170200107044

Page No.
Date / /

Assignment - 6

Q:1 Given a decision tree, you have the option of (a) Converting the decision tree to rules and then prunning the resulting rules, or (b) prunning the decision tree then converting pruned tree to rules. What advantage does (a) have over (b)?

=> with method (b) we would prune subtree of decision tree by removing the subtree Completely.

=> However, with method (a) prunning a rule may remove any precondition of it which is less restrictive then method (b).

Q:3 Elaborate the steps of ID_3 algorithm for generating Decision trees.

=> ID3 stands for 'Iterative Dichotomiser 3".
=> ID3 uses a top-down greedy approach to build a decision tree.
=> Most generally ID3 is only used for classification problems with nominal feature only.
=> ID3 uses information Gain as its attribute selection measure.
=> Let Node N represent or hold the tuples of Partition D.
=> The attribute with the highest information gain is

17020010704

Page No.
Date / /

choose as the spliting attribute for node N.

=) This attribute minimizes the information needed to classify the tuples in the resulting partitions and reflects the least randomness or impurity in these partitions.

=) The expected information needed to classify a tuple in D is given by

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2 (P_i)$$

where P_i is the probability that an arbitrary tuple D belongs to class C_i and is estimated

$$P_i = |C_{i,D}| / |D|$$

=) A log function to the base 2 is used, because the information is encoded in bits.

=) Info(D) is just the average amount of information needed to identify the class label of a tuple in D.

=) Info (D) is also known as the entropy of D.

* ID3 steps -

step I : Calculate the information Gain of each feature.

170200107044

Page No.
Date / /

Step 2 : Considering that all rows don't belong to the same class, split the dataset S into subsets using the feature for which the Information Gain is maximum.

Step 3 : Make a Decision tree node using the feature with the maximum information gain.

Step 4 : If all rows belongs to the same class, make the current node as a least node with the class at its label.

Step 5 : Repeat for the remaining feature until we run out of all feature, or the decision tree has all leaf nod

Q:04 Why naive Bayesian classification is called "naive" ?
- Apply Naive Bayesian classification on following dataset to classify given tuple.

X = (age <=30, Income = medium, student = yes, credit rating = fair)

1702001070044

Page No.

Date / /

Age	income	student	credit-rating	buys-computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31...40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31...40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31...40	high	yes	fair	yes
>40	medium	no	excellent	no

=) The given tuple have attribute like,

age <=30

Income = medium

student = yes

credit-rating = fair

170200107044

Page No.

Date / /

	class (9/14) Yes	class (5/14) NO
Attribute	**Age**	
<=30	2/9	3/5
31...40	4/9	0
>40	3/9	2/5
Attribute	**Income**	
Low	3/9	1/5
medium	4/9	2/5
High	2/9	2/5
Attribute	**Student**	
NO	3/9	4/5
Yes	6/9	1/5
Attribute	**Credit-rating**	
Fair	6/9	2/5
Excellent	3/9	5/5

$\Rightarrow$ P(C_i) : p(buys_ computer = "yes") = 9/14 = 0.643

P(buys _ computer = "no") = 5/14 = 0.357

$\rightarrow$ Compute p(X/C_i) for each class

class yes

○> $P(\text{age} \text{"} <=30 \text{"} \mid \text{buys\_computer} = \text{"yes"}) = 2/9 = 0.222$

$P(\text{income} = \text{"medium"} \mid \text{buys\_computer} = \text{"yes"}) = 4/9 = 0.444$

$P(\text{student} = \text{"yes"} \mid \text{buys\_computer} = \text{"yes"}) = 6/9 = 0.667$

$P(\text{credit\_rating} = \text{"fair"} \mid \text{buys\_computer} = \text{"yes"}) = 6/9 = 0.667$

class no

○> $P(\text{age} \text{"} <=30 \text{"} \mid \text{buys\_computer} = \text{"no"}) = 3/5 = 0.5$

$P(\text{income} = \text{"medium"} \mid \text{buys\_computer} = \text{"no"}) = 2/5 = 0.4$

$P(\text{student} = \text{"yes"} \mid \text{buys\_computer} = \text{"no"}) = 1/5 = 0.2$

$P(\text{credit\_rating} = \text{"fair"} \mid \text{buys\_computer} = \text{"no"}) = 2/5 = 0.4$

$X = (\text{age} <= 30, \text{income} = \text{medium}, \text{student} = \text{yes}, \text{credit\_rating} = \text{fair})$

$P(X/C_i)$: $P(X / \text{buys\_computer} = \text{"yes"})$

$= 0.222 * 0.444 * 0.667 * 0.667 = 0.044$

: $P(X / \text{buys\_computer} = \text{"no"})$

$= 0.6 * 0.4 * 0.2 * 0.4 = 0.019$

$P(X \mid C_i) * P(C_i)$: $P(X \mid \text{buys\_computer} = \text{"yes"}) *$

$(\text{buys\_computer} = \text{"yes"}) = 0.028$

: $P(X \mid \text{buys\_computer} = \text{"no"}) * P(\text{buys\_comp} = \text{"no"})$

$= 0.007$

=> Therefore, x belongs to class (computer_buys = yes)

170200107044

Page No.

