EE526 Homework 4

Xingche Guo

November 17, 2019

Problem 1

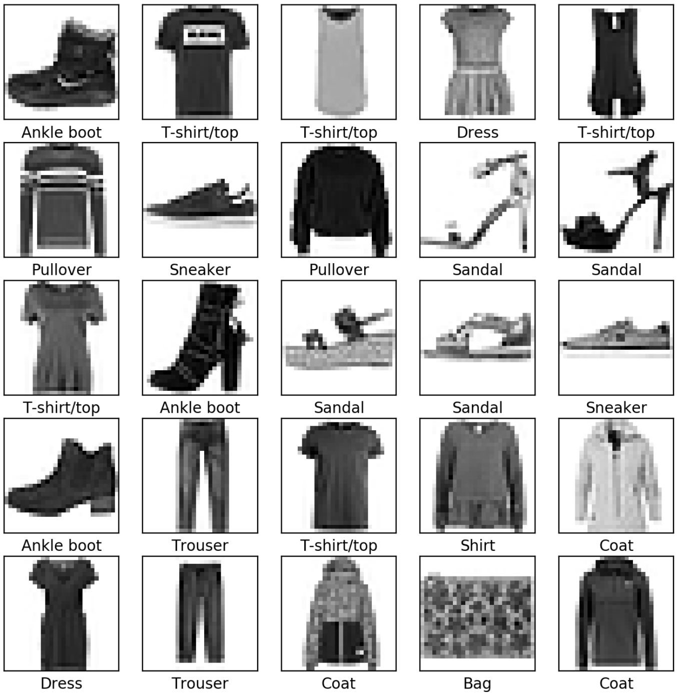

Figure 1

In this problem, CNN is implemented to the Fashion MNIST classification problem. See the visualization of the training data in Figure 1.

(a)

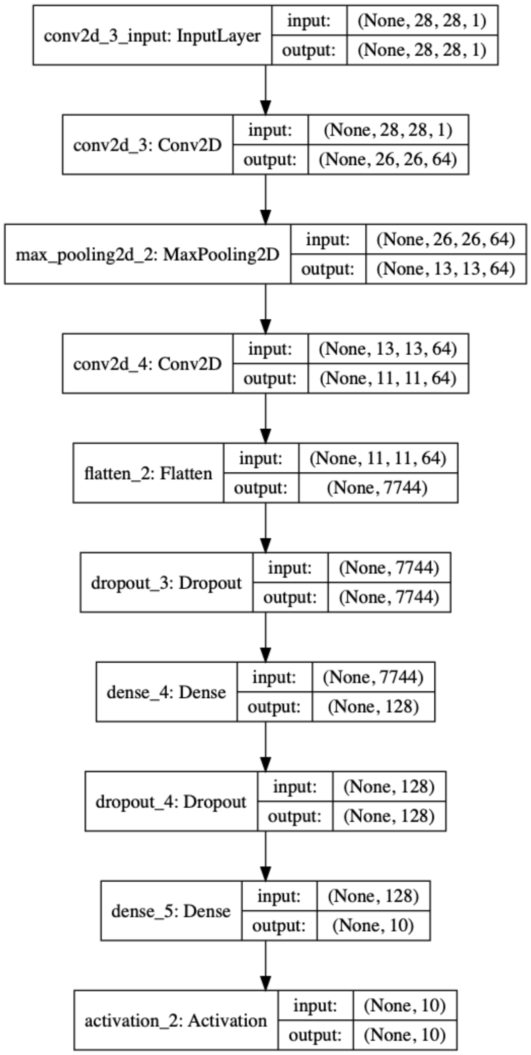

Figure 2

The CNN architecture can be found in Figure 2. Specifically, two convolutional layers and two fully connected layers are used. Batch size is chosen to be 128 and a total of 20 epochs are used. There are:

$$(1\times3\times3\times64+64)+(64\times3\times3\times64+64)+(7744\times128+128)+(128\times10+10) = 640+36928+991360+1290$$

parameters used.

