CCNU-UOW
CSCI964 Computational Intelligence
Spring 2020

Laboratory Exercise 2 (Week 2)

1 Task One: Delta Rule for Artificial Neural Networks Training

The delta rule is expressed as the following equation

$$w_i \leftarrow w_i - \eta \frac{dE}{dw} \tag{1}$$

In which

$$p = \sum_{i=1}^{n} x_i w_i - \theta \tag{2}$$

$$\widehat{y} = f(p) \tag{3}$$

$$f(p) = \frac{1}{1 + e^{-p}} \tag{4}$$

$$E = \frac{1}{2}(y - \widehat{y})^2 \tag{5}$$

So we can get the following quation by the chain rule

$$\frac{dE}{dw_i} = \frac{dE}{d\widehat{y}} \frac{d\widehat{y}}{dp} \frac{dp}{dw_i} \tag{6}$$

Where

$$\frac{dE}{d\hat{y}} = (\hat{y} - y) \tag{7}$$

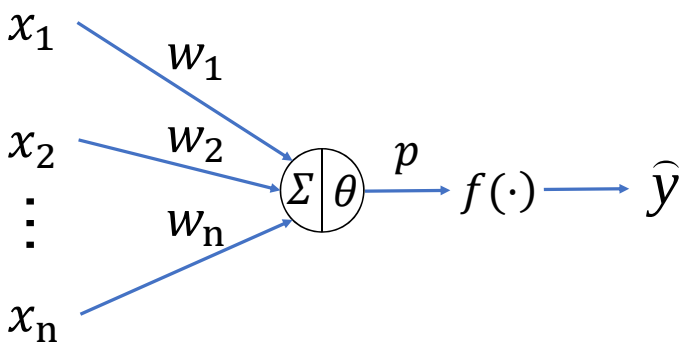

Figure 1: Perceptron with activation function

Algorithm 1 Delta Rule for learning single ANN

Input: training dataset $D = \{(\boldsymbol{x}_k, \boldsymbol{y}_k)\}_{k=1}^{m}$;
learning rate η
randomly initialize w_i and θ

Process:

1. **repeat**
2. for all $(\boldsymbol{x}_k, \boldsymbol{y}_k)$ do
3. compute $\widehat{y}$ through the forward process
4. update w_i through $w_i \leftarrow w_i - \eta(\widehat{y} - y)\widehat{y}(1 - \widehat{y})x_i$
5. update θ through $\theta \leftarrow \theta + \eta(\widehat{y} - y)\widehat{y}(1 - \widehat{y})$
6. end for
7. **until** $|E_{t+1} - E_t| \leq \varepsilon$ **or** $t < T$

Output:

$$\begin{aligned} \frac{d\widehat{y}}{dp} &= f(p)(1 - f(p)) \\ &= \widehat{y}(1 - \hat{y}) \end{aligned} \tag{8}$$

$$\begin{aligned} \frac{d\widehat{y}}{dp} &= f(p)(1 - f(p)) \\ &= \widehat{y}(1 - \hat{y}) \end{aligned} \tag{9}$$

So put (7)∼(9) into (6), we have

$$\begin{aligned} \frac{dE}{dw_i} &= \frac{dE}{d\hat{y}}\frac{d\hat{y}}{dp}\frac{dp}{dw_i} \\ &= (\hat{y} - y)\hat{y}(1 - \hat{y})x_i \end{aligned} \tag{10}$$

So for $w_i$updating we have

$$w_i \leftarrow w_i - \eta(\widehat{y} - y)\widehat{y}(1 - \widehat{y})x_i \tag{11}$$

In the same way, we can learn θ

$$\begin{aligned} \frac{dE}{d\theta} &= \frac{dE}{dy}\frac{d\widehat{y}}{dp}\frac{dp}{d\theta} \\ &= (\hat{y} - y)\hat{y}(1 - \hat{y})(-1) \end{aligned} \tag{12}$$

$$\theta \leftarrow \theta + \eta(\widehat{y} - y)\widehat{y}(1 - \widehat{y}) \tag{13}$$

Using C++ and delta rule to train the following ANNs by online learning
With the following input and output

$$\begin{array}{rcccc} x_1 & & 0 & 0 & 1 & 1 \\ x_2 & = & 0 & 1 & 0 & 1 \\ x_3 & & 1 & 1 & 1 & 1 \end{array} \tag{14}$$

$$y = \begin{array}{cccc} 0 & 0 & 1 & 1 \end{array} \tag{15}$$

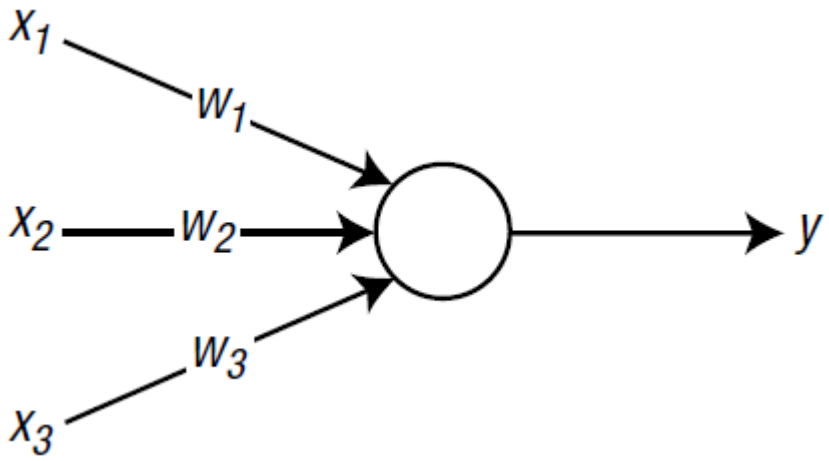

Figure 2: ANNS with 3 inputs

2 Task Two: Implementation of the Batch Method

Using batch method to train the above ANN and coding with C++.