EEP568

Varad Dhat

October 2023

1 Problem 1 (b)

The equation which we have for two variables:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_1^2 + a_4x_1x_2 + a_5x_2^2 + e \tag{1}$$

Now sum the equation for sum of residual square is calculated by:

$$Sr = \sum_i e_i^2 = \sum_i (Y_i - (a_0 + a_1x_{1i} + a_2x_{2i} + a_3x_{1i}^2 + a_4x_{1i}x_{2i} + a_5x_{2i}^2))^2 \tag{2}$$

$$\frac{\partial Sr}{\partial a_0} = -2\sum_i Y_i - (a_0 + a_1x_{1i} + a_2x_{2i} + a_3x_{1i}^2 + a_4x_{1i}x_{2i} + a_5x_{2i}^2)^2$$

$$\frac{\partial Sr}{\partial a_1} = -2\sum_i x_{1i}(Y_i - (a_0 + a_1x_{1i} + a_2x_{2i} + a_3x_{1i}^2 + a_4x_{1i}x_{2i} + a_5x_{2i}^2)$$

$$\frac{\partial Sr}{\partial a_2} = -2\sum_i x_{2i}(Y_i - (a_0 + a_1x_{1i} + a_2x_{2i} + a_3x_{1i}^2 + a_4x_{1i}x_{2i} + a_5x_{2i}^2)$$

$$\frac{\partial Sr}{\partial a_3} = -2\sum_i x_{1i}^2(Y_i - (a_0 + a_1x_{1i} + a_2x_{2i} + a_3x_{1i}^2 + a_4x_{1i}x_{2i} + a_5x_{2i}^2)$$

$$\frac{\partial Sr}{\partial a_4} = -2\sum_i x_{1i}x_{2i}(Y_i - (a_0 + a_1x_{1i} + a_2x_{2i} + a_3x_{1i}^2 + a_4x_{1i}x_{2i} + a_5x_{2i}^2)$$

$$\frac{\partial Sr}{\partial a_5} = -2\sum_i x_{1i}^2(Y_i - (a_0 + a_1x_{1i} + a_2x_{2i} + a_3x_{1i}^2 + a_4x_{1i}x_{2i} + a_5x_{2i}^2)$$

Now we have:

$$B \times a = c \tag{3}$$

$$\mathbf{B} = \begin{bmatrix} n & \sum x_{1i} & \sum x_{2i} & \sum x_{1i}^2 & \sum x_{1i}x_{2i} & \sum x_{2i}^2 \\ \sum x_{1i} & \sum x_{1i}^2 & \sum x_{1i}x_{2i} & \sum x_{1i}^3 & \sum x_{1i}^2x_{2i} & \sum x_{1i}x_{2i}^2 \\ \sum x_{2i} & \sum x_{1i}x_{2i} & \sum x_{2i}^2 & \sum x_{1i}^2x_{2i} & \sum x_{1i}x_{2i}^2 & \sum x_{2i}^3 \\ \sum x_{1i}^2 & \sum x_{1i}^3 & \sum x_{1i}^2x_{2i} & \sum x_{1i}^4 & \sum x_{1i}^3x_{2i} & \sum x_{1i}^2x_{2i}^2 \\ \sum x_{1i}x_{2i} & \sum x_{1i}^2x_{2i} & \sum x_{1i}x_{2i}^2 & \sum x_{1i}^3x_{2i} & \sum x_{1i}^2x_{2i}^2 & \sum x_{1i}x_{2i}^3 \\ \sum x_{2i}^2 & \sum x_{1i}x_{2i}^2 & \sum x_{2i}^3 & \sum x_{1i}^2x_{2i}^2 & \sum x_{1i}x_{2i}^3 & \sum x_{2i}^4 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} Y_i \\ X_{1i}Y_i \\ X_{2i}Y_i \\ X_{1i}^2Y_i \\ X_{1i}X_{2i}Y_i \\ X_{2i}^2Y_i \end{bmatrix} \quad \mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}$$

2 Problem 3

(1)

$$\mathbf{y_t} = (0, 0, 1, 0, 0) \quad \text{.... true label}$$
$$\hat{\mathbf{y_p}} = (0.15, 0.05, 0.6, 0.08, 0.12) \quad \text{.... predictions (trained classifier)}$$