(b)

In each forward pass, a total number of floating point operations needed is:

$$multiplication: \quad (3\times3\times26\times26\times64)+(3\times3\times11\times11\times64)+(7744\times128)+(128\times10) = 1451584$$

$$addition: \quad [(8{+}1)\times26\times26\times64]+[(8{+}1)\times11\times11\times64]+[(7743{+}1)\times128]+[(127{+}1)\times10] = 1451584$$

Thus, a total number of floating point operations needed $\approx 3 \times 10^6$.

(c)

In each mini-batch the total floating point operations needed $\approx 1.5 \times 10^9$.

(d)

The total floating point operations per epoch needed $\approx 7.2 \times 10^{11}$. The total floating point operations over the whole training process needed $\approx 1.4 \times 10^{13}$.

(e)

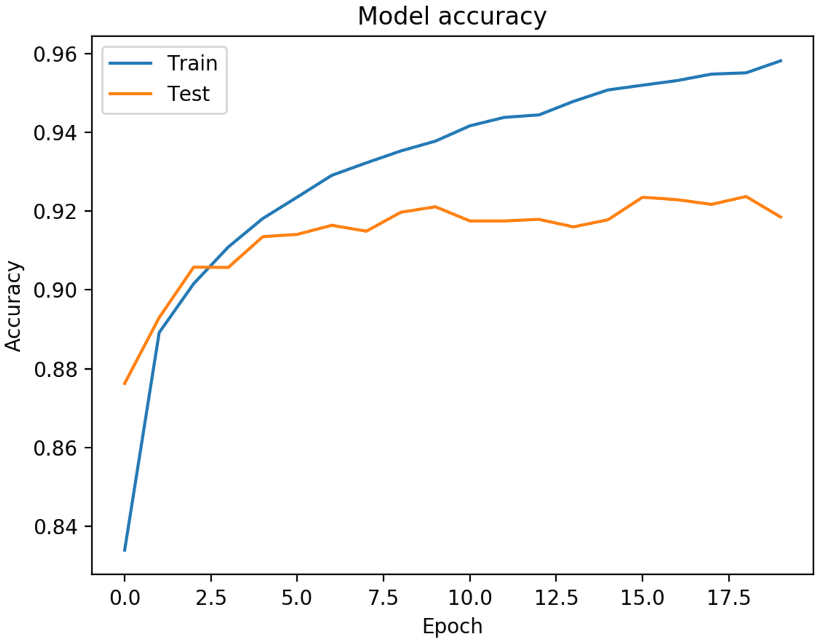

Figure 3

Train accuracy: 97.7%; Test accuracy: 92.5%. A plot of the training accuracy and test accuracy is shown in Figure 3.

