: Nhận dạng chữ số in trên bằng tốt nghiệp**

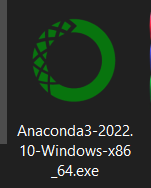
1. **Yêu cầu và chuẩn bị cơ bản**

Yêu cầu:

* Chạy chương trình trên Anaconda bằng máy tính (Anaconda là nền tảng mã nguồn mở về Khoa học dữ liệu trên Python thông dụng nhất hiện nay).
* Cấu hình máy:
  + 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++ CPU: Intel, Core i3 trở lên.
  + Ram tối thiểu 4GB.
  + Ổ cứng: HDD, SSD, ổ cứng trống tối thiểu 3GB để tải và cài đặt.

**Bước 1**: **Download Anaconda: Bạn click vào** [**đây**](https://repo.anaconda.com/archive/Anaconda3-2022.10-Windows-x86_64.exe)

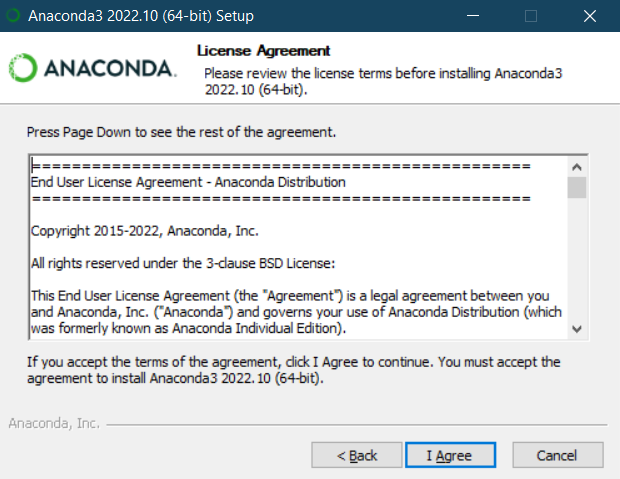
**Bước 2**: **Sau khi download Anaconda thành công, file sẽ được lưu trong mục download của máy tính**



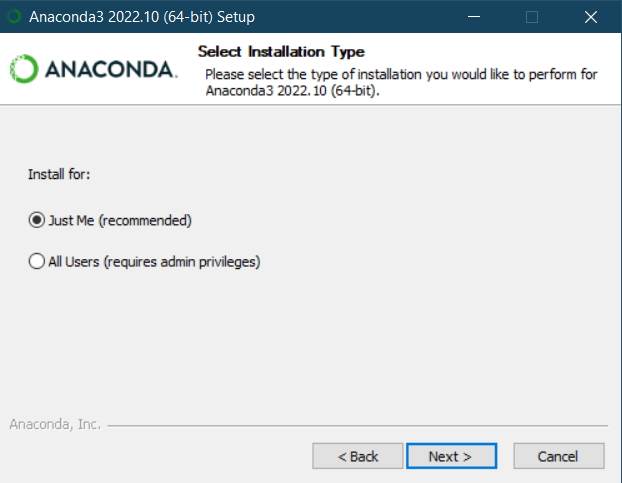
Hình 1.1. click vào bộ cài để cài đặt chương trình



