

Homework #5

1. Sigmoid function은 출력값이 $[0, 1]$ 사이의 값을 갖게 한다.

그러므로 출력이 원의 a, b $[a, b]$ 사이의 값을 가질 경우, $[0, 1]$ 사이의 값을

갖게끔 normalization을 거쳐야 sigmoid function을 사용하는 NN을 이용하여

regression 문제를 해결할 수 있다.

normalization을 하는 방법으로는 출력값 t 가 $a \leq t \leq b$ 일 때, $0 \leq \frac{t-a}{b-a} \leq 1$ 로 변형하는 것을 통하여

출력값이 $[0, 1]$ 사이의 값을 갖게 하는 방법이 있다.

단, sigmoid function을 activation function으로 사용하는 NN은 출력이 0 또는 1일 때

gradient가 0이 되므로 학습을 할 수가 없다.

2. Target값: 1

cross-entropy값: $-\{1 \times \log(0.7) + (1-1) \log(1-0.7)\} = -\log(0.7) = 0.15490195999$

$\therefore 0.15490195999$

\Rightarrow 수업을 들으면서 보완한 부분

cross-entropy : NN의 학습이 잘 이루어지지 않았을 확률

$-(t \log(y) + (t-1) \log(1-y))$: minimize 해야 NN이 학습을 잘할 확률이 maximize 됨.

3. $-\sum_{k=1}^{20} t_k \log(y_k) = -\{1 \times \log(\frac{1}{20})\} = 1.30102999566$

$\therefore 1.30102999566$

4. * ipynb 파일 참조

T	x_1	x_2	x_3	σ_1	σ_2	σ_3
0.01	6	3	2	1	0	0
1	6	3	2	0.9	0.05	0.02
10	6	3	2	0.4	0.3	0.2

5. *ipynb 파일참조

i) $(x_1, x_2) = (1, 0)$ 일때,

$$net_1 = W_{11}x_1 + W_{21}x_2 + W_{31} \cdot 1 = W_{11} + W_{31} = 0.6$$

$$net_2 = W_{12}x_1 + W_{22}x_2 + W_{32} \cdot 1 = W_{12} + W_{32} = 0.6$$

$$h_1 = \text{sigmoid}(0.6) = 0.6457$$

$$h_2 = \text{sigmoid}(0.6) = 0.6457$$

$$\text{Softmax Layer 첫번째노드로의 입력값} = V_{11}h_1 + V_{21}h_2 + V_{31} \cdot 1 = 1.0165$$

$$\text{두번째노드로의 입력값} = V_{21}h_1 + V_{22}h_2 + V_{32} \cdot 1 = 1.5457$$

$$\text{세번째노드로의 입력값} = V_{31}h_1 + V_{32}h_2 + V_{33} \cdot 1 = 0.4874$$

$$y_1 = \frac{\exp(1.0165)}{\exp(1.0165) + \exp(1.5457) + \exp(0.4874)} = 0.3043$$

$$y_2 = \frac{\exp(1.5457)}{\exp(1.0165) + \exp(1.5457) + \exp(0.4874)} = 0.5165$$

$$y_3 = \frac{\exp(0.4874)}{\exp(1.0165) + \exp(1.5457) + \exp(0.4874)} = 0.1792$$

ii) $(x_1, x_2) = (0, 1)$ 일때,

$$net_1 = W_{11}x_1 + W_{21}x_2 + W_{31} \cdot 1 = W_{21} + W_{31} = 0.9$$

$$net_2 = W_{12}x_1 + W_{22}x_2 + W_{32} \cdot 1 = W_{22} + W_{32} = 1$$

$$h_1 = \text{sigmoid}(0.9) = 0.7109$$

$$h_2 = \text{sigmoid}(1) = 0.7311$$

$$\text{Softmax Layer 첫번째노드로의 입력값} = V_{11}h_1 + V_{21}h_2 + V_{31} \cdot 1 = 1.0768$$

$$\text{두번째노드로의 입력값} = V_{21}h_1 + V_{22}h_2 + V_{32} \cdot 1 = 1.625$$

$$\text{세번째노드로의 입력값} = V_{31}h_1 + V_{32}h_2 + V_{33} \cdot 1 = 0.5286$$

$$y_1 = \frac{\exp(1.0768)}{\exp(1.0768) + \exp(1.625) + \exp(0.5286)} = 0.3023$$

$$y_2 = \frac{\exp(1.625)}{\exp(1.0768) + \exp(1.625) + \exp(0.5286)} = 0.5230$$

$$y_3 = \frac{\exp(0.5286)}{\exp(1.0768) + \exp(1.625) + \exp(0.5286)} = 0.1747$$

다음페이지 계속 >>

iii) $(x_1, x_2) = (1, 1)$ 일 때,

$$net_1 = W_{11}x_1 + W_{21}x_2 + W_{31} \cdot 1 = W_{11} + W_{21} + W_{31} = 1$$

$$net_2 = W_{12}x_1 + W_{22}x_2 + W_{32} \cdot 1 = W_{12} + W_{22} + W_{32} = 1.2$$

$$h_1 = \text{sigmoid}(1) = 0.7311$$

$$h_2 = \text{sigmoid}(1.2) = 0.7685$$

Softmax Layer 첫번째노드로의 입력값 $= V_{11}h_1 + V_{21}h_2 + V_{31} \cdot 1 = 1.0998$

두번째노드로의 입력값 $= V_{21}h_1 + V_{22}h_2 + V_{32} \cdot 1 = 1.6573$

세번째노드로의 입력값 $= V_{31}h_1 + V_{32}h_2 + V_{33} \cdot 1 = 0.5424$

$$y_1 = \frac{\exp(1.0998)}{\exp(1.0998) + \exp(1.6573) + \exp(0.5424)} = 0.3013$$

$$y_2 = \frac{\exp(1.6573)}{\exp(1.0998) + \exp(1.6573) + \exp(0.5424)} = 0.5261$$

$$y_3 = \frac{\exp(0.5424)}{\exp(1.0998) + \exp(1.6573) + \exp(0.5424)} = 0.1725$$

x_1	x_2	net_1	net_2	h_1	h_2	y_1	y_2	y_3
1	0	0.6	0.6	0.6457	0.6457	0.3043	0.5165	0.1792
0	1	0.9	1	0.7109	0.7311	0.3023	0.523	0.1747
1	1	1	1.2	0.7311	0.7685	0.303	0.5261	0.1725