Deep learning for big visual data

HW1

student ID: NM6111035

Name: 吳建澄

1.

$$L_1 = |0.15-0| + |0.7-1| + |0-0.1| + |0-0.05|$$
$$= 0.15 + 0.3 + 0.1 + 0.05$$
$$= 0.6$$

$$L_2 = (0.15-0)^2 + (0.7-1)^2 + (0.1)^2 + (0.05)^2$$
$$= 0.125$$

$$L_2 = \sqrt{0.125} = 0.3536$$

$$L_{ce} = -1 \log 0.7 = 0.3567$$

2.

C_1: $TP = 68$

$FP = 14 + 12 + 6 = 32$

$FN = 12 + 9 + 11 = 32$

$Precision = \frac{68}{32+68} = 0.68$

$Recall = \frac{68}{32+68} = 0.68$

$F_1 = \frac{2 \times 0.68 \times 0.68}{0.68 + 0.68} = 0.68$

C_2: $TP = 74$

$FP = 12 + 3 + 10 = 25$

$FN = 14 + 5 + 7 = 26$

$Precision = \frac{74}{74+25} = 0.7475$

$Recall = \frac{74}{74+26} = 0.74$

$F_1 = \frac{2 \times 0.74 \times 0.7475}{0.7475 + 0.74} = 0.744$

C3: $TP = 82$

$FP = 9+5+12 = 26$

$FN = 12+3+3 = 18$

$Precision = \frac{82}{26+82} = 0.787$

$Recall = \frac{82}{82+18} = 0.82$

$F_1 = \frac{2 \times 0.82 \times 0.787}{0.82+0.787} = 0.803$

C4: $TP = 72$

$FP = 21$

$FN = 28$

$Precision = 0.7742$

$Recall = 0.72$

$F_1 = 0.746$

Overall average accuracy = 0.74

micro-averge precision = 0.747

3. $S_r = \sum e_i^2 = \sum (y_i - b_0 - b_1 x_1 - b_2 x_2)^2$

$$\frac{\partial S_r}{\partial b_0} = -2 \sum (y_i - b_0 - b_1 x_1 - b_2 x_2)$$

$$\frac{\partial S_r}{\partial b_1} = -2 x_1 \sum (y_i - b_0 - b_1 x_1 - b_2 x_2)$$

$$\frac{\partial S_r}{\partial b_2} = -2 x_2 \sum (y_i - b_0 - b_1 x_1 - b_2 x_2)$$

$$\begin{bmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \end{bmatrix}$$

$\sum x_1 = 555$ $\quad \sum y = 1452$

$\sum x_2 = 145$ $\quad \sum x_1 y = 101895$

$\sum x_1^2 = 38767$ $\quad \sum x_2 y = 2536$

$\sum x_2^2 = 2823$ $\quad \sum x_1 x_2 = 9854$

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} -6.8675 \\ 3.1478 \\ -1.6561 \end{bmatrix} \#$$

4. SVM works by finding the best line or hyperplane that separates two classes in high-dimention.

Non linear SVM is when the data can't be separated by hyperplan or line. In this case, kernel Trick is used to map the data into higher dimention space to make data can be separated.

5. discriminative classifier learn to separate different classes of data.

generate classifier model the joint probability distribution of the input features and the target variable