$$a = \sum_{i=1}^{n} w_i x_i + b \tag{1}$$

$$y = f(a) \tag{2}$$

$$\mathbf{W}^{(l)} \leftarrow \mathbf{W}^{(l)} - \alpha \frac{\partial L}{\partial \mathbf{W}^{(l)}} \tag{3}$$

$$\mathbf{b}^{(l)} \leftarrow \mathbf{b}^{(l)} - \alpha \frac{\partial L}{\partial \mathbf{b}^{(l)}} \tag{4}$$

$$\Phi(\boldsymbol{x}) = \phi_K(\phi_{K-1}(\dots \phi_2(\phi_1(\boldsymbol{x}))\dots)) \tag{5}$$

$$\frac{\partial L_i^k}{\partial w_{pq}^k} \tag{6}$$

$$H(p, q) = -\sum_{i=1}^{n} p_i \log q_i \tag{7}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{8}$$

$$z = F(x) = f_K(f_{K-1}(\dots f_2(f_1(x))\dots)) \tag{9}$$

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda \sum_{i=1}^{N} |w_i| \tag{10}$$

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda \sum_{i=1}^{N} |w_i|^2 \tag{11}$$