ARTIFICIAL INTELLIGENCE

Linear regression $y = \beta_0 + \beta_1 x$

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \qquad \beta_o = \bar{y} - \beta_1\bar{x}$$

1. Consider the following set of points: {(-2 , -1) , (1 , 1) , (3 , 2)}. Find the least square regression line for the given data points.

Solution

x_i	y_i	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})$ * $(y_i - \bar{y})$	$(x_i - \bar{x})^2$
-2	-1	-8/3	-5/3	40/9	64/9
1	1	1/3	1/3	1/9	1/9
3	2	7/3	4/3	28/9	49/9
$\bar{x} = 2/3$	$\bar{y} = 2/3$			Sum = 69/9	Sum = 114/9

$\beta_1 = 69/114$ and $\beta_o = \frac{2}{3} - \left(\frac{69}{114} * \frac{2}{3}\right) = \frac{2}{3}\left(1 - \frac{69}{114}\right) = \frac{45}{114} * \frac{2}{3} = \frac{90}{342}$

The linear regression line is y = $\frac{90}{342} + \frac{69}{114}\,x$

2. Find the least square regression line for the following set of data {(-1 , 0),(0 , 2),(1 , 4),(2 , 5)}

   Solution: 1.7x + 1.9

3. The values of y and their corresponding values of y are shown in the table below

x	0	1	2	3	4
y	2	3	5	4	6

   a) Find the least square regression line y = a x + b.
   b) Estimate the value of y when x = 10.

   Solution: a) 0.9 x + 2.2
   b) y = 11.2

Naïve Bayes classification

1.

Color	Type	Origin	Stolen
Red	Sports	Domestic	Yes
Red	Sports	Domestic	No
Red	Sports	Domestic	Yes
Yellow	Sports	Domestic	No
Yellow	Sports	Imported	Yes
Yellow	SUV	Imported	No
Yellow	SUV	Imported	Yes
Yellow	SUV	Domestic	No
Red	SUV	Imported	No
Red	Sports	Imported	Yes

Predict if a Red, SUV and Domestic vehicle will be stolen or not.

P(Stolen = Yes | X)

= P(X | Stolen = Yes) * P(Stolen = Yes)

= P(Color = Red | Stolen = Yes) * P(Type = SUV | Stolen = Yes) * P(Origin = Domestic | Stolen = Yes) * P(Stolen = Yes)

= 3/5 * 1/5 * 2/ 5 * 5/10

= 3/125

P(Stolen = No | X)

= P(X | Stolen = No) * P(Stolen = No)

= P(Color = Red | Stolen = No) * P(Type = SUV | Stolen = No) * P(Origin = Domestic | Stolen = No) * P(Stolen = No)

= 2/5 * 3/5 * 3/5 * 5/10

= 9/125

P(Stolen = No | X) > P(Stolen = Yes|X) therefore predicted class is Stolen = No

2.

Refund	Marital Status	Taxable Income	Evade
Yes	Single	125K	No
No	Married	100K	No
No	Single	70K	No
Yes	Married	120K	No
No	Divorced	95K	Yes
No	Married	60K	No
Yes	Divorced	220K	No
No	Single	85K	Yes
No	Married	75K	No
No	Single	90K	Yes

X = (Refund = No, Marital status = Married ,Income = 120K)

Solution: Class = No. P(Income = 120k| No) = 0.0072

$$P(A_i \mid c_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} e^{-\frac{(A_i-\mu_{ij})^2}{2\sigma_{ij}^2}}$$

Sample mean = 110

Sample variance = 2975

k-Nearest Neighbour

X1	X2	Y
4	21	0
5	19	0
10	24	1
11	25	1

Find Class (8,21)

Decision Tree Construction

X	Y	Z	C
1	1	1	I
1	1	0	I
0	0	1	II
1	0	0	II

Performance of classification

		Predicted	
		CAT	**DOG**
Actual	**CAT**	1984	447
	DOG	336	107

Calculate True Positive, True Negative, False Positive, False Negative, TPR, FPR, Sensitivity, Specificity, Precision, Recall and F1-score.