Problem 2

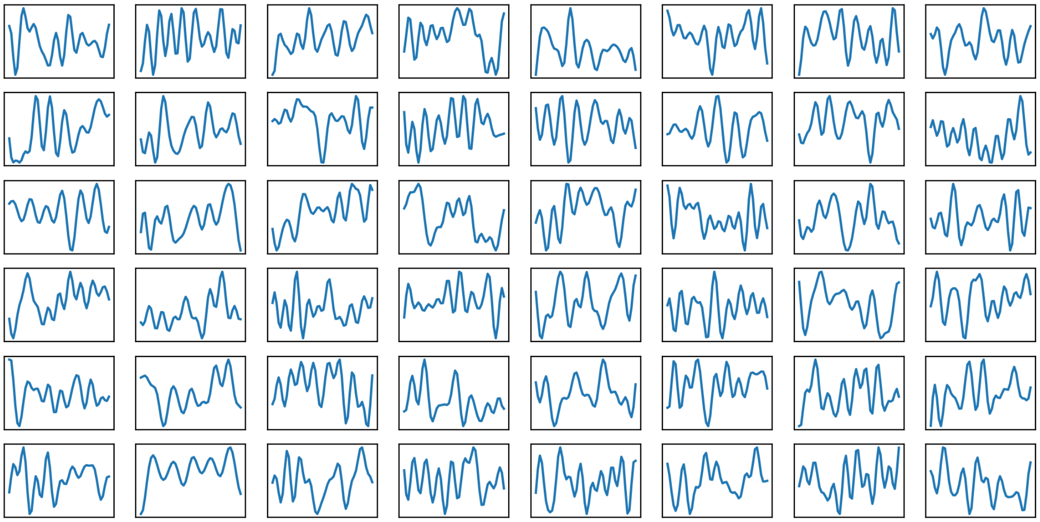

Figure 4

In this problem, RNN is implemented to the time series prediction problem. The dataset contains 100 signals (the network should be trained based on the first 66 signals and tested using the last 34 signals), for each signal, there are 500 (equally spaced) measurements. See the visualization of 48 randomly chosen signals (first 50 measurements) in Figure 4. We regard the the $1:499$ measurements as X, and the $2:500$ measurements as Y, and apply the RNN model with 3 recurrent layers and one final non-recurrent layer, the details can be found in Figure 5. We let the number of epochs to be $5,000$, the training and testing loss for each epochs are plot in Figure 6. We also compared the true Y_{test} to the prediction of Y_{test}, the prediction results can be found in Figure 7.

One strange thing is that train loss is larger than the testing loss, I guess the problem is that we do not run enough iterations. I also tried to hold out more time measurements (say, predict the last 5 measurements instead of the last one measurement). However, it requires a more complex model and the RNN will converge much slower.

