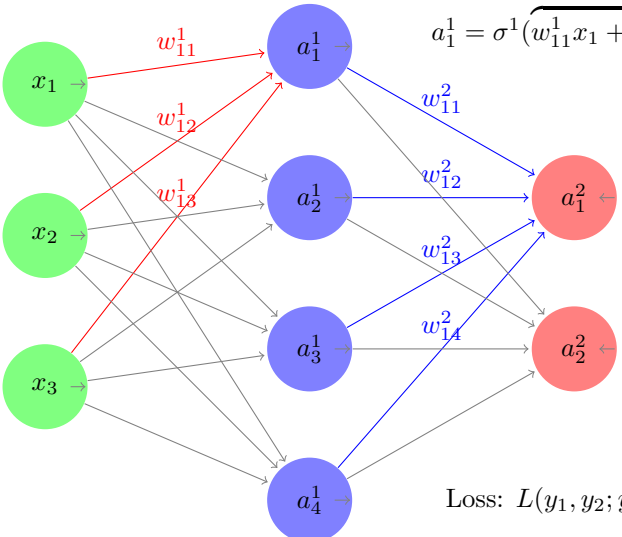
0: Input layer 1: Hidden layer 2: Output layer



$$a_1^1 = \sigma^1(\overbrace{w_{11}^1 x_1 + w_{12}^1 x_2 + w_{13}^1 x_3 + b_1^1}^{z_1^1})$$

$$a_1^2 = \sigma^2(\overbrace{\sum_{j=1}^{4} w_{1j}^2 a_j^1 + b_1^2}^{z_1^2}) \equiv \hat{y}_1$$

$$a_2^2 = \sigma^2(\overbrace{\sum_{j=1}^{4} w_{2j}^2 a_j^1 + b_2^2}^{z_2^2}) \equiv \hat{y}_2$$

Loss: $L(y_1, y_2; \hat{y}_1, \hat{y}_2)$; $L = 2$ is the last layer