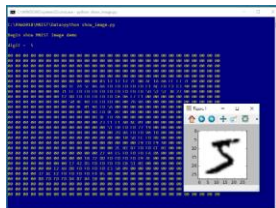


Predictive Analytics World / Deep Learning World

Exercises - The MNIST Dataset

1. Prepare the MNIST image dataset for use by a Keras/TensorFlow program. Locate the zipped or gzipped binary files and extract them. Write a Python program that converts and merges the four binary files into two text files (a training file and a test file). The target format for the two files should be: 5 ** 0 0 152 27 . . 0. Each line of data is one 28x28 image. There are 786 values on each line. The values are delimited by a single space character. The first value is the class label ('0' to '9'). The second value is a dummy ** separator. The next 784 values are the pixel values for the image.

2. Write a utility program that can read and display a specified image. For example:



3. Write a Keras/TensorFlow program to create a neural network classification model for the MNIST image dataset. There are thousands of reasonable design choices but consider two convolution layers, a pooling layer, and using dropout.

4. Which statement about image convolution is most accurate?

- a.) Convolution achieves feature extraction, translation invariance, and parameter reduction.
- b.) Convolution achieves feature extraction and translation invariance, but increases parameters.
- c.) Convolution achieves translation invariance but not feature extraction or parameter reduction.

5. If image size is W , filter size is F , padding is P , stride is S , which statement is most accurate?

- a.) $W + S$ must be greater than $F * P$
- b.) $W + S$ must be less than or equal to $F * P$
- c.) $W - F + 2P$ must be evenly divisible by S

6. Which statement about neural pooling layers is most accurate?

- a.) Pooling reduces parameters but tends to increase model overfitting.
- b.) Pooling reduces parameters and also tends to prevent model overfitting.
- c.) In general, average-pooling tends to work better than max-pooling.

7.) Which statement about deep image classification is most accurate?

- a.) ReLU activation is often used because its Jacobian matrix can be calculated efficiently.
- b.) ReLU activation is often used because it tends to limit the vanishing gradient effect.
- c.) ReLU activation is rarely used because it is not differentiable at all of its domain values.