1.5em 0pt

Chapter 5

Adam p19@ipm.edu.mo
20200429

Keywords: *ReLU function; overfitting; Dropout; Vanishing Gradient*

1. Summary

- Deep Learning can be simply defined as a Machine Learning technique that employs the deep neural network.
- The previous neural networks had a problem where the deeper (more) hidden layers were harder to train and degraded the performance. Deep Learning solved this problem.
- The outstanding achievements of Deep Learning were not made by a critical technique but rather are due to many minor improvements.
- The poor performance of the deep neural network is due to the failure of proper training. There are three major showstoppers: the vanishing gradient, overfitting, and computational load.Phil Kim (2017)
- The vanishing gradient problem is greatly improved by employing the ReLU activation function and the cross entropy-driven learning rule. Use of the advanced gradient descent method is also beneficial.
- The deep neural network is more vulnerable to overfitting. Deep Learning solves this problem using the dropout or regularization.Microstrong (2018)
- The significant training time is required due to the heavy calculations. This is relieved to a large extent by the GPU and various algorithms.

Each hidden layer contains 20 nodes. The network has 25 input nodes for the matrix input and five output nodes for the five classes. The output nodes employ the softmax activation function.

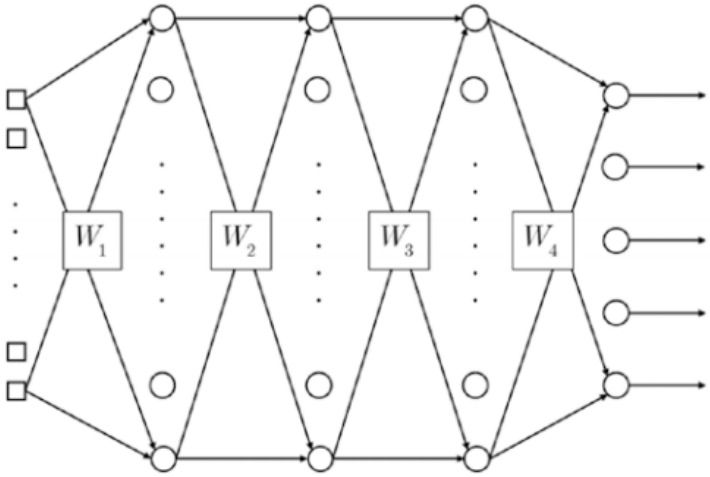

Figure 1: The Deep Neural Network with Three Hidden Layers

References

Microstrong. dropout. `https://zhuanlan.zhihu.com/p/38200980`, 2018.

Phil Phil Kim, Kim. *MATLAB Deep Learning*. Apress, 2017.