

NAME : Muhammad Yamin Afsidi

I.D : 12701

Subject : Data mining.

Teacher : Zain Shaukel.

(2)

Question No 1

Distinguish between classification and regression with the help of relevant scenarios.

Ans

Classification and regression are two major's prediction problems which are usually dealt in data mining. Predictive modeling is the technique of developing a model of function using the historic data to predict the new data. The significant difference between classification and regression is that classification maps the input data object to some discrete labels. on the other hand, regression maps the input data object to the continuous real values.

(2)

Definition of Classification;

Classification is the process of finding or discovering a model which helps in separating the data into multiple categorical classes. In classification the group membership of the problem is identified, which means the data categorize under different labels according to some parameters and then labels are predicted for the data.

Example,

Suppose, we want to predict the possibility of rain in some region on the basis of some parameters. Then there would be two labels rain and no rain under different region can be classified.

Definition of regression

Regression is the process of finding a model for distinguishing the data into continuous real values instead of using classes. Mathematically, with a regression problem, one is trying to find the function approximation with the minimum error deviation. In regression, the data numeric dependency is predicted to distinguish it.



Question No: 2

"SSN"	"Test "1"	Test "2"	Final	"Grade"
123-45-6787	100	83	49	"D"
123-12-1234	96	97	48	"D"
567-89-0123	60	40	44	"C"
087-65-4321	36	45	47	"B-"
456-78-9012	88	77	45	"A-"
234-56-7890	80	90	46	"C-"
789-67-8901	-1	4	43	"F"
632-79-9989	30	40	80	"B+"

⑤

Ans $P(123-46-4455 / \text{Grade "D"})$

$$P(\text{Test 1} = 95 / \text{Grade} = \text{D})$$

$$\mu_1 = \frac{100 + 96}{2} \Rightarrow 98$$

$$\sigma^2 = \frac{(100 - 98)^2 + (96 - 98)^2}{2 - 1}$$
$$= 8$$

$$\frac{1}{\sqrt{2\pi(8)}} e^{-\frac{(95 - 98)^2}{2(8)}}$$

$$= 0.063$$

$$P(\text{Test 2} = 31 / \text{Grade "D"})$$

$$\mu_2 = \frac{83 + 97}{2} \Rightarrow 90$$

Sigma

$$\sigma^2 = \frac{(83 - 90)^2 + (97 - 90)^2}{2 - 1}$$
$$= 98$$

$$\frac{1}{\sqrt{2\pi(98)}} e^{-\frac{(31 - 90)^2}{2(98)}}$$

$$= 0.029$$

$$P(\text{final} = 30 / \text{Grade} = \text{D})$$

$$\mu_1 = \frac{49 + 48}{2}$$

$$\Rightarrow 48.5$$

$$s^2 = \frac{(49 - 48.5)^2 + (48 - 48.5)^2}{2-1}$$

$$= 0.5$$

$$\frac{1}{\sqrt{2\pi(0.5)}} e^{-\frac{(38-48.5)^2}{2(0.5)}}$$

$$= 0.000000005 \cdot (9.155 \times 10^{-13})$$

$P(123.46.4455 / \text{Grade} = C)$

$P(\text{Test 1} = 85 / \text{Grade} = C)$

$$\mu_i = \frac{60}{2} \Rightarrow 30$$

$$s^2 = \frac{(60 - 30)^2}{2-1} \Rightarrow 900$$

$$\frac{1}{\sqrt{2\pi(900)}} e^{-\frac{(85-30)^2}{2(900)}}$$

$$= 0.013$$

$P(\text{Test 2} = 31 / \text{Grade} = C)$

$$\mu_i = \frac{40}{2} \Rightarrow 20$$

$$s^2 = \frac{(40 - 20)^2}{2-1} \Rightarrow 400$$

$$\frac{1}{\sqrt{2\pi(400)}} e^{-\frac{(31-20)^2}{2(400)}}$$

$$= 0.021$$

(7)

