

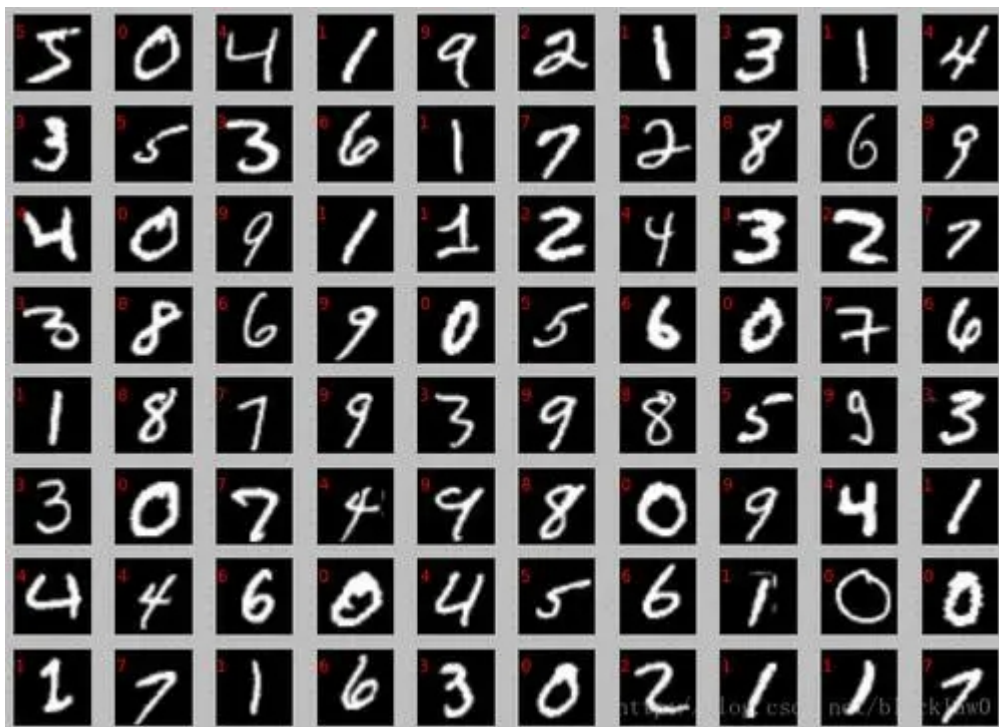
MNIST

MNIST数据集是最常用的基准测试数据集之一。

该数据集由手写数组的灰度图片组成。每张图28*28像素。

MNIST数据集共有60000个训练样本和10000个测试样本。每个样本都有对应的数字标签。

该数据集可以通过PyTorch框架进行下载。



读取MNIST数据集中的图片

直接下载的数据是无法用看图软件打开的。

```
import torchvision

# torchvision.datasets.MNIST("下载路径",train=True表示训练集False表示测试集,download=True表示下载)
# 下载训练集
train_set = torchvision.datasets.MNIST("", train=True, download=True)
# 下载测试集
test_set = torchvision.datasets.MNIST("", train=False, download=True)

print(train_set)
print(test_set)

# 取训练集中的第一个样本，返回图片和标签
image, label = train_set[0]
import matplotlib.pyplot as plt

plt.imshow(image, cmap="gray")
plt.title(label)
plt.show()
```

每一张图片，都是由一个28*28的像素矩阵表示。如图所示，每个像素点保存的是一个灰度(亮度)值

