$$L = \left(\hat{X} - X \right)^2$$

$$\hat{X} = non(X * W1) * W2$$

For non function:

$$\frac{\partial L}{\partial W1} = \frac{\partial L}{\partial \hat{X}} * \frac{\partial \hat{X}}{\partial non(XW1)} * \frac{\partial non(XW1)}{\partial XW1} * \frac{\partial XW1}{\partial W1}$$

$$\frac{\partial L}{\partial \hat{X}} = 2 * (non(X * W1) * W2 - X)$$

$$\frac{\partial \hat{X}}{\partial non(XW1)} = W2^T$$

$$\frac{\partial non(XW1)}{\partial XW1} = non'(X * W1)^T$$

$$\frac{\partial XW1}{\partial W1} = X$$

$$\frac{\partial L}{\partial W1} = 2 * (non(X * W1) * W2 - X) * W2^T * non'(X * W1)^T * X$$

$$\frac{\partial L}{\partial W2} = \frac{\partial L}{\partial \hat{X}} * \frac{\partial \hat{X}}{\partial W2}$$

$$\frac{\partial \hat{X}}{\partial W2} = non(X * W1)^T$$

$$\frac{\partial L}{\partial W2} = 2 * non(X * W1)^T * (non(X * W1) * W2 - X)$$

For tanh function:

$$\frac{\partial L}{\partial W1} = \frac{\partial L}{\partial \hat{X}} * \frac{\partial \hat{X}}{\partial tanh(XW1)} * \frac{\partial tanh(XW1)}{\partial XW1} * \frac{\partial XW1}{\partial W1}$$

$$\frac{\partial L}{\partial \hat{X}} = 2 * (tanh(X * W1) * W2 - X)$$

$$\frac{\partial \hat{X}}{\partial tanh(XW1)} = W2^T$$

$$\frac{\partial tanh(XW1)}{\partial XW1} = (1 - \tanh^2(X * W1))^T$$

$$\frac{\partial XW1}{\partial W1} = X$$

$$\frac{\partial L}{\partial W1} = 2 * (tanh(X * W1) * W2 - X) * W2^T * (1 - \tanh^2(X * W1))^T * X$$

$$\frac{\partial L}{\partial W2} = \frac{\partial L}{\partial \hat{X}} * \frac{\partial \hat{X}}{\partial W2}$$

$$\frac{\partial \hat{X}}{\partial W2} = tanh(X * W1)^T$$

$$\frac{\partial L}{\partial W2} = 2 * tanh(X * W1)^T * (tanh(X * W1) * W2 - X)$$