62020
SEM: 06

ML

BARFI Page No.
Date

Reference

7/12

Q3]

A]

- iii]

① covariance is a measure of how changes in one variable are associated with changes in a second variable

② specifically, covariance measures the degree to which two variables are linearly associated. However, it is also often used informally as a general measure of how monotonically related the two variables are.

Sol^n:-

$$Xm = \frac{\Sigma X}{N} = 1.82$$

$$Ym = \frac{\Sigma Y}{N} = 1.91$$

Covariance matrix $C = \begin{bmatrix} Cxx & Cxy \\ Cyx & Cyy \end{bmatrix}$

$$Cxx = \frac{\Sigma (x-Xm)^2}{N-1} = 0.6165$$

$$Cxy = Cyx = \frac{\Sigma (x-Xm)(y-Ym)}{N-1} = 0.61544$$

$$Cyy = \frac{\Sigma (y-Ym)^2}{N-1} = 0.7165$$

$$C = \begin{bmatrix} 0.6165 & 0.61544 \\ 0.61544 & 0.7165 \end{bmatrix}$$

62020 ML
SEM: 06

BARFI Page No.
Date

Reference

8/12

$$C_{xx} = \sum \frac{(x-x_m)^2}{N-1}$$

$$= 0.6012$$

$$C_{xy} = \sum \frac{(x-x_m)(y-y_m)}{N-1} \quad \because \{C_{xy} = C_{yx}\}$$

$$= \frac{5.438}{9}$$

$$= 0.6042$$

$$C_{yy} = \sum \frac{(y-y_m)^2}{N-1}$$

$$= \boxed{0.7032}$$

$x_i - x_m$	$y - y_m$	$(x-x_m) \times (y-y_m)$
0.68	0.49	0.333
-1.32	-1.21	1.597
0.38	0.99	0.3762
0.08	0.29	0.023
1.28	1.09	1.395
0.48	0.79	0.379
0.18	-0.31	-0.055
-0.82	-0.81	0.664
-0.32	-0.31	0.0992
-0.62	-1.01	0.626
Σ		$\Sigma = 5.438$

$$\boxed{\therefore C = \begin{bmatrix} 0.6012 & 0.6042 \\ 0.6042 & 0.7032 \end{bmatrix}}$$

G2020
SEM: 06

ML

BARFI Page No.
Date

Rebecca

9/12

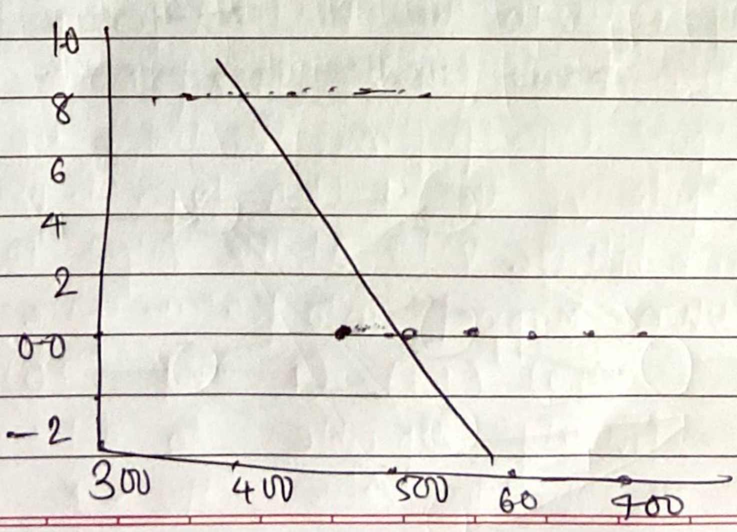
Q3]

A]
i]

① The logit function is the natural log of the odds that Y equals one of the categories.
② The log of the link function is what the logit function is.
③

$$Ln\left(\frac{P}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \cdots \beta_K X_K$$

④ The logit function is particularly popular because its results are relatively easy to interpret. But many of the others work just as well
⑤ Once we fit the model, we can then back-transform the estimated regression coefficients off of a log scale so that we can interpret the conditional effects of each X.
⑥ If we use Y as the outcome variable and tried to fit a line, it wouldn't be very good representation of the relationship.