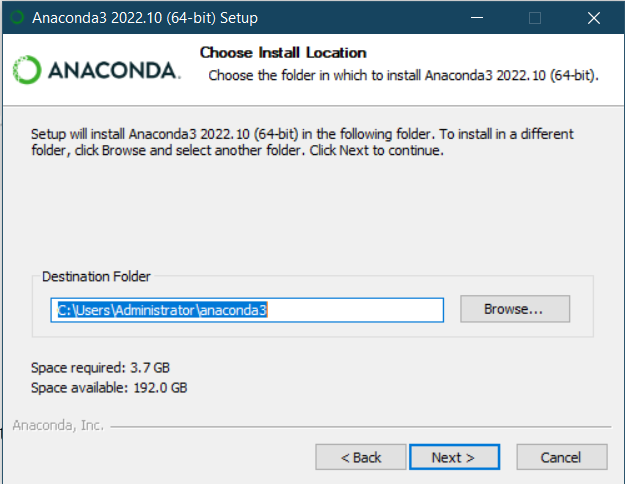
Hình 1.2. Click Next



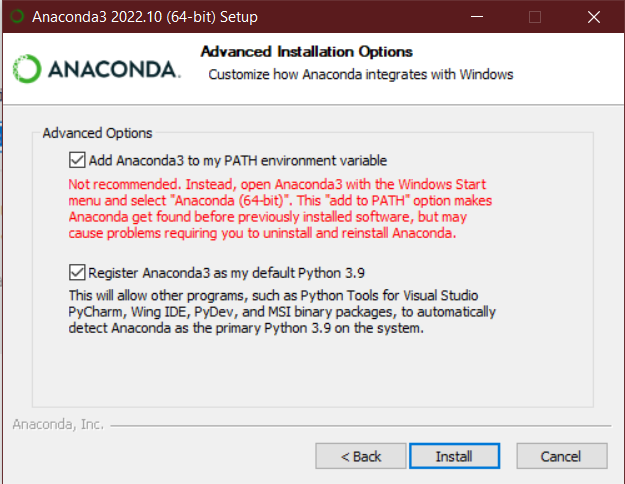
Hình 1.3. Click I Agree



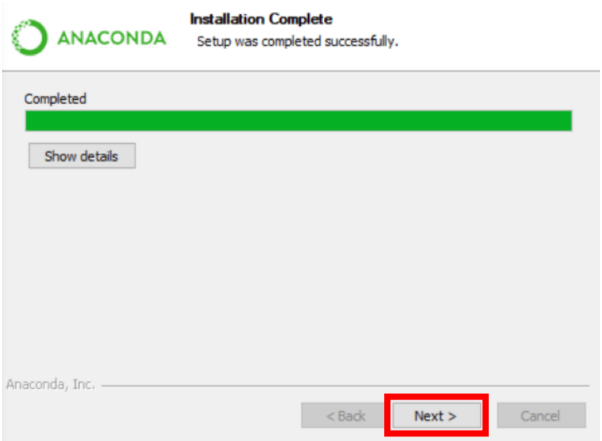
Hình 1.4: Chọn Just me và Click Next



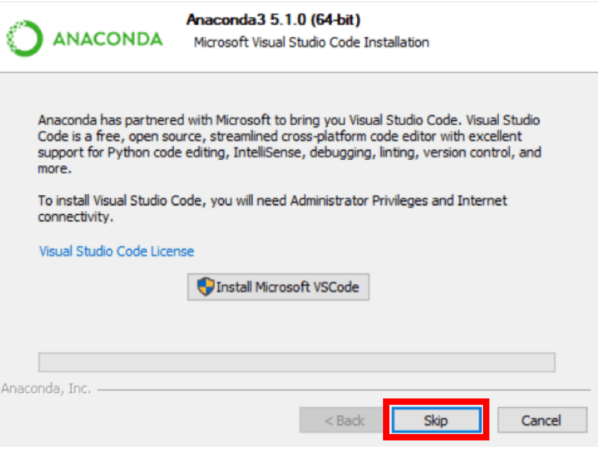
Hình 1.5: Chọn đường dẫn và Click Next



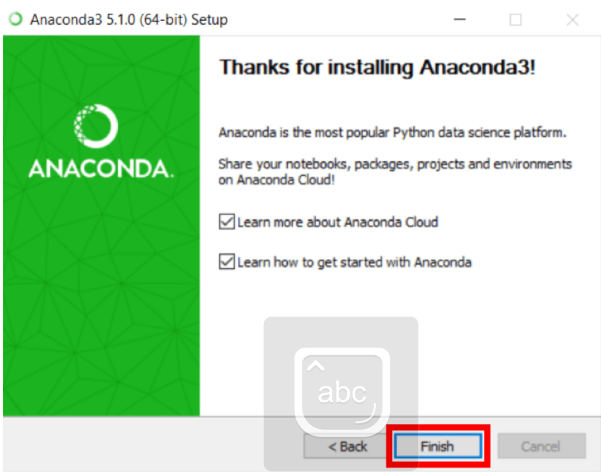
Hình 1.6: Tích vào add Anaconda3 to my PATH environment variable và Click Install để cài đặt



