

A Preliminary Study on Deep Learning in Data Analysis

Advisor: Dr. rer. nat. Chuan-Kang Ting Name: Liang-Yu Chiu

Introduction

•What is Deep Learning?

Deep learning is the research on learning models with multilayer representations. The architecture has the features following. It has many sub-functions to solve the problem by dividing the problem into many sub-problems and be solve by these sub-functions. Then, solve the whole problem by using the answer from the sub-functions.

The architecture is often represented by the multi-layer neural networks because of the motivation from visual cortex.

•Motivation

The motivation of deep learning is from visual cortex. If we see a picture of human face, we might see the points, the lines, and colors, etc. Then the thing will change into the useful information likes eyes, nose, mouth, etc. Finally, we know that it is a human face.

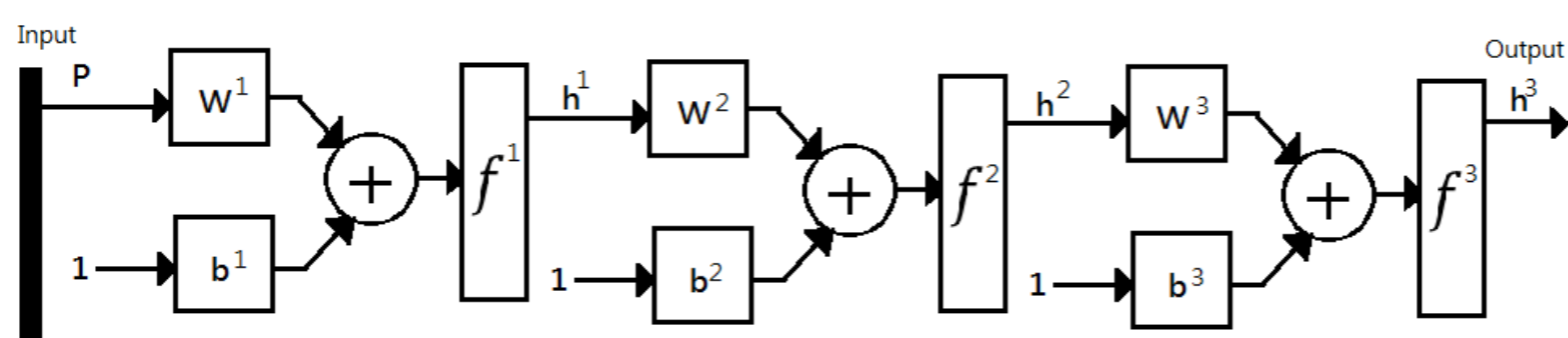
Neural Networks for Deep Architecture

•Universality Theorems

It is well known by people who use neural networks. The theorems is that neural can compute any problem. If adding enough units and layers, it can solve more complex problems.

• Multi-Layer Neural Networks

The figure following shows the multi-layer neural networks with two hidden layer



and can write by the equations

$$\mathbf{h}^k = f^k(\mathbf{b}^k + \mathbf{W}^k \mathbf{h}^{k-1})$$

where \mathbf{W} is the matrix of weights, \mathbf{b} is the vector of offsets, \mathbf{h} is the output of early layer, k is the number of layer, f is the activation function, and \mathbf{h}^0 is the input \mathbf{p} .

•Something hard...

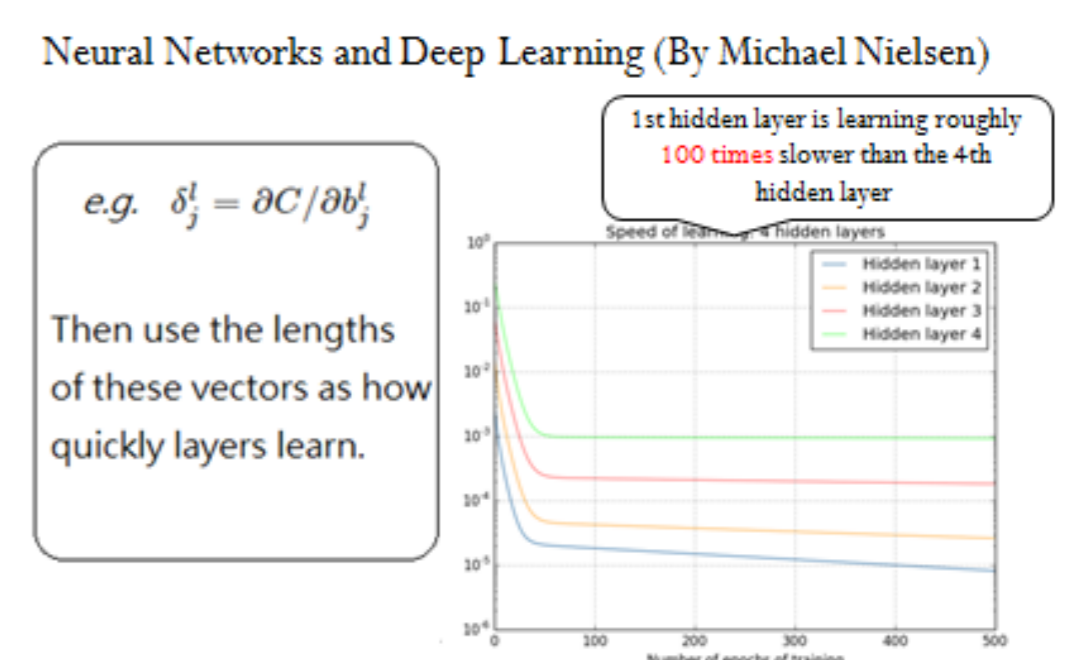
If using the back propagation to train the deep architecture with random initialization, the result is poor. The reason might be that random

initialization leave more useless information into later layer.

Another problem named **vanishing gradient problem**. It is mean that when the early layer converges, the later layer does not. The problem would happen by the very different learning speed showing by the figure following.

Because of these problems, many different architectures are designed to achieve better quality likes

convolutional networks, deep belief network, auto-encoder, and others.



Experiment

Trying to train by MNIST handwritten digit data set and the code in the Internet has good quality, so I try to train other different data set. However, the result has large test error showing by the table.

architecture\data	MNIST	Credit	Arcene
MLP	1.65%	34.29%	56%
CNN	0.92%	34.29%	56%
DBN	1.17%	34.29%	56%

Then, it is strange that other data with same test error by different architecture. The reason could be that the predictions are same (other datas are binary classification).

Conclusion

Deep learning is a powerful way to learn from data, but if the data set is not the scale of 'deep', the performance might not be good.

Reference

- [1] Neural Networks and Deep Learning (Michael Nielsen / Aug 2015)
- [2] Learning Deep Architectures for AI (Yoshua Bengio Dept. IRO, Universit'e de Montr'eal C.P. 6128, Montreal, Qc, H3C 3J7, Canada)