Aidan Fischer ~~MA~~ CS 584 HW1 2/14/2024

I pledge my honor that I have abided by the Stevens Honor System

$$J = \sum_{i=1}^{n} L(x_i, y_i) w, b) = -\frac{1}{n}\sum_{i=1}^{n}\left(y\log\hat{y} + (1-y)\log(1-\hat{y})\right) + \lambda\sum_{j=1}^{d} w_j^2$$

$$\frac{\partial J}{\partial w} = \frac{d}{dw}\left(-\frac{1}{n}\sum_{i=1}^{n}\left(y\log(\hat{y}) + (1-y)\log(1-\hat{y})\right) + \lambda\sum_{j=1}^{d} w_j^2\right)$$

Expand $\hat{y}$ to its definition

$$= \frac{d}{dw}\left(-\frac{1}{n}\sum_{i=1}^{n}\left(y\log\left(\frac{1}{1+e^{-(xw+b)}}\right) + (1-y)\log\left(1-\frac{1}{1+e^{-(xw+b)}}\right)\right) + \lambda\sum_{j=1}^{d} w_j^2\right)$$

Bring ~~Apply~~ derivative operator inside

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(y\frac{d}{dw}\log\left(\frac{1}{1+e^{-(xw+b)}}\right) + (1-y)\frac{d}{dw}\log\left(1-\frac{1}{1+e^{-(xw+b)}}\right)\right) + \lambda\sum_{j=1}^{d}\frac{d}{dw}w_j^2$$

Apply derivative operator

$$\cancel{= -\frac{1}{n}\left(y\left(\frac{1}{\frac{1}{1+e^{-(xw+b)}}}\right)\left(0-1\cdot(-xw+b)e^{-(xw+b)}\right)\right)}$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(y\left(1+e^{-(xw+b)}\right)\left(\frac{0-(-xe^{-(xw+b)})}{(1+e^{-(xw+b)})^2}\right) + (1-y)\left(\frac{1}{1-\frac{1}{1+e^{-(xw+b)}}}\right)\left(-\left(\frac{0-(-xe^{-(xw+b)})}{(1+e^{-(xw+b)})^2}\right)\right)\right) + 2\lambda w$$

$(\log'(f(x)) = \frac{f'(x)}{f(x)})$

Simplify

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(y\left(\frac{xe^{-(xw+b)}}{1+e^{-(xw+b)}}\right) + (1-y)\left(\frac{-xe^{-(xw+b)}}{\left(1-\frac{1}{1+e^{-(xw+b)}}\right)(1+e^{-(xw+b)})^2}\right)\right) + 2\lambda w$$

Simplify

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{xye^{-(xw+b)}}{1+e^{-(xw+b)}} + \frac{-(1-y)xe^{-(xw+b)}}{(1+e^{-(xw+b)})^2-(1+e^{-(xw+b)})}\right) + 2\lambda w$$

Simplify

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{xye^{-(xw+b)}}{1+e^{-(xw+b)}} + \frac{-(1-y)xe^{(xw+b)}}{(1+e^{-(xw+b)})(1+e^{+(xw+b)}-1)}\right) + 2\lambda w$$

Simplify

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{xye^{-(xw+b)}}{1+e^{-(xw+b)}} + \frac{-(1-y)xe^{-(xw+b)}}{(1+e^{-(xw+b)})(e^{-(xw+b)})}\right) + 2\lambda w$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{xye^{-(xw+b)}}{1+e^{-(xw+b)}} + \frac{-(1-y)x}{1+e^{-(xw+b)}}\right) + 2\lambda w$$

$$\boxed{= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x(ye^{-(xw+b)}-(1-y))}{1+e^{-(xw+b)}}\right) + 2\lambda w} \quad \left(\frac{\partial J}{\partial w}\right)$$

$$\frac{\partial J}{\partial b} = \frac{\partial}{\partial b}\left(-\frac{1}{n}\sum_{i=1}^{n}\left(y\log\left(\frac{1}{1+e^{-(xw+b)}}\right) + (1-y)\log\left(1-\frac{1}{1+e^{-(xw+b)}}\right)\right) + \lambda\sum_{j=1}^{d} w_j^2\right)$$

Bring in derivative

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(y\frac{d}{db}\log\left(\frac{1}{1+e^{-(xw+b)}}\right) + (1-y)\frac{d}{db}\log\left(1-\frac{1}{1+e^{-(xw+b)}}\right)\right) + \lambda\sum_{j=1}^{d}\frac{\partial}{\partial b} w_j^2$$

Apply derivative

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(y\left(1+e^{-(xw+b)}\right)\left(\frac{-(-e^{-(xw+b)})}{(1+e^{-(xw+b)})^2}\right) + (1-y)\left(\frac{1}{1-\frac{1}{1+e^{-(xw+b)}}}\right)\left(-\frac{0-(-e^{-(xw+b)})}{(1+e^{-(xw+b)})^2}\right)\right) + 0$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{ye^{-(xw+b)}}{1+e^{-(xw+b)}} + \frac{-(1-y)e^{-(xw+b)}}{\left(1-\frac{1}{1+e^{-(xw+b)}}\right)(1+e^{-(xw+b)})^2}\right)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{ye^{-(xw+b)}}{1+e^{-(xw+b)}} - \frac{(1-y)e^{-(xw+b)}}{((1+e^{-(xw+b)})-1)(1+e^{-(xw+b)})}\right)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{ye^{-(xw+b)}}{1+e^{-(xw+b)}} - \frac{(1-y)e^{-(xw+b)}}{e^{-(xw+b)}(1+e^{-(xw+b)})}\right)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\left(\frac{ye^{-(xw+b)}}{1+e^{-(xw+b)}} - \frac{(1-y)}{1+e^{-(xw+b)}}\right)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\frac{ye^{-(xw+b)}-1+y}{1+e^{-(xw+b)}} = -\frac{1}{n}\sum\frac{y(e^{-(xw+b)}+1)-1}{1+e^{-(xw+b)}}$$

$$\boxed{= -\frac{1}{n}\sum_{i=1}^{n} y - \frac{1}{1+e^{-(xw+b)}}} \quad \left(\frac{\partial J}{\partial b}\right)$$