Hình 1.7: Click Next



Hình 1.8. Click Skip

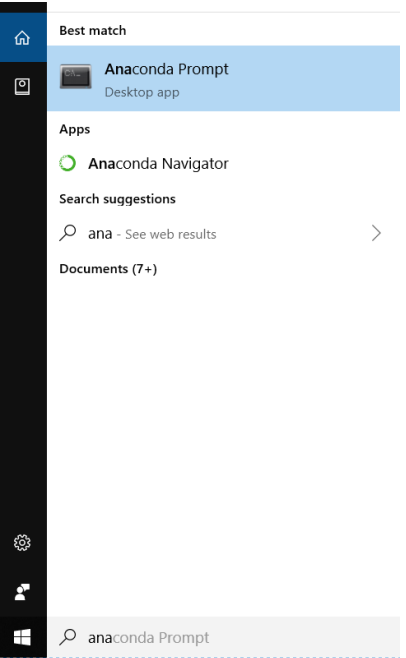


Hình 1.9. Click Finish.

Như vậy ta đã cài đặt xong Anaconda.

**Bước 3:** **Cài đặt một số library cơ bản**

Mở anaconda command line và copy lệnh sau để cài library.



Hình 1.10. Mở Anaconda prompt

Một số thư viện cơ bản

• Numpy: conda install -c anaconda numpy

• Pandas: conda install -c anaconda pandas

• Matplotlib: conda install -c conda-forge matplotlib

• Keras: conda install -c conda-forge keras

Hoặc bạn cũng có thể cài bằng pip trên anaconda prompt

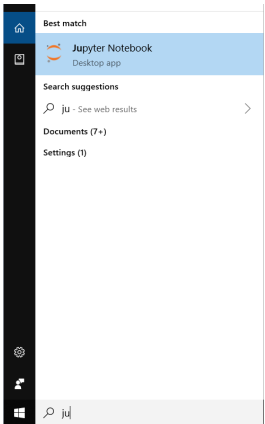
• Numpy: pip install numpy

• Pandas: pip install pandas

• Matplotlib: pip install matplotlib

• Keras: pip install Keras

**Bước 4:** **Hướng dẫn sử dụng Jupyter notebook.**



Hình 1.11. Mở jupyter notebook

Giao diện mở ra, click vào new chọn Python3 để tạo 1 notebook mới.

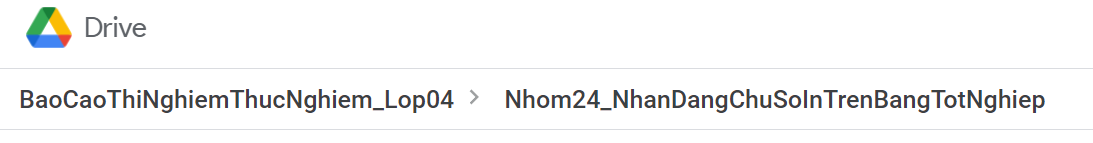


Hình 1.12. Tạo file mới trên jupyter notebook

* Sau đấy thì dùng giống như google colab, nhưng điểm khác biệt là jupyter notebook sẽ chạy trên máy của bạn chứ không phải cloud như google colab.

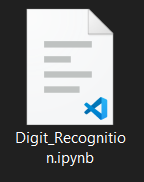
1. **Thực hiện tải và chạy đề tài.**

**Bước 1: Click vào** [**đây**](https://drive.google.com/drive/folders/1yQOhEvtigHkzZJb1WARO0pXzU17PXvcd) **để lấy file code đề tài.**