$$P(\text{Final} = 30 / \text{Grade C})$$

$$\mu_j = \frac{44}{2} = 22$$

$$s^2 = \frac{(44 - 22)^2}{2 - 1} = 484$$

$$\frac{1}{\sqrt{2\pi(484)}} e^{-\frac{(30-22)^2}{2(484)}}$$

$$= 0.018$$

$$= 0.013 \times 0.021 \times 0.018$$

$$= 0.0000049 \quad (4.914 \times 10^{-6})$$

$$P(123.46 - 44.55 / \text{Grade "B-"})$$

$$P(\text{Test 1} = 85 / \text{Grade "B-"})$$

$$\mu_j = \frac{36}{2} \Rightarrow 18$$

$$s^2 = \frac{(36 - 18)^2}{2 - 1} \Rightarrow 324$$

$$\frac{1}{\sqrt{2\pi(324)}} e^{-\frac{(85-18)^2}{2(324)}} \Rightarrow 0.0295$$

$$P(\text{Test 2} = 31 / \text{Grade "B-"})$$

$$\mu_j = \frac{45}{2} = 22.5$$

$$s^2 = (45 - 22.5)^2 \Rightarrow 506.25$$

$$= \frac{1}{\sqrt{2\pi(506.25)}} e^{-\frac{(31-22.5)^2}{2(506.25)}} \Rightarrow 0.0178$$

Date: / /

(8)

$$P(\text{SCA}) = 123.46.4455$$

$$P(\text{final} = 30 / \text{Grade "B-"})$$

$$\mu_i = \frac{47}{2} \Rightarrow 23.5$$

$$\begin{aligned} \sigma^2 &= \frac{(47 - 23.5)^2}{2-1} \\ &= 552.25 \end{aligned}$$

$$\frac{1}{\sqrt{2\pi(552.25)}} e^{-\frac{(30-23.5)^2}{2(552.25)}}$$

$$= 0.017$$

$$0.0245 \times 0.0178 \times 0.017$$

$$= 0.000007 \quad (7.4137 \times 10^{-6})$$

$$P(123.46.4455 / \text{Grade "A-"})$$

$$P(\text{Test 1} = 85 / \text{Grade "A-"})$$

$$\mu_i = \frac{88}{2} \Rightarrow 44$$

$$\sigma^2 = \frac{(88 - 44)^2}{2-1}$$

$$= 1936$$

$$\frac{1}{\sqrt{2\pi(1936)}} e^{-\frac{(85-44)^2}{2(1936)}}$$

$$= 0.009$$

Date: / /

9

$$P(\text{Test 2} = 31 / \text{Grade "A-"})$$

$$\mu_i = \frac{77}{2} \Rightarrow 38.5$$

$$\sigma^2 = \frac{(77 - 38.5)^2}{2-1}$$

$$= 1482.25$$

$$\frac{1}{\sqrt{2\pi(1482.25)}} e^{-\frac{(31-38.5)^2}{2(1482.25)}}$$

$$= 0.0103$$

$$P(\text{Final} = 30 / \text{Grade "A-"})$$

$$\mu_i = \frac{45}{2} \Rightarrow 22.5$$

$$\sigma^2 = \frac{(45 - 22.5)^2}{2-1}$$

$$= 506.25$$

$$\frac{1}{\sqrt{2\pi(506.25)}} e^{-\frac{(30-22.5)^2}{2(506.25)}}$$

$$= 0.0178$$

$$0.009 \times 0.0103 \times 0.0178$$

$$= 0.000001 \left(1.65006 \times 10^{-6} \right)$$

$$P(123.46.4455 / \text{Grade "C-"})$$

$$P(\text{Test 1} = 85 / \text{Grade "C-"})$$

$$\mu_i = \frac{80}{2} \Rightarrow 40$$

$$s^2 = \frac{(80 - 40)^2}{2-1} \Rightarrow 1600$$

$$\frac{1}{\sqrt{2\pi(1600)}} e^{-\frac{(85-40)^2}{2(1600)}}$$

$$= 0.0101$$

$$P(\text{Test 2} = 31 / \text{Grade "C-"})$$

$$\mu_i = \frac{90}{2} \Rightarrow 45$$

$$s^2 = \frac{(90 - 45)^2}{2-1}$$

$$= 2025$$

$$\frac{1}{\sqrt{2\pi(2025)}} e^{-\frac{(31-45)^2}{2(2025)}}$$

$$= 0.008$$

$$P(\text{Final} = 30 / \text{Grade "C-"})$$

$$\mu_i = \frac{46}{2} = 23$$

$$s^2 = \frac{(46 - 23)^2}{2-1}$$

$$= 529$$

Date: / /

(11)

~~0.0174~~

$$\frac{1}{\sqrt{2\pi(529)}} e^{\frac{(30-25)^2}{2(29)}}$$

$$= 0.0174$$

$$0.0101 \times 0.008 \times 0.0174$$

$$= 0.000001 (1.40592 \times 10^{-6})$$

P(123.46.4455 / Grade F)

