APOORVA ARYA
2020032

SML Assignment - 3 - Theory

Ans2 We get the best accuracy when we take n_components = 15 because that way we have used the maximum number of features in this case for analysis.

Ans 5 (a) $d_i = y_i(\beta^T x_i + \beta_0)$

Changing activation function does not affect the update rule.

→ Distance does not depend on the activation function.

$$\beta_{new} \Rightarrow \beta - \eta \frac{\partial d_i}{\partial \beta}$$

$$\beta_{0new} \Rightarrow \beta_0 - \eta \frac{\partial d_i}{\partial \beta}$$

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

we know that $\beta^T \beta = 1$

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

$$\alpha(\beta, \beta_0) = \phi(\beta, \beta_0) + \lambda(g(\beta))$$

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

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

$$= \sum_{i=1}^{n} x_i \, y_i + 2\lambda \beta = 0$$

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

$$= \sum_{i=1}^{N} y_i = 0$$

Ans6 $\quad y_1 = \sigma(\beta_{11} x + \beta_{01})$

$$y_2 = \text{sigt}(\beta_2 y_1 + \beta_{02})$$

$$x \longrightarrow \bigcirc \longrightarrow \bigcirc \; y_2$$
$$\uparrow y_1$$

For misclassified samples

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

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

$$\sigma'(x) = \sigma(x)\left[1 - \sigma(x)\right]$$

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

$$= -y_i\left(\beta_{12}\left[\sigma(\beta_{11} x_i + \beta_{01}) \cdot (1 - \sigma(\beta_{11} x_i + \beta_{01}))\right] x_i\right)$$

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

Now,

$$\frac{\partial d_i}{\partial \beta_{11}} = -y_i\left(\beta_{12} \frac{\partial y_{1i}}{\partial \beta_{01}}\right)$$

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

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

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

$$\frac{\partial d_i}{\partial \beta_{02}} = -y_i \, (y_{1i})$$

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

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

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

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

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

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