62020
SEM:06 ML

BARFI Page No.
Date

Replace

10 / 12

(7) The graph shows an attempt to fit a line between x variable and a binary outcome Y you can see a relationship there - higher values of x are associated with more 0s and lower values of x have more 1's. But its not a linear relationship

(8) Spam detection is an example where we need to classify it as spam or no spam.

0 = Not spam
1 = Spam.

In order to apply logistic regression to the spam detection problem, sender of email, occurrence of words / phrases like "offr", "prize", "gift" are checked.

62020 ML

Sem: 06

Rebecca

11 / 12

Q3]

B]
ii]

① In single perceptron / Multi layer perceptron) we only have linear seperability because they are composed of input & output layers.

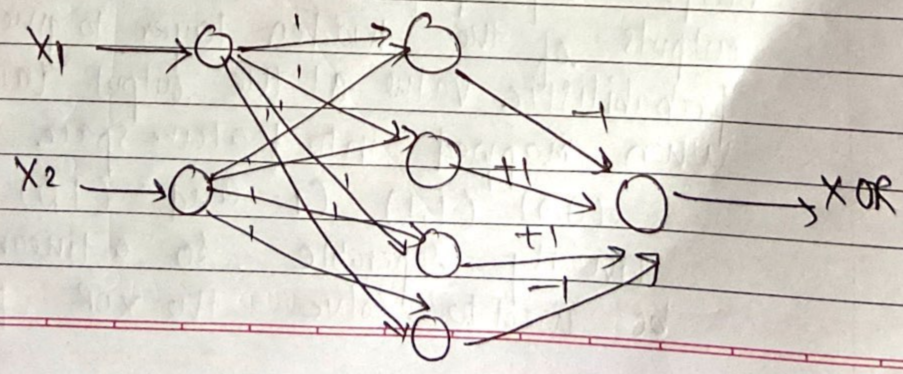
② Radial basis function, we define a receptor = t, we draw a confrontal map around the receptor and gaussian function are generally used for Radial Basis function. we define the radial distance
$r = ||x - t||$

③ Radial Basis Neural Network composes of input, hidden and output layer. RBNN is strictly limited to have exactly one hidden layer called as feature vector.

④ EOX function
we have 4 inputs and will not increase dimension at the feature vector hear. we select $\phi(x)$, we will have each receptor t

⑤ p = # of input features
M = # of transformed vector dimensions
each node in the hidden layer, performs a set of non-linear radian basis function.

⑥

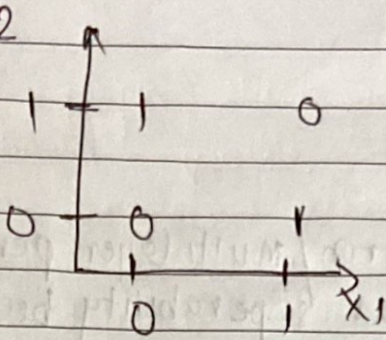


6 2020
8EM:06
Rebecca

ML

BARFI Page No.
Date

12/12



$p=2$

$\phi_1 \rightarrow t_1\ (0,0);\ \sigma_1 = 1$ $\quad \phi_1(x) = e^{-\frac{\|x-t_1\|^2}{2}}$

$\phi_2 \rightarrow t_2\ (0,1);\ \sigma_2 = 1$ $\quad \phi_2(x) = e^{-\frac{\|x-t_2\|^2}{2}}$

$\phi_3 \rightarrow t_3\ (1,0);\ \sigma_3 = 1$ $\quad \phi_3(x) = e^{-\frac{\|x-t_3\|^2}{2}}$

$\phi_4 \rightarrow t_4 = (1,1),\ \sigma_4 = 1$ $\quad \phi_4(x) = e^{-\|x-t_4\|^2/2}$

Input		ϕ_1	ϕ_2	ϕ_3	ϕ_4	$\Sigma W \cdot \phi$	output
0	0	1.0	0.6	0.6	0.4	−0.2	0
0	1	0.6	1.0	0.4	0.6	0.2	1
1	0	0.6	0.4	1.0	0.6	0.2	1
1	1	0.4	0.6	0.6	1.0	−0.2	0
		−1	+1	+1	−1		

- Only nodes in the hidden layer perform the radian basis ~~function~~ transformation function
- output layer performs the linear combination of the outputs of the hidden layer to give a final probabilistic value at the output layer.
- When mapped into feature space (0,0) (1,1) C1 and (1,0) (0,1) become linearly seperable so a linear classifier can be used to solve the XOR problem.