P(Test 1 = 85 / Grade F)

$$\mu_i = \frac{-1}{2} \Rightarrow -0.5$$

$$\sigma^2 = \frac{(-1 - 0.5)^2}{2-1}$$

$$= 2.25$$

$$\frac{1}{\sqrt{2\pi(2.25)}} e^{\frac{(85+0.5)^2}{2(2.25)}}$$

474.59

P(Test 2 = 31 / Grade = f)

$$\mu_i = \frac{4}{2} \Rightarrow 2$$

$$\sigma^2 = \frac{(4-2)^2}{2-1}$$

$$= 4$$

(12)

$$\frac{1}{\sqrt{2\pi(4)}} e^{-\frac{(31-2)}{2(4)}}$$

5.2255

$P(\text{Final} = 21 / \text{Grade} = F)$

$$\mu_1 = \frac{43}{2} \Rightarrow 21.5$$

$$b^2 = \frac{(43 - 21.5)^2}{2-1}$$
$$= 462.25$$

$$\frac{1}{\sqrt{2\pi(462.25)}} e^{-\frac{(31-21.5)}{2(462.25)}}$$
$$= 0.0187$$

$$\frac{474.59 \times 5.2255 \times 0.0187}{46.37}$$

$P(123.46.4455 / \text{Grade "B+"})$

$P(\text{Test 1} = 85 / \text{Grade "B+"})$

$$\mu_1 = \frac{30}{2} \Rightarrow 15$$

$$b^2 = \frac{(30 - 15)^2}{2-1}$$

$$= 225$$

13

$$\frac{1}{\sqrt{2\pi(225)}} e^{-\frac{(85-15)^2}{2(225)}}$$

0.0310

$P(\text{Test 2} = 31 \mid \text{Grade "B+"})$

$$\mu_i = \frac{40}{2} \Rightarrow 20$$

$$b^2 = \frac{(40-20)^2}{2-1}$$

$$= 400$$

$$\frac{1}{\sqrt{2\pi(400)}} e^{-\frac{(31-20)^2}{2(400)}}$$

= 0.0202

$P(\text{Final} = 30 \mid \text{Grade "B+"})$

$$\mu_i = \frac{50}{2} \Rightarrow 25$$

$$b^2 = \frac{(50-25)^2}{2-1}$$

$$= 625$$

$$\frac{1}{\sqrt{2\pi(625)}} e^{-\frac{(30-25)^2}{2(625)}}$$

= 0.0160

$$0.0310 \times 0.0202 \times 0.0160$$

$$= 0.0001 \quad (1.06192 \times 10^{-4})$$

We have a greater probability
for Grade F.

So Grade is F.

←————→
END

Question No: 3

Ans . Naive Bayes :

Fruit	Yellow	Sweet	long	Total
Mango	350	450	0	800
Banana	400	300	350	1050
others	50	100	50	200
Total	800	850	400	2050

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

prob. (A) when B is True

i) $P(x/\text{mango}) =$

$$P(\text{yellow/mango}) = P(\text{mango/yellow}) \cdot P(\text{yellow})$$

(15)

$$= \frac{350}{800} \cdot \frac{800}{1200}$$

$$= \frac{350}{650}$$

$$P(\text{yellow/mango}) = 0.53$$

$$P(S/\text{Mango}) = \frac{P(\text{mango}/\text{sweet}) \cdot P(\text{sweet})}{P(\text{mango})}$$

$$= \frac{450}{850} \cdot \frac{850}{650}$$

$$P(\text{sweet/mango}) = 0.69$$

$$P(L/\text{mango}) = \frac{P(\text{mango}/\text{long}) \cdot P(\text{long})}{P(\text{mango})}$$

$$= \frac{0}{400} \cdot \frac{400}{1200}$$

$$P(L/\text{mango}) = 0$$

$$P(X/\text{mango}) = P(Y/M) \cdot P(S/M) \cdot P(L/M)$$

$$= 0$$

(ii) Banana = $P(Y/B) = 1$

$$P(S/B) = 0.75$$

$$P(L/B) = 0.875$$

$$P(X/\text{Banana}) = P(Y/B) \cdot P(S/B) \cdot P(L/B)$$

$$= 1 \times 0.75 \times 0.875$$

$$= 0.65$$

(8)

others

$$P(Y/O) = 0.33$$

$$P(S/O) = 0.66$$

$$P(L/O) = 0.33$$

$$P(X/Others) = P(Y/O) \cdot P(S/O) \cdot P(L/O)$$

$$= 0.33 \times 0.66 \times 0.33$$

$$= 0.072$$

The Type of fruit class is
belong to bananas

Ans