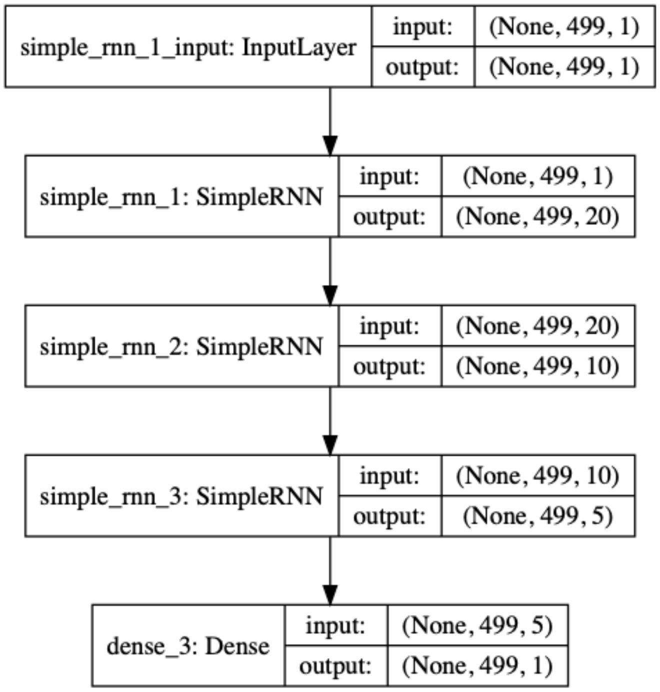

Figure 5

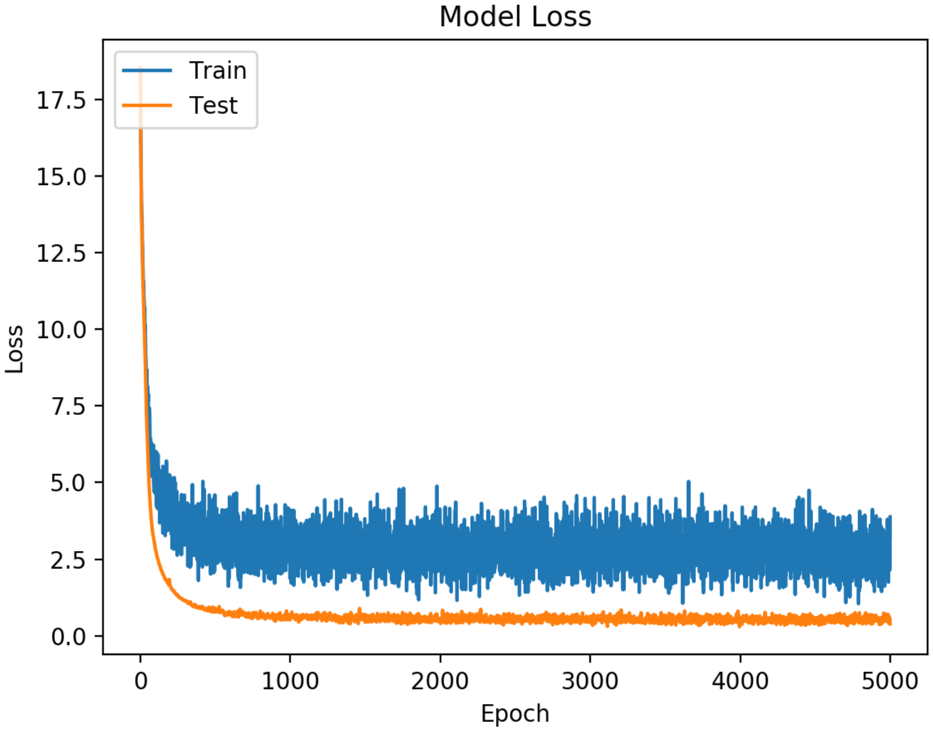

Figure 6

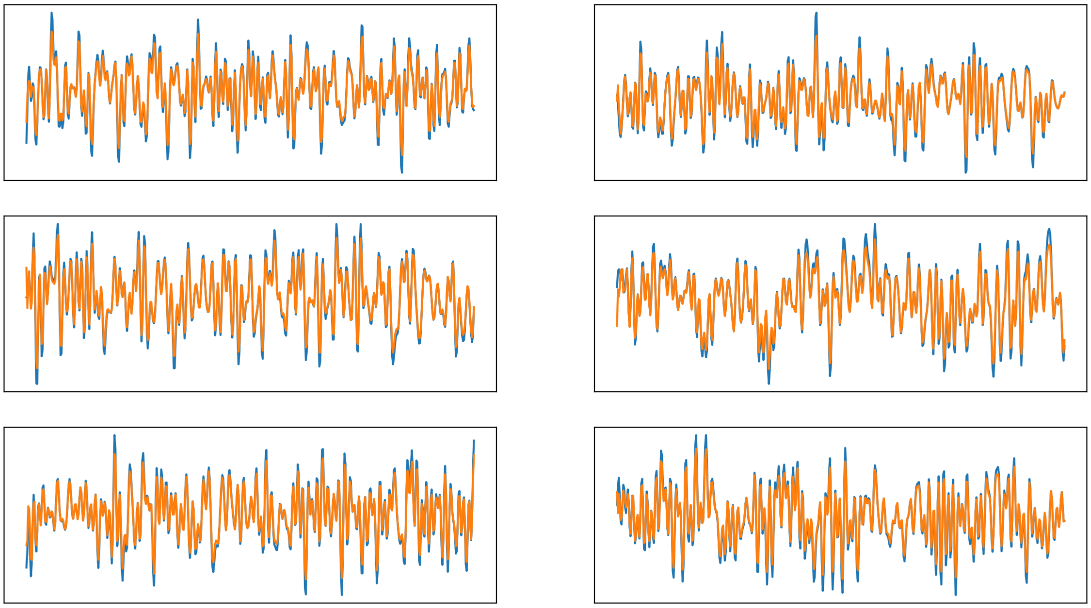

Figure 7