

In this homework assignment, you will download the MNIST DATABASE of handwritten digits in the green box below on website

<http://yann.lecun.com/exdb/mnist/> to solve the following problems.

(Hint: Try to download the dataset directly in the program with code will be easier. )

## THE MNIST DATABASE

### of handwritten digits

[Yann LeCun](#), Courant Institute, NYU  
[Corinna Cortes](#), Google Labs, New York  
[Christopher J.C. Burges](#), Microsoft Research, Redmond

The MNIST database of handwritten digits, available from this page, has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger set available from NIST. The digits have been size-normalized and centered in a fixed-size image.

It is a good database for people who want to try learning techniques and pattern recognition methods on real-world data while spending minimal efforts on preprocessing and formatting.

Four files are available on this site:

<code>train-images-idx3-ubyte.gz</code>	training set images (9912422 bytes)
<code>train-labels-idx1-ubyte.gz</code>	training set labels (28881 bytes)
<code>t10k-images-idx3-ubyte.gz</code>	test set images (1648877 bytes)
<code>t10k-labels-idx1-ubyte.gz</code>	test set labels (4542 bytes)

The explanation of database as follow :

The MNIST database of handwritten digits, has a training set of 60,000 examples, and a test set of 10,000 examples. It is a subset of a larger set available from NIST. The digits have been size-normalized and centered in a fixed-size image.



Analysis the data.

PLZ Mark The Question Number Clearly!

(Hint: For doing this homework, you can use “Google Colaboratory” for free GPU to overcome the difficulties of computer hardware if you need it !)

1. Try to print out 0~9 so that knowing the things you are going to input in the model.
2. Using CNN for handwriting recognition. PLZ give the reason why you choose the specific Loss function, Optimizer, etc.
3. Try to print out some output to make sure it is right.
4. Use ROC, recall... to see the model performance.