Homework5

Xiongjiao

$$x_{1:n} = x_1 \oplus x_2 \oplus \ldots \oplus x_n$$

$$c_i = f\left(\mathbf{w} \cdot \mathbf{x}_{i:i+h-1} + b_1\right)$$

$$O = f\left(MU + b_2\right)$$

$$\hat{y} = \text{softmax}(O)$$

$$J = CE(y, \hat{y})$$

Assume, the function of f is REUL

when $w > 0,$

$$\frac{\partial J}{\partial w} = \frac{\partial J}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial O} \cdot \frac{\partial O}{\partial f} \cdot \frac{\partial f}{\partial M} \cdot \frac{\partial M}{\partial c} \cdot \frac{\partial c}{\partial w} = -\frac{1}{\hat{y}} \hat{y}(1-\hat{y}) U^T I \; x^T = (1-\hat{y}) U^T I \; x^T$$

when $w \le 0,$

$$\frac{\partial J}{\partial w} = \frac{\partial J}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial O} \cdot \frac{\partial O}{\partial f} \cdot \frac{\partial f}{\partial M} \cdot \frac{\partial M}{\partial c} \cdot \frac{\partial c}{\partial w} = 0$$