Multi-Layer Perceptron

Zhuoran Liu

January 30, 2017

1 Introduction

In this report, we will consider the neural network multi-layer perceptron. Multi-layer perceptron is a neural network with multiple hidden layers and in each hidden layer there are multiple neurons. Between different neurons in adjacent layers, there exist weights to connect different neurons. In every hidden layer, there exist biases to tune some outcomes of neurons. The aim of training the network is to find the proper weights and biases which can be used to predict the new data.

1.1 Underlying Theory

Firstly, we will talk about the working mechanism of the MLP. Given the neural network structure(weights matrix W and biases vector b) and input data vector a, we can calculate the feed forward process by formula

$$\mathbf{z} = g(\mathbf{w}a + \mathbf{b})$$

Here g is the activation function and output z is a vector. Use z as the input of the next layer, we can do similar calculation again until we get the output. Given the last output z, we will calculate the $argmax(z)$. The category of the $argmax(z)$ is the predication of the input a. Known the right category, we will tune the weights and biases to predict better and better. This is the learning process of MLP. Given the input

1.2 Learning Algorithm

Backpropagation algorithm is the learning algorithm we will use. It consists of 4 steps.

$$\delta_j^L = \frac{\partial C}{\partial a_j{}^L} \sigma'(z_j^L)$$

2 Problem statement

3 Results

3.1 The structure of Multi-Layer Perceptron

3.1.1 Different layers

3.1.2 Different neurons

3.2 The initialization of the weights and biases

3.2.1 tanh distribution initialization

3.2.2 Boltzmann machine initialization

3.3 Different learning methods

3.3.1 Stochastic gradient descent

3.3.2 Momentum learning

4 Discussion

5 Conclusion

6   Appendix