COEN 240
HW6

Bhanu Prakash Naredla
00001630571

$$l(w) = y \log(\sigma(w^T x)) + (1-y) \log(1-\sigma(w^T x))$$

$$\nabla_{w_j} l(w) = \nabla_{w_j}\left(y \log(\sigma(w^T x))\right) + \nabla_{w_j}\left((1-y)\log(1-\sigma(w^T x))\right)$$

$$\nabla_{w_j}\left(y \log(\sigma(w^T x))\right)$$

$$\Rightarrow y \cdot \frac{1}{\sigma(w^T x)} \cdot \sigma(w^T x) \cdot (1-\sigma(w^T x)) \cdot x_j$$

$$\Rightarrow y \cdot (1-\sigma(w^T x)) \cdot x_j$$

$$\Rightarrow x_j \left(y - y\sigma(w^T x)\right) \quad \text{—— (1)}$$

$$\nabla_{w_j}\left((1-y)\log(1-\sigma(w^T x))\right)$$

$$\Rightarrow (1-y) \cdot \frac{1}{1-\sigma(w^T x)} \cdot -(\sigma(w^T x) \cdot (1-\sigma(w^T x)) \cdot x_j$$

$$\Rightarrow (1-y) \cdot -\sigma(w^T x) \cdot x_j$$

$$\Rightarrow x_j \left(y\sigma(w^T x) - \sigma(w^T x)\right) \quad \text{— (2)}$$

add ①, ②

$$\nabla_{w_j} l(w) = x_j\left(y - y\sigma(w^T x)\right) + x_j\left(y\sigma(w^T x) - \sigma(w^T x)\right)$$

$$\Rightarrow x_j\left(y - \cancel{y\sigma(w^T x)} + \cancel{y\sigma(w^T x)} - \sigma(w^T x)\right)$$

$$\boxed{\Rightarrow \left[y - \sigma(w^T x)\right] \cdot x_j}$$

NOTE: we directly took $\nabla_{w_j}(w^T x) = x_j$

$$w^T x = w_1 x_1 + w_2 x_2 + \ldots + w_j x_j + \ldots + w_k x_k$$

$$\Rightarrow \nabla_{w_j}(w^T x) = x_j$$