1-25

Deep Learning Using Noise for Back Propagation

Kazuki NAGAO[1], Shinsaburo KITTAKA[1], Yoko UWATE[1] and Yoshifumi NISHIO[1]
([1] Tokushima University)

1. Introduction

Deep Learning from 2012 has been focused again. Because deep learning automatically learns valid features of discrimination. Recently, it is used in image recognition, module of self-driving and so on.

In this study, new system of deep learning is proposed. This system is noisy feed back. Purposes of this study are to improve learning accuracy and to reduce learning loops. Conventional system and proposed system are compared.

2. Proposed method

Figure 1 shows proposed and conventional system. Arrows of left to right indicate propagation of input signal and right to left indicate learning of back propagation. We use Gaussian distribution (0, 0.01) for noise and add it to equations (2). In addition, the putting place of Gaussian distribution is changed. We experiment 8 patterns about conventional system without noise, (h_{1back}), (h_{2back}), (y_{back}), (h_{1back}, h_{2back}), (h_{1back}, y_{back}), (h_{2back}, y_{back}) and (h_{1back}, h_{2back}, y_{back}).

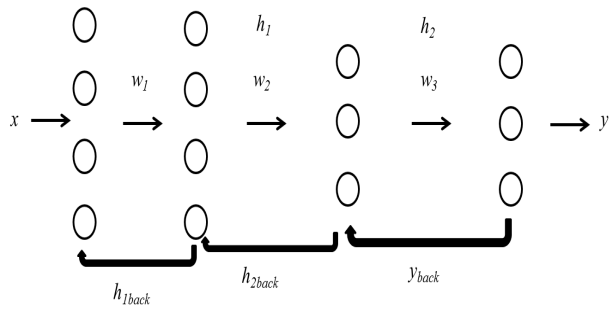

Figure 1: Schematic diagram of deep learning.

$$\begin{cases} h_1 = \dfrac{1}{1+e^{(-\sum w_1 x)}} \\ h_2 = \dfrac{1}{1+e^{(-\sum w_2 h_1)}} \\ y = \dfrac{1}{1+e^{(-\sum (w_3 h_2))}} \end{cases} \tag{1}$$

$$\begin{cases} y_{back} = (y-ty)(1-y)y \\ h_{2back} = \sum w_3 y_{back}(1-h_2)h_2 \\ h_{1back} = \sum w_2 h_{2back}(1-h_1)h_1 \end{cases} \tag{2}$$

$$\begin{cases} {w_1}^{(l+1)} = {w_1}^{l} - \varepsilon\ x h_{1back} \\ {w_2}^{(l+1)} = {w_2}^{l} - \varepsilon\ h_1 h_{2back} \\ {w_3}^{(l+1)} = {w_3}^{l} - \varepsilon\ h_2 y_{back} \end{cases} \tag{3}$$

ε is the rate of decay which is 0.01. h means propagation. w means weights between neurons. The initial value of the weights is chosen by random from 0.01 to 0.04.

3. Simulation result

We define as the learning loops = 25000 and the number of learning data sets = 150.

Figure 2 shows the learning accuracy and learning loops of conventional system. Figure 3 shows the learning accuracy and learning loops of the proposed system. Table 1 shows learning accuracy for all kinds of patterns. Table 2 shows learning loops when learning accuracy is 0.15 or less.

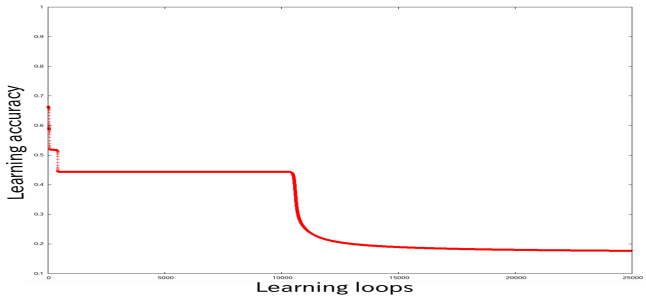

Figure 2: Conventional system.

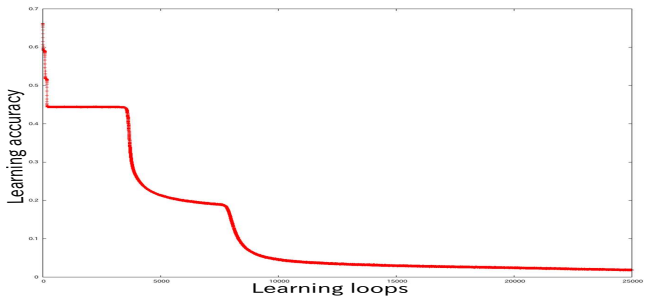

Figure 3: Proposed system.

Table 1: Learning accuracy.

	accuracy
Conventional	0.1771
h_{1back}	0.1787
h_{2back}	0.1727
y_{back}	0.0269
$h_{1back} + h_{2back}$	0.1750
$h_{1back} + y_{back}$	0.0178
$h_{2back} + y_{back}$	0.1708
$h_{1back} + h_{2back} + y_{back}$	0.0261

Table 2: Learning loops.

	loops
y_{back}	12612
$h_{1back} + y_{back}$	8002
$h_{1back} + h_{2back} + y_{back}$	11103

From Figs. 2 and 3, the learning accuracy and learning loops of using Gaussian distribution are better than not using it.

4. Conclusion

We confirmed that it is good to use noise at back propagation of the output layer to final intermediate layer. Because back propagation of the output layer is compared with inputs and outputs, so it is properly carried out weight adjustment.

In the future, we use a variety of noise and want to obtain better results.