Machine Learning Homework 7

Hairui Yin

1.

Given input x_1, x_2, the hidden units is given by

$$\begin{cases} a_1^{(2)} = g(20x_1 + 20x_2 - 30) \\ a_2^{(2)} = g(-20x_1 - 20x_2 + 10) \end{cases}$$

where g is the sigmoid activation functions. And the final output $h_\theta(x)$ is given by

$$h_\theta(x) = g(20a_1^{(2)} + 20a_2^{(2)} - 10)$$

To verify the neural network, we input all possible x_1, x_2:

1. $x_1 = 0, x_2 = 0$

$$\begin{cases} a_1^{(2)} = g(0 + 0 - 30) = 9.357622968839299 \times 10^{-14} \\ a_2^{(2)} = g(0 - 0 + 10) = 0.9999546021312976 \end{cases}$$

$$\begin{aligned} h_\theta(x) &= g(20a_1^{(2)} + 20a_2^{(2)} - 10) \\ &= 0.9999545608951235 \\ &\approx 1 \end{aligned}$$

2. $x_1 = 0, x_2 = 1$

$$\begin{cases} a_1^{(2)} = g(0 + 20 - 30) = 4.5397868702434395 \times 10^{-5} \\ a_2^{(2)} = g(0 - 20 + 10) = 4.5397868702434395 \times 10^{-5} \end{cases}$$

$$\begin{aligned} h_\theta(x) &= g(20a_1^{(2)} + 20a_2^{(2)} - 10) \\ &= 4.548037850511231 \times 10^{-5} \\ &\approx 0 \end{aligned}$$

3. $x_1 = 1, x_2 = 0$

$$\begin{cases} a_1^{(2)} = g(20 + 0 - 30) = 4.5397868702434395 \times 10^{-5} \\ a_2^{(2)} = g(-20 - 0 + 10) = 4.5397868702434395 \times 10^{-5} \end{cases}$$

$$\begin{aligned} h_\theta(x) &= g(20a_1^{(2)} + 20a_2^{(2)} - 10) \\ &= 4.548037850511231 \times 10^{-5} \\ &\approx 0 \end{aligned}$$

4. $x_1 = 1, x_2 = 1$

$$\begin{cases} a_1^{(2)} = g(20 + 20 - 30) = 0.9999546021312976 \\ a_2^{(2)} = g(-20 - 20 + 10) = 9.357622968839299 \times 10^{-14} \end{cases}$$

$$\begin{aligned} h_\theta(x) &= g(20a_1^{(2)} + 20a_2^{(2)} - 10) \\ &= 0.9999545608951235 \\ &\approx 1 \end{aligned}$$

Therefore, the neural network is correct.

2.

(a) The zero-padding matrix is as following, and we do convolution in it with kernel

$$\text{zero-padding: } \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{Kernel: } \begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

The result is

$$\text{Output: } \begin{bmatrix} 3 & 3 & 5 & 2 & 5 & 3 \\ 3 & 8 & 5 & 6 & 5 & 5 \\ 6 & 5 & 8 & 8 & 8 & 4 \\ 4 & 7 & 6 & 8 & 6 & 5 \\ 3 & 8 & 7 & 7 & 7 & 5 \\ 4 & 3 & 5 & 4 & 5 & 3 \end{bmatrix}$$

(b) With 2×2 max-pooling layer with stride 2 after the result of (a), the result is

$$\text{Output: } \begin{bmatrix} 8 & 6 & 5 \\ 7 & 8 & 8 \\ 8 & 7 & 7 \end{bmatrix}$$