Artificial Inteligent

2013042776
컴퓨터공학과
남궁 선

1. Introduction

Backpropagation technique was used. Internally, Normalization, Minibatch, Affline, ReLU, Softmax, and Cross Entropy Error techniques were used. The accuracy of the training set was 97% and the accuracy of the test set was 97%. No overfitting occurred.

2. Archiitecture and Order

1. Normalization
2. Minibatch
3. Affline
4. ReLU
5. Softmax
6. Cross Entropy Error
7. Backpropagation

I have divided the files by function. Execution commands are the same as before.

3. Normalization

When retrieving data, normalization improves data throughput and data identification capabilities. Since the range of pixels is 255, it is divided by 255.

4. Minibatch

Because the library is optimized for batch calculations, the data is grouped together and processed to reduce the load on the bus. As a result, the processing speed is improved. Learning all the data is very slow and there are lots of difficulties in real life, so I used the Minibatch technique to extract data randomly. In this project I choose 100 samples randomly.

5. Affilne

The result of multiplying the weight with the data is passed to the ReLU layer.

6. ReLU (Activation function)

It is a nonlinear function, it is simpler than Sigmoid, and I used it as a function that is used frequently in recent years. It is efficient because it is simple to implement and fast.

7. Softmax

Each factor of the softmax function is the probability that the corresponding index will occur. This can be combined with argmax to effectively yield the highest probability result. Entire summation is 1.

8. Cross Entropy Error

Loss function. Combining CEE with one-hot-encoding will result in only the index corresponding to the correct answer. The output at this time becomes the reference value, and there is no unnecessary additional calculation.

9. Backpropagation

It is a technique to update the weight by transferring the error of the result and the correct answer to the previous step. Compared to general neural networks, the speed is dramatically improved and accuracy is also improved. Also, if you have back propagation characteristics of each function, it is a very good technique if you learn well because it is easy to implement backward as well as forward.

10. Hyperparameter - Learning rate.

The learning rate was set at 0.1. It is also related to the number of iterations. However, if the learning rate is too low, the learning is not done properly within the limit range. If the learning rate is too high, the weighting error is large. Even if it was 0.01, it was not learned properly when it was done 10000 times. If it exceeded 0.2, accuracy is overfitting about 98%

`0.87235` `0.9816`

11. Hyperparameter - iterate

10000. Even if the repetition rate exceeded 15000 times with the refresh rate of 0.1, the accuracy of the training set became 98% and the overfitting with the error of the test set occurred. In the case of 20,000 times, it increased to 98.5% and overfitting occurred.

`0.979533333333` `0.985783333333`

12. Execution screen

13. Feedback

I originally implemented CNN. However, the project was abandoned because I had just followed the code without understanding it properly. The power of CNN was enormous. CNN easily exceeded the accuracy of 98%. I will have to study after that. I think I know the professor last lectured the neural network lecture. Perhaps because CNN is so powerful and attractive. Then professor worried that students use neural networks without the basics. I tried hard to try new things besides what I already learned in the lecture. As a result, ReLU was introduced and CEE was introduced. The more various functions according to purpose, the better the performance and accuracy of machine learning can be improved.

Thank you for this semester.