Assignment 2 (Sandipan Haldar 18MF10050)

1) $l(y_i, \hat{y}_i) = \log(1+\exp(-y_i \cdot \hat{y}_i))$

Loss function

$$\sum_{i=1}^{n} \log(1+\exp(-y_i \cdot \hat{y}_i)) + \lambda \|w\|^2$$

$$\sum_{i=1}^{n} \log(1+\exp(-y_i \cdot \hat{y}_i)) + \lambda \sum_{i=1}^{d} w_i^2$$

We know $w_{t+1} = w_t - \eta \frac{\partial Loss}{\partial w}$

let us take a single component of w.

$$w_{t+1,i} = w_{t,i} - \eta \frac{\partial loss}{\partial w_{t,i}}$$

$$loss = \sum_{j=1}^{n} \log(1+\exp(-y_j \cdot \hat{y}_j)) + \lambda \sum_{j=1}^{d} w_j^2$$

$$\frac{\partial loss}{\partial w_j} = \sum \frac{\partial l(y_i, \hat{y}_i)}{\partial w_j} + 2\lambda w_j$$

Now $\frac{\partial l(y_i, \hat{y}_i)}{\partial w_j} =$

PRIORITY	APPOINTMENTS	IMPORTANT NOTES

$$\frac{\partial}{\partial w_j} \log\left(1+\exp(-y_i \cdot \hat{y}_i)\right)$$

$$= \frac{1}{\log(1+\exp(-y_i \cdot \hat{y}_i))} \cdot \frac{\partial}{\partial w_j} e^{-y_i \hat{y}_i}$$

$$= \frac{1}{1+\exp(-y_i \hat{y}_i)} \cdot e^{-y_i \hat{y}_i} \cdot (-y_i) \cdot \frac{\partial \hat{y}_i}{\partial w_j}$$

Now $\hat{y}_i = \tanh(w^T x_i) = \tanh\left(\sum_{j=1}^{d} w_j x_{ij}\right) \Rightarrow \frac{\partial \hat{y}_i}{\partial w_j} = x_{ij} \cdot (1 - \hat{y}_i^2)$

$$\therefore \frac{\partial l(y_i, \hat{y}_i)}{\partial w_j} = \frac{e^{-y_i \cdot \hat{y}_i}}{1+\exp(-y_i \hat{y}_i)} \cdot (-y_i) \cdot x_{ij} \cdot (1 - \hat{y}_i^2)$$

$$\therefore \frac{\partial \text{loss}}{\partial w_j} = \sum_{i=1}^{n} \frac{-y_i e^{-y_i \hat{y}_i} (1-\hat{y}_i^2)}{1+\exp(-y_i \hat{y}_i)} x_{ij} + 2\lambda w_j$$

PRIORITY	APPOINTMENTS	IMPORTANT NOTES

09

For $j = 1$ to d

10 thus $$w_{t+1,j} = w_{tj} - \eta\left(\sum_{i=1}^{\hat{n}} \frac{-y_i e^{-y_i \hat{y}_i}(1-\hat{y}_0)^2}{1+\exp(-y_i \hat{y}_i)} x_{ij} + 2\lambda w_j\right)$$

11

12 ~~$$w_{t+1} = w_t - \eta\left(\sum_{i=1}^{\hat{n}} \frac{-y_i e^{-y_i \hat{y}_i}}{1+e^{-y_i \hat{y}_i}} \cdot x_t + 2\lambda w_t\right)$$~~

01

02 ind

$$\boxed{w_{t+1} = w_t - \eta\left(x^T \cdot l + 2\lambda w_t\right)}$$

03

04 Where $l = [l_1\ l_2\ l_3 \dots l_n]^T$

05 $$l_i = \frac{-y_i e^{-y_i \hat{y}_i}}{1+e^{-y_i \hat{y}_i}}(1-\hat{y}_i^2)$$

06

07

PRIORITY	APPOINTMENTS	IMPORTANT NOTES