Hình 2.1. Tải file code

**Bước 2: Nhấn download để tải file code về máy, file tải về sẽ được lưu tại thư mục download trên máy tính của bạn.**



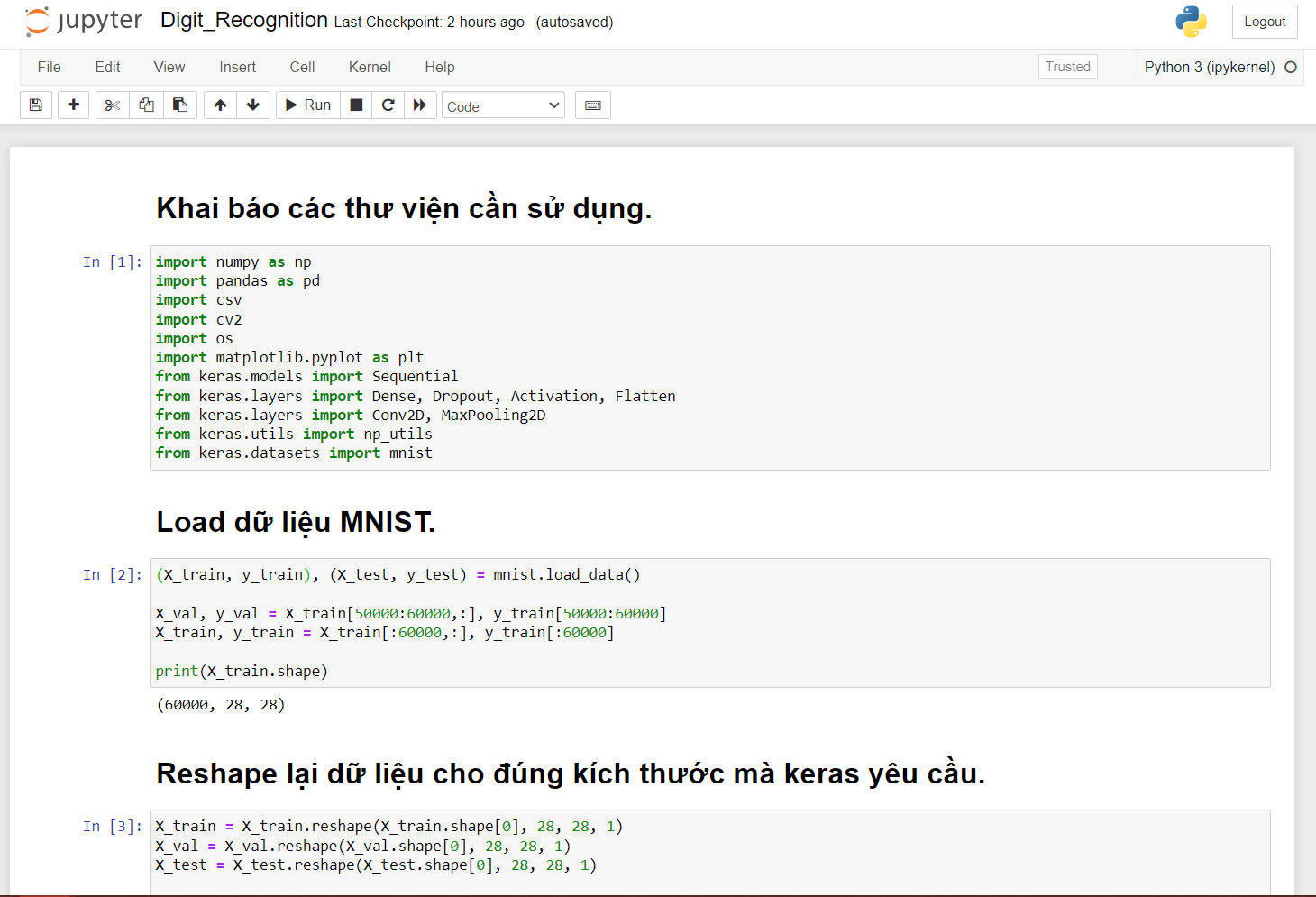
Hình 2.2. File được lưu tại mục download

