written4

October 26, 2023

1 Building CNN for MNIST data

1.1 STP598 Machine Learning & Deep Learning

1.2 Written Assignment 4

1.3 Due 11:59pm Sunday 11/05/2023

In the class, we have introduced TensorFlow and built a DNN for MNIST data to classify those hand written digits. In this short exerciese, you are asked to build a CNN for the same classification task.

1.4 Set up TensorFlow

Import TensorFlow into your program to get started:

[1]:
```python
import tensorflow as tf
print("TensorFlow version:", tf.__version__)

import numpy as np
import matplotlib.pyplot as plt
```

```
TensorFlow version: 2.0.0
```

1.5 Load a dataset

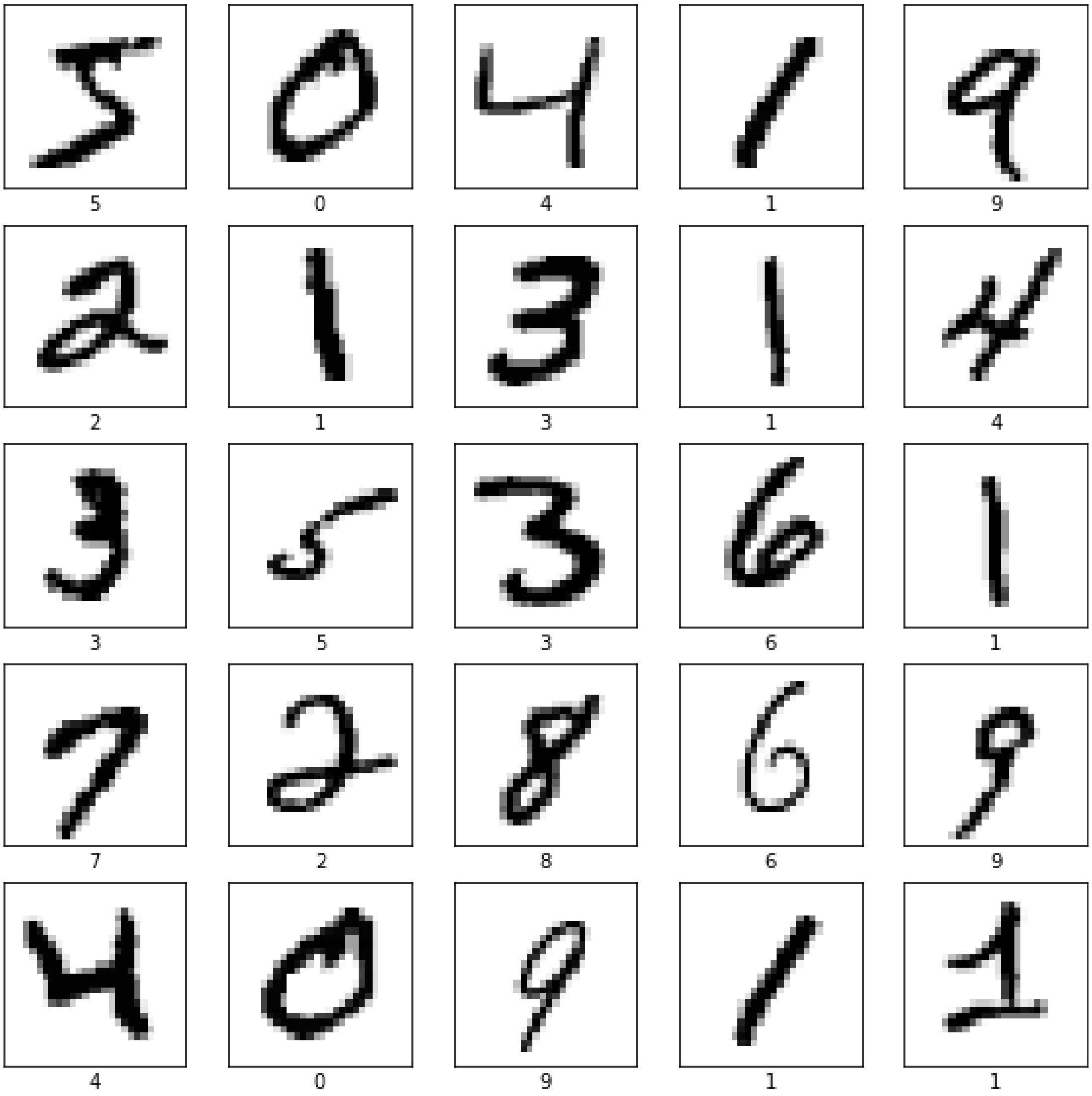
Load and prepare the MNIST dataset. Convert the sample data from integers to floating-point numbers:

[2]:
```python
mnist = tf.keras.datasets.mnist

(x_train, y_train), (x_test, y_test) = mnist.load_data()
x_train, x_test = x_train / 255.0, x_test / 255.0
```

Let's take a peek at these images.

[3]:
```python
plt.figure(figsize=(10,10))
for i in range(25):
    plt.subplot(5,5,i+1)
    plt.xticks([])
    plt.yticks([])
    plt.grid(False)
    plt.imshow(x_train[i], cmap=plt.cm.binary)
    plt.xlabel(y_train[i])
plt.show()
```



2 Build a CNN model

1. First, build a 2-layer CNN `model` by stacking `tf.keras.layers.Cov2D` and `tf.keras.layers.MaxPooling2D` layers, followed by a `tf.keras.layers.Flatten` and `tf.keras.layers.Dense` with output dimension being 10 (classes). Start with `tf.keras.Sequential`.

[ ]:

2. We use `adam` for the optimizer, `tf.keras.losses.SparseCategoricalCrossentropy` (with `from_logits=True`) for the loss function, and `accuracy` as the metrics. Compile the model.

[ ]:

2.1 Train and evaluate your model

3. Before we train the model. It is important to make sure the input size is complient. Use the `Model.fit` method to adjust your model parameters and minimize the loss (you can set `epochs=5`):

[ ]:

4. The `Model.evaluate` method checks the models performance, usually on a “Validation-set” or “Test-set”.

Save testing loss and accuracy output by `evaluate` function in `test_loss` and `test_acc` respectively!

[ ]:

2.2 bonus question

5* Now build a 4-layer CNN model similarly as above. What problem could you run into? How to resolve it? Compare the results of the two models.

[ ]: