Question 5

@ $\boxed{\text{Lecture 16}}$

to prove :- if an activation function is sign function if we replace it with sigmoid function; we need to prove update parameters don't change

distance equ. is as follows.

$$d_i = -y_i \left(\beta^T x_i + \beta_0 \right)$$

a pt. x_i with true label is y_i

prediction of perceptron

$$\beta^T x_i + \beta_0 \quad \geq 0 \quad +1$$

$$< 0 \quad -1$$

changing activation fn will not influence update parameters because distance is not dependent on the activation fn.

$$\beta^{new} \leftarrow \beta_{old} - \eta \frac{\partial L_i}{\partial \beta}$$

$$= \beta_{old} - \eta\,(-y_i x_i)$$

$$\beta_0^{new} \leftarrow \beta_0^{old} - \eta\,(-y_i)$$

classmate

Date ______
Page ______

(B) Given :

$$\phi(\beta, \beta_0) = -\sum_{i=1}^{N} y_i(\beta^T x_i + \beta_0)$$

$$\min \phi(\beta, \beta_0) \quad s.t \quad ||\beta|| = 1$$

$$\beta^T \beta = 1$$

$$L(\beta, \beta_0, \lambda) = \phi(\beta, \beta_0) - \lambda g(\beta)$$

$$[g(\beta) = \beta^T \beta - 1]$$

$$= -\sum_{i=1}^{N} y_i(\beta^T x_i + \beta_0) - \lambda(\beta^T \beta - 1)$$

$$\frac{\partial L}{\partial \beta} = -\sum_{i=1}^{N} y_i(x_i) - \lambda(2\beta) = 0$$

$$= \sum_{i=1}^{N} x_i y + 2\lambda\beta = 0$$

$$\frac{\partial L}{\partial \beta_0} = -\sum_{i=1}^{N} y_i = 0$$

$$\frac{\partial L}{\partial \lambda} = -(\beta^T \beta - 1) = 0$$

$$= \beta^T \beta = 1$$

Scanned with CamScanner

$$\beta_{new} \leftarrow \beta - \eta \frac{\partial L}{\partial \beta}$$

$$\beta_{0\,new} \leftarrow \beta_0 - \eta \frac{\partial L}{\partial \beta_0}$$

classmate

Date

Page

6 $y_1 = \sigma(\beta_{11} x + \beta_{01})$

$y_2 = \text{sign}(\beta_{12} y_1 + \beta_{02})$

$d_i^2 = -y_i(\beta_{12} y_{1i} + \beta_{02})$

$$\frac{\partial d_i}{\partial \beta_{11}} = -y_i\left(\beta_{12} \frac{\partial y_{1i}}{\partial \beta_{11}}\right)$$

$$= -y_i\left(\beta_{12}\, \sigma(\beta_{11} x_1 + \beta_{01}) . x_i\right)$$

$$= \frac{-y_i x_i \beta_{12} . e^{-v_i}}{(1+e^{-v_i})^2}$$

$$\beta_{11} - \eta \frac{\partial d_i}{\partial \beta_{11}} \longrightarrow \beta_{11}$$

$$\partial d_i = -y_i\left(\beta_{12} \frac{\partial y_{1i}}{\partial \beta_{01}}\right)$$

$$\frac{\partial y_{1i}}{\partial \beta_{01}} = \sigma(\beta_1 + x_i + \beta_{01})$$

$$= \frac{e^{-v_i}}{(1+e^{-v_i})^2}$$

$$\frac{\partial l_i}{\partial \beta_{01}} = -y_i \frac{\beta_{12}\, e^{-v_i}}{(1+e^{-v_i})^2}$$

$$\frac{\partial l_i}{\partial \beta_{12}} = -y_i \, (y_{11}')$$

$$= -y_i \, (\sigma \, (\beta_{11} x + \beta_{01}))$$

$$\frac{\partial l_i}{\partial \beta_{02}} = -y_i$$

$$\beta_{01} \rightarrow \beta_{01} - \eta \frac{\partial l_i}{\partial \beta_{01}}$$

$$\beta_{12} \rightarrow \beta_{12} - \eta \frac{\partial l_i}{\partial \beta_{12}}$$

$$\beta_{02} \rightarrow \beta_{02} - \eta \frac{\partial l_i}{\partial \beta_{02}}$$

Scanned with CamScanner