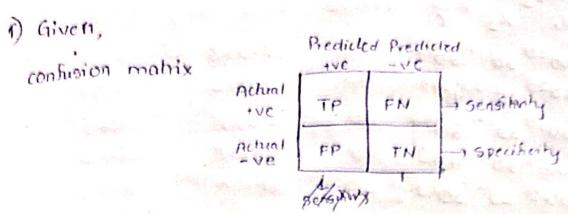
Machine learning - E1 SLOT



people who really have diabetes = achol postive = 6%

probability that a person who lest positive = ? = TP activity has the diabeles

From
$$\textcircled{2}$$
, $TP = 0.9TP + 0.9FN$

$$TP = 9FN$$

$$TP = 9FN - \textcircled{9}$$

From $\textcircled{3}$, $TN = 0.85TN + 0.35FP$

$$0.15FN = 0.85FP$$

$$19TN = 25FF \implies 3TN = 18FP$$

$$TN = 6FP$$

WE Know that TP+FN+FP+TN=100%=1 and TP+FN=0.06

: probability that a person who Tests the have diabetes is 0.054

) given, data

We take Regression line as Y=a,x+a,x2+c

substitute data in points, we get

into the mohix form as

$$\begin{bmatrix} 200 & 30 & 1 \\ 150 & 25 & 1 \\ 250 & 35 & 1 \\ 300 & 50 & 1 \\ 100 & 20 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ c \end{bmatrix} = \begin{bmatrix} 50 \\ 40 \\ 60 \\ 70 \\ 30 \end{bmatrix}$$

It is in the form of Ax=8

$$A^{T}A = \begin{bmatrix} 200 & 150 & 250 & 300 & 100 \\ 30 & 25 & 35 & 50 & 20 \\ 1 & 1 & 1 & 1 \\ 250 & 35 & 1 \\ 300 & 50 & 1 \\ 100 & 20 & 1 \end{bmatrix}$$

$$X = (A^{T}A)^{-1} \cdot A^{T}B = \begin{bmatrix} 0.1692 \\ 0.5128 \\ -1.282 \end{bmatrix}$$

Regression line, Y = 0.1692 x, + 0.5128 x2 - 1.282

$$y = 0.1692 \times 255 + 0.5128 \times 40 - 1.282$$

 $y = 62.376$

これをからなるとことができましたがなるとこと

The state of the s

これが、これを対すったし、

c)
$$X_1 \quad X_2 \quad \text{Ypred} \quad \text{Yince} \quad (\text{Ybrue-Ypred}) \quad (\text{Ybrue-Ypred})^2$$
 $200 \quad 30 \quad 47.942 \quad 50 \quad 2.058 \quad 4.235$
 $150 \quad 25 \quad 36.918 \quad 40 \quad 3.082 \quad 9.498$
 $250 \quad 35 \quad 58.966 \quad 60 \quad 1.034 \quad 1.069$
 $300 \quad 50 \quad 75.118 \quad 70 \quad -5.118 \quad 26.193$
 $100 \quad 20 \quad 25.894 \quad 30 \quad 4.106 \quad 16.859$

$$MSE = \frac{1}{n} \sum_{i=1}^{5} (\text{Ybrue-Ypred})^2 = \frac{58.304}{5}$$

$$= 11.660 \quad 8$$

2) Given

$$X_1$$
 X_2 X_1^2 X_2^2 X_1X_2 $X_1-\overline{X_1}$ $X_2-\overline{X_1}$
 5 7 25 49 35 -3.5 -2.75
 4 6 16 36 24 -4.5 -3.75
 11 17 121 289 187 2.5 7.25
 14 9 196 81 126 5.5 -0.75

a)
$$\overline{X}_1 = 8.5$$

 $\overline{X}_2 = 9.75$
 $\overline{X}_1 \times 2 = 93$

b) covariance matrix =
$$\begin{bmatrix} cov(x_1, x_2) & cov(x_1, x_2) \\ cov(x_2, x_1) & cov(x_2, x_2) \end{bmatrix}$$

$$Cov(x_1, x_1) = \sigma^2(x_1) = 17.25$$

 $Cov(x_1, x_2) = \sigma^2(x_2) = 18.6875$
 $Cov(x_1, x_2) = \overline{x_1.x_2} - \overline{x_1.\overline{x_2}}$

() Eigen values and eigen vectors of 5 are

$$\lambda_1 = -\frac{575 + \sqrt{105505}}{32}$$

$$X_1 = \begin{bmatrix} -\frac{2}{3} + \sqrt{105505} \\ 324 \end{bmatrix}$$
 $X_2 = \begin{bmatrix} -\frac{23}{324} + \sqrt{105505} \\ 324 \end{bmatrix}$

$$\lambda_{2.} = \sqrt{\frac{105505 + 575}{32}}$$

$$\times_2 = \begin{bmatrix} -23 + \sqrt{105505} \\ -23 + \sqrt{105505} \\ 1 \end{bmatrix}$$

$$X_1 = \begin{bmatrix} -1.073 \end{bmatrix}$$

$$n_2 = 28.119$$

$$X_2 \supset \begin{bmatrix} 0.931 \\ 1 \end{bmatrix}$$

Since 22 has the highest value

x2 is the principle component.

$$X_{\text{scaled}} = \begin{bmatrix} -3.5 & -2.75 \\ -4.5 & -3.75 \\ 2.5 & 7.25 \\ 5.5 & -0.75 \end{bmatrix}$$

dimensionally reduced dala = Xscaled x PCI