Yash Butala
17CS30038

Q1.1) $x_i \in \mathbb{R}^d$

$y_i \in \{-1, 1\}$

$w \in \mathbb{R}^d$

$\hat{y}_i = \tanh(w \cdot x_i)$

$$\text{Loss} = \sum_{i=1}^{N} l(y_i, \hat{y}_i) + \lambda \|w\|^2$$

step size $= h$

$$l(y_i, \hat{y}_i) = \log_e \left(1 + e^{-y_i \hat{y}_i}\right)$$

$\Rightarrow$

$$\frac{\partial \text{Loss}}{\partial w} = \frac{\partial}{\partial w}\left[\sum_{i=1}^{N} l(y_i, \hat{y}_i) + \lambda \|w\|^2\right]$$

$$= 2\lambda(w) + \sum_{i=1}^{N}\left[\frac{\partial}{\partial w}\left(\log_e\left(1 + e^{-y_i \hat{y}_i}\right)\right)\right]$$

$$= 2\lambda(w) + \sum_{i=1}^{N} \frac{(-y_i)\, e^{-y_i \hat{y}_i}}{1 + e^{-y_i \hat{y}_i}} \cancel{(1-\hat{y}^2)} \times \frac{\partial}{\partial w} \tanh(x_i \cdot w)$$

$$= 2\lambda w + \sum_{i=1}^{N} \frac{(-y_i)\, e^{-y_i \hat{y}_i}}{1 + e^{-y_i \hat{y}_i}} \left(1 - \tanh^2(x_i \cdot w)\right)(x_i)$$

$$= 2\lambda w + \sum_{i=1}^{N} \frac{(-y_i)\, e^{-y_i \hat{y}_i}}{1 + e^{-y_i \hat{y}_i}} \left(1 - \hat{y}_i^2\right)(x_i)$$

$$\therefore w^{t+1} = w_t - h\left[2\lambda w_t + \sum_{i=1}^{N} \frac{(-y_i)\, e^{-y_i \hat{y}_i}}{1 + e^{-y_i \hat{y}_i}} \left(1 - \hat{y}_i^2\right)(x_i)\right]$$

Yash Butala
17CS30038

ANS : $$w_{t+1} = w_t[1-2\lambda\eta] + \sum_{i=1}^{N} \eta y_i \frac{e^{-y_i\hat{y}_i}}{1+e^{-y_i\hat{y}_i}} (1-\hat{y}_i^{\,2})(x_i)$$

[Note: for SGD loss, in above equation: N = 1.]
[I have directly used loss given in the question.]