

CT2 Questions with Answers.docx - Microsoft Word

Home Insert Page Layout References Mailings Review View Add-Ins

Clipboard Font Paragraph Styles Editing

Machine Learning Techniques and Practices – CT2 Date: 09.04.2022

Duration: 1 Hr Time: 10:00 – 11:00 AM

Total Marks: 20

**Note:** Some MCQs may have multiple answers. In such case, you have to write all the correct choices. Otherwise, no marks will be provided for that question.

1. (i) Write the various methods/values that can be utilized to replace the NULL values.  
(ii) Write the various function(s) that can be utilized to define the datatype of a particular feature. (1 + 1 = 2 M)

Default, 0, Mean, Median, Mode  
astype(); dtype()

2. If a model performs well for training dataset and produces a lot of errors during testing time, then the model is said to be (1 M)

(a) Underfitting (b) Overfitting (c) Perfect fit (d) Having High Bias  
and Low Variance (e) Having Low Bias and High Variance

3. What does the term “k” signify in k-means Algorithm? (1 M)

(a) Only the no. of samples that has to be considered  
(b) Only the no. of clusters to be formed  
(c) Both (a) and (b)

Page: 1 of 3 Words: 915 English (India) 100%



CT2 Questions with Answers.docx - Microsoft Word

Home Insert Page Layout References Mailings Review View Add-Ins

Clipboard: Paste, Cut, Copy, Format Painter

Font: Calibri (Body), 11, Bold, Italic, Underline, Text Color, Background Color

Paragraph: Bulleted List, Numbered List, Decrease Indent, Increase Indent, Line and Paragraph Spacing, Paragraph Style

Styles: Normal, No Spacing, Heading 1, Heading 2, Title, Subtitle, Subtle Emphasis, Emphasis

Find, Replace, Select, Editing

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

2. If a model performs well for training dataset and produces a lot of errors during testing time, then the model is said to be **(1 M)**

(a) Underfitting and Low Variance (b) Overfitting and High Variance (c) Perfect fit (d) Having High Bias

3. What does the term "k" signify in k-means Algorithm? **(1 M)**

(a) Only the no. of samples that has to be considered (b) Only the no. of clusters to be formed (c) Both (a) and (b) (d) None of the above

4. Name the three ways by which the ensemble models try to consolidate the predicted outcomes. **(2 M)**

Maximum Voting, Averaging, Weighted Average

5. Which of the following information is correct about stratified cross-validation **(1 M)**

(a) Splits the data into "k" random folds (b) Split the data into "k" folds ensuring that the number of representation from each class is similar in every fold

1 | Page

Dr. R. Bala Krishnan, Assistant Professor, Department of CSE

Page: 1 of 3 Words: 915 English (India) 100%

+39 RN TA HV RS PY SS PG MB BK

Pushpanjali Yerranagu Rakesh Chauhan Evin Xavier Surudhi Sakthivel Pavithra Gurushankar Madhavan B Bala Krishnan

02:20 03:19 / 35:44







1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

(d) None of the above

6. Write the names of the three distance metrics that are used to evaluate the distance between the data points in ML algorithms. **(2 M)**

**Euclidean Distance; Manhattan Distance; Minkowski Distance**

7. Write the name of the plot which is used to identify the value of "k" in k-Means clustering algorithm. **(1 M)**

**Elbow Plot**

8. Assume that you are going to build a random forest comprising of two models internally. Now, I want to use the following dataset to train my models. Identify the possible training datasets for each model by considering both row and feature sampling with replacement technique. **(5 M)**

Shape	Diag	# of Diag	Length (in cm)	Height (in cm)	Target Class
Rectangle	Yes	2	100	10	1
Rectangle	Yes	2	1000	100	1
Square	Yes	2	100	100	0
Square	No	2	10000	10000	0
Triangle	No	0	100	1000	0
Triangle	Yes	0	50	1000	0
Circle	No	0	100	100	1
Circle	Yes	0	50	50	1