Date / /

Q:5 Calculate the weights using neural network single layer perceptron model. Three inputs are x_0, x_1, x_2 bias and weights are as follows:

$W_1(0) = 30$ $\quad W_2(0) = 300$

$b(a) = 50$ $\quad \eta = 0.01,$ $\quad x_0 = +1$

Activation function is

$sgn(x) = +1$; if $x >= 0$

$sgn(x) = -1$; if $x < 0$

a) Calculate the x_2 for $x_1 = 100$ and 200

$\Rightarrow 30x_1 + 300x_2 + 50 = 0$

$\therefore x_2 = \dfrac{-50 - 30x_1}{300}$ For $x_1 = 100$

$= \dfrac{-500 - 3000}{300}$

$= -10.1667$

for $x_2 = 200$

$x_2 = \dfrac{-50 - 6000}{300}$

$= -20.1667$

170200107044

Page No.

Date / /

b) For bias $b(0) = -1230$ recalculate the weights w_1 and w_2.

$\Rightarrow$ For $b = -1230$, $w_1 = 30$, $w_2 = 300$

(100, −10.1667)

(200, −20.1667)

$x(0) = [1, 100, -10.1667]^T$

$x(0) = (-1230, 30, 300]^T$

$y(0] = sgn(w^T(0) x (0))$

$= sgn(-1230 + 3000 * -3050.01)$

$= -1 \neq d(0)$

we need to recalculate weights

$(x)(1) = [+1, 200, -20.1667]^T$

$w(1) = [-1230, 30, 300]^T$

$y(1) = sgn(w^T(1)^T x(1))$

$= sgn(-1230 + 6000 - 6050.01)$

$= -1 \neq d(0)$

we need to recalculate weights

$w(n+1) = w(n) + \eta [d(n) - y(n)] x(n)$

$w(1) = [-1230, 30, 300]^T$

$x(1) = [+1, 200, -20.1667]^T$

$d(1) = +1$, $y(1) = -1$, $\eta = 0.01$.

170200107044

Page No.
Date / /

$W(1+1) = W(2) = (-1230, 30, 300)^T + (0.012 \; [+1 \; +1]$
$[+1, 200, -20, 1667]$
$= [-1230, 30, 300]^T + [0.02, 4, -0.4033]^T$
$W(1) = [-1229.98; 34, 299.597]$

hence $w_1 = 34$
$w_2 = 299.597$

Q:2 Define linear and nonlinear regression using figures. Calculate the value of y for x=100 based on Linear regression prediction method.

x	y	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i - \bar{x})^2$
4	390	-4.75	-165	783.75	22.5625
9	580	0.25	25	6.25	0.0625
10	650	1.75	95	166.25	3.0625
14	730	5.25	175	918.75	27.5625
4	410	-4.75	-145	688.75	22.5625
7	530	-1.75	-25	43.75	3.0625
12	600	3.25	45	146.25	10.5625
22	790	13.25	235	3113.75	175.5625
1	350	-7.75	-205	1588.75	60.0625
3	400	-5.75	-155	891.25	33.0625
8	590	-0.75	35	-26.25	0.5625
11	640	2.25	85	191.25	5.0625

170200107044

Page No.

Date / /

⇒ For linear regression

$y = b_0 + b_1 x$ where $b_1 = \frac{\Sigma [(x_i - \bar{x})(y_i - \bar{y})]}{\Sigma (x_i - \bar{x})^2}$ $b_0 = \bar{y} - \bar{b} * \bar{x}$

$\bar{x} = \frac{\Sigma x}{n} = \frac{105}{12} = 8.75$ $\bar{y} = \frac{\Sigma y}{n} = \frac{6660}{12} = 555$

$b_1 = \frac{[\Sigma (x_i - \bar{x})(y_i - \bar{y})]}{\Sigma (x_i - \bar{x})^2}$

$= \frac{8512.5}{363.75}$

$= 23.4020$

$b_0 = \bar{y} - b_1 * \bar{x}$

$= 55 - 23.4020 \times 8.75$

$= 555 - 204.7675$

$= 350.2325$

For $x = 100$ $y = b_0 + b_1 x$

$= 350.2325 + (23.4020) \times 100$

$= 350.2325 + 2340.2$

$= 2690.4325$

simple linear regression relates two variables (x & y) with a straight line ($y = mx + b$), while non-linear regression relates the two variable (x & y) in a nonlinear (curved) relationship.

Nonlinear equation is different for every curve.