Quiz 6

Overfitting has always been the enemy of generalization. Dropout is very simple and yet very effective way to regularize networks by reducing coadaptation between the neurons.

Let **Y** be the intermediate activation of the neuron before dropout and **H** be the output of dropout.

1. Write the forward propagation of dropout in mathematical formate.
2. Write the backward propagation of dropout in mathematical formate.

1. Forward Propagation : $H = D \odot Y$

$$\begin{bmatrix} H_{11} & \cdots & H_{1J} \\ \vdots & H_{ij} & \vdots \\ H_{I1} & \cdots & H_{IJ} \end{bmatrix} = \begin{bmatrix} D_{11} & \cdots & D_{1J} \\ \vdots & D_{ij} & \vdots \\ D_{I1} & \cdots & D_{IJ} \end{bmatrix} \odot \begin{bmatrix} Y_{11} & \cdots & Y_{1J} \\ \vdots & Y_{ij} & \vdots \\ Y_{I1} & \cdots & Y_{IJ} \end{bmatrix}$$

$$H_{ij} = D_{ij} \times Y_{ij} = \begin{cases} Y_{ij} , & \text{if } D_{ij} = 1 \\ 0 , & \text{if } D_{ij} = 0 \end{cases} \text{※}$$

2. Back Propagation : $\frac{\partial L}{\partial Y} = \frac{\partial L}{\partial H} \odot \frac{\partial H}{\partial Y}$

$$\frac{\partial H}{\partial Y} = \begin{bmatrix} \frac{\partial H_{11}}{\partial Y_{11}} & \cdots & \frac{\partial H_{1J}}{\partial Y_{1J}} \\ \vdots & \frac{\partial H_{ij}}{\partial Y_{ij}} & \vdots \\ \frac{\partial H_{I1}}{\partial Y_{I1}} & \cdots & \frac{\partial H_{IJ}}{\partial Y_{IJ}} \end{bmatrix}$$

$$\frac{\partial H_{ij}}{\partial Y_{ij}} = \begin{cases} 1 ; & D_{ij} = 1 \\ 0 ; & D_{ij} = 0 \end{cases} = D_{ij} \Rightarrow \frac{\partial H}{\partial Y} = D$$

$$\Rightarrow \frac{\partial L}{\partial Y} = \frac{\partial L}{\partial H} \odot D \text{※}$$