**Train Dataset**



CT2 Questions with Answers.docx - Microsoft Word

Home Insert Page Layout References Mailings Review View Add-Ins

Clipboard: Paste, Cut, Copy, Format Painter

Font: Calibri (Body), 11, Bold, Italic, Underline, Text Color, Background Color

Paragraph: Bulleted List, Numbered List, Decrease Indent, Increase Indent, Line and Paragraph Spacing, Paragraph Style

Styles: Normal, No Spacing, Heading 1, Heading 2, Title, Subtitle, Subtle Emphasis, Emphasis

Find, Replace, Select, Editing

Video settings

Shape, Diag, # of Diag, Length (in cm), Height (in cm), Target Class

Train Dataset

Test Dataset

Shape	Diag	# of Diag	Length (in cm)	Height (in cm)	Target Class
Rectangle	Yes	2	100	10	1
Rectangle	Yes	2	1000	100	1
Square	Yes	2	100	100	0
Square	No	2	10000	10000	0
Triangle	No	0	100	1000	0
Triangle	Yes	0	50	1000	0
Circle	No	0	100	100	1
Circle	Yes	0	50	50	1
Diamond	Yes	2	100	100	0
Rectangle	Yes	2	100	5	1

Shape	Length	Target Class
Rectangle	100	1
Rectangle	100	1
Rectangle	1000	1
Square	100	0
Triangle	100	0
Triangle	50	0
Circle	50	1

Shape	Height	Target Class
Rectangle	10	1
Rectangle	10	1
Square	100	0
Triangle	100	0
Triangle	1000	0
Circle	100	1
Circle	50	1

Page: 02:20 Words: 915 English (United States) 08.19

00 +49 16 16 08:18 / 35:44

Jeremiah Thomas Rakesh Chauhan Evin Xavier Surudha Sakthivel PG Bala Krishnan

1x

Dr. R. Bala Krishnan, Assistant Professor, Department of CSE

9. Suppose you are given the following set of data with three Boolean input variables a, b, and c, and a single Boolean output variable K. (5 M)

a	b	c	K
1	0	1	1
1	1	1	1
0	1	1	0
1	1	0	0
1	0	1	0
0	0	0	1
0	0	0	1
0	0	1	1

Using Naive Bayes Classifier, identify the value of K for the given input values, a = 1, b = 1 and c = 0?

[Hint: Consider, Alpha = 1]



c, and a single Boolean output variable K.

(5 M)

a	b	c	K
1	0	1	1
1	1	1	1
0	1	1	0
1	1	0	0
1	0	1	0
0	0	0	1
0	0	0	1
0	0	1	1

Using Naive Bayes Classifier, identify the value of K for the given input values, a = 1, b = 1 and c = 0?

[Hint: Consider, Alpha = 1]



a	b	c	K
1	0	1	1
1	1	1	1
0	0	0	1
0	0	0	1
0	0	1	1

Total = 8; No. of 1 = 5

a = 0 → 3; a = 1 → 2; b = 0 → 4; b = 1 → 1; c = 0 → 2; c = 1 → 3

a	b	c	K
0	1	1	0
1	1	0	0
1	0	1	0

Total = 8; No. of 0 = 3

a = 0 → 1; a = 1 → 2; b = 0 → 1; b = 1 → 2; c = 0 → 1; c = 1 → 2

Set: Alpha = 1



[Hint: Consider, Alpha = 1]

a	b	c	K
1	0	1	1
1	1	1	1
0	0	0	1
0	0	0	1
0	0	1	1

Total = 8; No. of 1 = 5

$a = 0 \rightarrow 3; a = 1 \rightarrow 2; b = 0 \rightarrow 4; b = 1 \rightarrow 1; c = 0 \rightarrow 2; c = 1 \rightarrow 3$

Set: Alpha = 1

$a = 0 \rightarrow 4; a = 1 \rightarrow 3; b = 0 \rightarrow 5; b = 1 \rightarrow 2; c = 0 \rightarrow 3; c = 1 \rightarrow 4$

Total = 21

$P(K = 1) = 5 / 8 = 0.625$

$P(a = 1 | K = 1) = 3 / 21 = 0.143$

$P(b = 1 | K = 1) = 2 / 21 = 0.095$

$P(c = 0 | K = 1) = 4 / 21 = 0.19$

New Value =  $a = 1, b = 1$  and  $c = 0$

$\Rightarrow P(K = 1) \times P(a = 1 | K = 1) \times P(b = 1 | K = 1) \times P(c = 0 | K = 1)$

