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 conditions (2). In addition, the putting place of Gaussian distribution is changed. We experiment 8 patterns about conventional system without noise, \((h_{\text{1back}}, h_{\text{2back}}, y_{\text{back}}), (h_{\text{1back}}, h_{\text{2back}}, y_{\text{back}})\) and \((h_{\text{1back}}, h_{\text{2back}}, y_{\text{back}})\).

\[
\begin{align*}
  h_1 & = \frac{1}{1 + e^{-\left(\sum w_1 x\right)}} \\
  h_2 & = \frac{1}{1 + e^{-\left(\sum w_2 h_1\right)}} \\
  y & = \frac{1}{1 + e^{-\left(\sum w_y h_2\right)}} \\
  y_{\text{back}} & = (y - ty)(1 - y)y \\
  h_{\text{2back}} & = \sum w_3 y_{\text{back}}(1 - h_2)h_2 \\
  h_{\text{1back}} & = \sum w_2 h_{\text{2back}}(1 - h_1)h_1 \\
  w_1^{(i+1)} & = w_1^i - \varepsilon \times h_{\text{1back}} \\
  w_2^{(i+1)} & = w_2^i - \varepsilon \times h_{\text{2back}} \\
  w_3^{(i+1)} & = w_3^i - \varepsilon \times h_{\text{2back}}
\end{align*}
\]

\(i\) 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.

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.