L1:

$$\begin{aligned} \text{Loss} &= \frac{1}{5}\sum_{i=1}^{n} |y_{ti} - \hat{y}_{pi}| \\ &= \frac{1}{5}(0.15 + 0.05 + 0.4 + 0.08 + 0.12) \approx 0.16 \end{aligned}$$

L2:

$$\begin{aligned} \text{Loss} &= \frac{1}{5}\sum_{i=1}^{n} (y_{ti} - \hat{y}_{pi})^2 \\ &= \frac{1}{5}(0.0025 + 0.0025 + 0.16 + 0.0064 + 0.0144) \\ &\approx 0.0412 \end{aligned}$$

Cross entropy loss:

$$\begin{aligned} \text{Loss} &= -\sum_{i=1}^{n} y_{ti}\log(\hat{y}_{pi}) \\ &= -\log(0.6) \\ &\approx 0.737 \end{aligned}$$

KL divergence:

$$\begin{aligned} D_{KL}(y_{ti}||\hat{y}_{pi}) &= \sum_{i=1}^{n} y_{ti}(\log(y_{ti}) - \log(\hat{y}_{pi})) \\ &= \log(1) - \log(0.6) \\ &\approx 0.737 \end{aligned}$$

(2) Total = 468

$$\begin{aligned} TP &= 55 + 80 + 98 + 87 = 320 \\ TN_{c1} &= 468 - (100 + 101) + 55 = 322 \\ TN_{c2} &= 468 - (113 + 105) + 80 = 330 \\ TN_{c3} &= 468 - (124 + 100) + 98 = 302 \\ TN_{c4} &= 468 - (131 + 122) + 87 = 302 \end{aligned}$$

$$\begin{array}{cccc} FP_{c1} = 46, & FP_{c2} = 25, & FP_{c3} = 42, & FP_{c4} = 49 \\ FN_{c1} = 45, & FN_{c2} = 33, & FN_{c3} = 26, & FN_{c4} = 49 \end{array}$$

Overall average accuracy:

$$\begin{aligned} Acc &= \frac{TP + TN}{TP + TN + FP + FN} \\ &= \frac{320}{468} = 68.37\% \approx 0.68376 \end{aligned}$$

Per class precision:

$$
\begin{aligned}
P_{c1} &= \frac{TP_1}{TP_1 + FP_1} = \frac{55}{55+46} = 0.5445 \\
P_{c2} &= \frac{80}{80+25} = 0.7619 \\
P_{c3} &= \frac{98}{98+42} = 0.7 \\
P_{c4} &= \frac{87}{87+35} = 0.7131
\end{aligned}
$$

Per class Recall:

$$
\begin{aligned}
Rc_1 &= \frac{TP_1}{TP_1 + FN_1} = \frac{55}{55+45} = 0.55 \\
Rc_2 &= \frac{TP_2}{TP_2 + FN_2} = \frac{80}{80+33} = 0.708 \\
Rc_3 &= \frac{98}{98+26} = 0.7903 \\
Rc_4 &= \frac{87}{87+49} = 0.664
\end{aligned}
$$

F1 score:

$$F1_4 = 0.6877$$

Micro Average precision:

$$\text{Micro Average precision} = \frac{\sum_{i=c1}^{4} TP_i}{\sum_{i=c1}^{4} TP_i + \sum_{i=c1}^{4} FP_i}$$

$$= \frac{320}{320+148} = 68.376\% \approx 0.68376$$

(3)

Discriminative classifiers focus on directly modeling the decision boundary between classes. They achieve this by learning the conditional probability P(yx), where y represents the label and x the input data. Some well-known examples of discriminative classifiers include Logistic Regression, Support Vector Machines, and Neural Networks. A key advantage of discriminative models is that they often yield higher accuracy in classification tasks, especially when there's an abundance of data. However, their primary limitation is that they don't provide insights into how the data is generated.

On the other hand, generative classifiers aim to understand how the data for each class might have been generated. This involves learning the joint probability distribution P(x,y) and then deriving P(yx) for classification. Generative classifiers give insights into the underlying structure of the data and can even generate new samples similar to the input data. However, they might not always be as accurate as discriminative models when it comes to pure classification tasks.