$\Rightarrow 0.625 \times 0.143 \times 0.095 \times 0.19 = 1.61 \times 10^{-3}$

Hence, for the values ( $a = 1, b = 1$  and  $c = 0$ ),  $K = 0$

a	b	c	K
0	1	1	0
1	1	0	0
1	0	1	0

Total = 8; No. of 0 = 3

$a = 0 \rightarrow 1; a = 1 \rightarrow 2; b = 0 \rightarrow 1; b = 1 \rightarrow 2; c = 0 \rightarrow 1; c = 1 \rightarrow 2$

Set: Alpha = 1

$a = 0 \rightarrow 2; a = 1 \rightarrow 3; b = 0 \rightarrow 2; b = 1 \rightarrow 3; c = 0 \rightarrow 2; c = 1 \rightarrow 3$

Total = 15

$P(K = 0) = 0.375$

$P(a = 1 | K = 0) = 3 / 15 = 0.2$

$P(b = 1 | K = 0) = 3 / 15 = 0.2$

$P(c = 0 | K = 0) = 3 / 15 = 0.2$

New Value =  $a = 1, b = 1$  and  $c = 0$

$\Rightarrow P(K = 0) \times P(a = 1 | K = 0) \times P(b = 1 | K = 0) \times P(c = 0 | K = 0)$

$\Rightarrow 0.375 \times 0.2 \times 0.2 \times 0.2 = 3 \times 10^{-3}$



[Hint: Consider, Alpha = 1]

a	b	c	K
1	0	1	1
1	1	1	1
0	0	0	1
0	0	0	1
0	0	1	1

Total = 8; No. of 1 = 5

$a = 0 \rightarrow 3; a = 1 \rightarrow 2; b = 0 \rightarrow 4; b = 1 \rightarrow 1; c = 0 \rightarrow 2; c = 1 \rightarrow 3$

Set: Alpha = 1

$a = 0 \rightarrow 4; a = 1 \rightarrow 3; b = 0 \rightarrow 5; b = 1 \rightarrow 2; c = 0 \rightarrow 3; c = 1 \rightarrow 4$

Total = 21

$P(K=1) = 5 / 8 = 0.625$

$P(a=1 | K=1) = 3 / 21 = 0.143$

$P(b=1 | K=1) = 2 / 21 = 0.095$

$P(c=0 | K=1) = 4 / 21 = 0.19$

New Value =  $a = 1, b = 1$  and  $c = 0$

$\Rightarrow P(K=1) \times P(a=1 | K=1) \times P(b=1 | K=1) \times P(c=0 | K=1)$

$\Rightarrow 0.625 \times 0.143 \times 0.095 \times 0.19 = 1.61 \times 10^{-3}$

Hence, for the values ( $a = 1, b = 1$  and  $c = 0$ ): **K = 0**

a	b	c	K
0	1	1	0
1	1	0	0
1	0	1	0

Total = 8; No. of 0 = 3

$a = 0 \rightarrow 1; a = 1 \rightarrow 2; b = 0 \rightarrow 1; b = 1 \rightarrow 2; c = 0 \rightarrow 1; c = 1 \rightarrow 2$

Set: Alpha = 1

$a = 0 \rightarrow 2; a = 1 \rightarrow 3; b = 0 \rightarrow 2; b = 1 \rightarrow 3; c = 0 \rightarrow 2; c = 1 \rightarrow 3$

Total = 15

$P(K=0) = 0.375$

$P(a=1 | K=0) = 3 / 15 = 0.2$

$P(b=1 | K=0) = 3 / 15 = 0.2$

$P(c=0 | K=0) = 3 / 15 = 0.2$

New Value =  $a = 1, b = 1$  and  $c = 0$

$\Rightarrow P(K=0) \times P(a=1 | K=0) \times P(b=1 | K=0) \times P(c=0 | K=0)$

$\Rightarrow 0.375 \times 0.2 \times 0.2 \times 0.2 = 3 \times 10^{-3}$