**Machine Learning & Data Mining, Spring 2020**

**Homework 7**

Due May 14

**이름: 김도운**

**학과: 사이버보안학과**

**학번: 201520893**

1. **Build a neural network that has two binary inputs x1 and x2, and output Y, where Y implements the OR function: Y=(x1 OR x2). Assume that the activation function is a threshold function such that h(a)=1 if a >=0, and 0 otherwise.**

* WT=
* a =

1

w0

h(a)

a

w1

X1

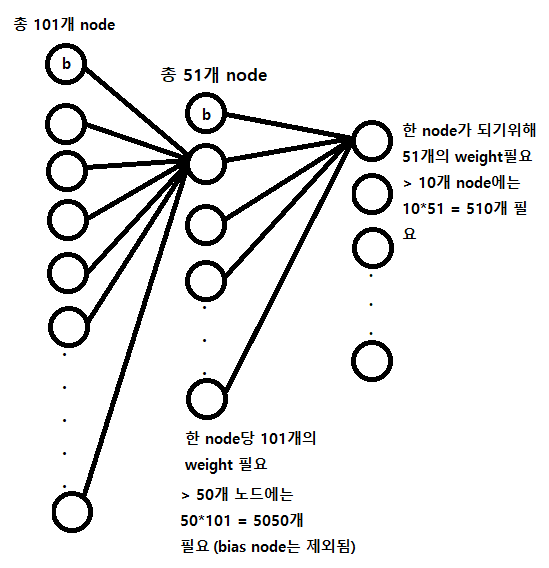
w2

X2

* 예를 들어, (x1, x2)쌍이 (0,0) 이면, a = -0.5이므로 h(a) < 0이 되고, 0으로 분류하고, (0,1) 또는 (1,0)인 경우 h(a) = 0.5 >0 이므로 1로 분류하게 되며, (1,1)일 때에는 h(a)=1.5 >0이므로 1로 분류하게 된다. OR 연산에 대해 제대로 분류하도록 weight값들이 잘 설정되었음을 알 수 있다.

1. **Suppose a neural network has 100 input features, one hidden layer with 50 hidden nodes, and an output layer with 10 output nodes. How many weight parameters (including bias) should be learned?**

* 첫번째 layer에서 50개의 node로 이루어진 one hidden layer의 각 node로 계산되기위해 101개(bias 포함)의 weight가 필요하므로, 50\*101 = 5050개, 다음으로 hidden layer에서 10개의 output node의 각 node로 계산되기 위해 51개의 weight가 필요하므로 이 때 51\*10 = 510개가 학습되어야 한다. 따라서 총 5050+510=5560개의 weight값들의 학습이 필요하다.



1. **Discuss how Convolutional neural networks differ from the traditional neural networks in terms of the convolution operations.**

* Fully connected layer만으로 구성된 인공 신경망의 입력 값은 Affine즉 행렬 곱셈 값이며, convolutional layer와 pooling layer(optional)로 구성된 CNN의 경우 1차원 뿐만 아니라 N차원의 입력 값에 대해서도 적용 가능하다. 즉 이미지나 동영상 데이터 등의 처리를 위한 인공신경망이라고 할 수 있다.
* Convolution layer에서의 연산은 Input으로 주어진 입력 데이터(이미지)의 local region마다 filter라는 파라미터와의 dot product를 수행하며, 이 때 local region과 내적을 수행하는 filter값은 공유되므로 traditional NN에서 많은 weight 파라미터들이 필요한 것과 다르게 파라미터 수가 현저하게 줄어들게 된다.