$$L(w) = -\frac{1}{N}\sum_{i=1}^{N} \log P[Y = y_i \mid x_i, w]$$

$$\frac{dL(w)}{dw} = -\frac{1}{N}\sum_{i=1}^{N} \frac{1}{P[Y = y_i \mid x_i, w]} \cdot \frac{dP[Y = y_i \mid x_i, w]}{dw}$$

$$\frac{dP[Y = y_i \mid x_i, w]}{dw} = P(Y = y_i \mid x_i, w)\,(1 - P(Y = y_i \mid x_i, w))\,x_i$$

$$\therefore \frac{dL(w)}{dw} = -\frac{1}{N}\sum_{i=1}^{N} (1 - P(Y = y_i \mid x_i, w)) \cdot x_i$$