**Bước 3: Click vào upload trên jupyter notebook để up file lên và sử dụng**



Hình 2.3. Chọn file code

**Bước 4: Sau khi upload file code lên, kích đúp để chạy file và Nhấn Run để chạy code.**



Hình 2.4. Mở file code

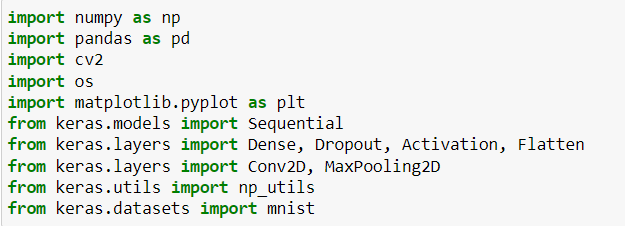
* Kết quả ví dụ:



Hình 2.5. Ví dụ nhận dạng

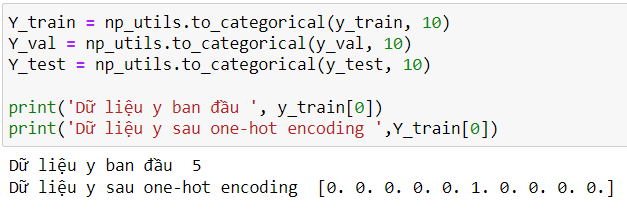
1. **Một số giải thích.**

* Các thư viện được sử dụng:



Hình 3.1. Khai báo các thư viện cần sử dụng.

* Vì mỗi ảnh sẽ thuộc 1 class từ 0 đến 9, nên tất cả sẽ có 10 class. Nên output layer sẽ có 10 node để tương ứng với phần trăm ảnh là số 0, 1,… , 9. Softmax function - thay vì chỉ muốn kết quả là ảnh là số mấy, ta muốn dự đoán phần trăm của ảnh là số nào. Ví dụ: 90% ảnh là số 5, 1% ảnh là số 1, ...



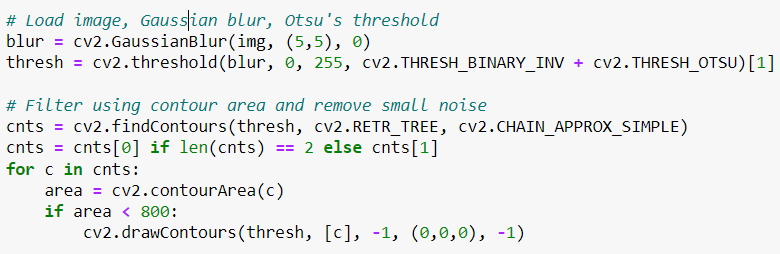
Hình 3.2. Áp dụng one hot encoding

* Thực hiện định nghĩa và huấn luyện mô hình



Hình 3.3. Mô hình

* Sử dụng Gaussian blur và Otsu's threshold để lọc các điểm ảnh nhỏ gây nhiễu trong quá trình nhận diện

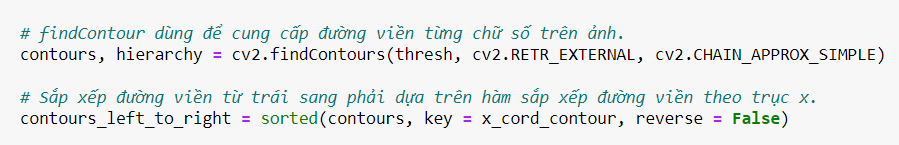


Hình 3.4. Xử lý ảnh

* Sắp xếp đường viền chữ số từ trái qua phải dựa trên trục x



Hình 3.5. Sắp xếp viền



Hình 3.6. Xử lý viền

Một vài nhận xét về chương trình:

* Chương trình đã nhận diện được số hiệu trên bằng tốt nghiệp. Đôi khi chữ số bị khuất một chút vẫn đọc được.
* Tuy nhiên vẫn còn một số nhược điểm:
* Khi ảnh đầu vào bị đặt một góc quá nghiêng thì một vài chữ số không nhận diện được.
* Đôi khi bị nhận dạng nhầm giữa 1, 4 và 7
* Hoạt động kém khi ảnh mờ và có độ phân giải thấp.