**Supervised**

**MNIST Dataset**

## **MNIST Dataset:**

## The MNIST dataset contains 60,000 images of handwritten digits (zero to nine) Each image is a 28 × 28 × 1 array of floating-point numbers representing grayscale intensities ranging from 0 (black) to1 (white). The target data consists of one-hot binary vectors of size 10.

**KNN:**

K nearest neighbors is a classification algorithm. It classifies the new data point (test data) into some category by nearest class.

**centroid features:**

function that split image to small windows like 5\*5 and compute center to each window and put it in vector.

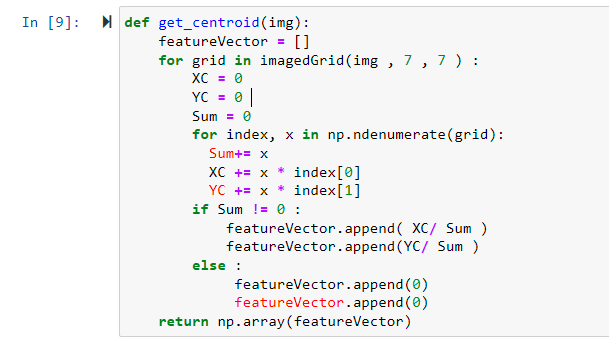
**For my assignment:**

I split the dataset into two subsets:

10,000 for training (10,000 to 20,000)

, 1000 images for testing.

I get centroid of window with size (7\*7)



Then use KNN to classify and change k value like table:

**Accuracy:**

|  |  |
| --- | --- |
| K | Accuracy |
| 290 | **64.3 %** |
| 200 | **67.5 %** |
| 120 | **72.8 %** |
| 70 | **74.1 %** |
| 30 | **79.9 %** |
| 13 | **82.39 %** |
| 7 | **84.3 %** |
| 1 | **86.7 %** |