

第一題

weka.gui.GenericObjectEditor

weka.filters.unsupervised.instance.Resample

About

Produces a random subsample of a dataset using either sampling with replacement or without replacement.

More

Capabilities

debug False

doNotCheckCapabilities False

invertSelection False

noReplacement False

randomSeed 1

sampleSizePercent 60.0

Open... Save... OK Cancel

Filter

Choose Resample -S 1 -Z 60.0

Current relation

Relation: customer_churn-weka.filters.unsupervised.insta... Attributes: 20

Instances: 3384 Sum of weights: 3384

第二題(先轉 nominal 後才截圖)

Selected attribute

Name: Churn	Distinct: 2	Type: Nominal
Missing: 0 (0%)		Unique: 0 (0%)

No.	Label	Count	Weight
1	0	2816	2816
2	1	568	568

weka.gui.GenericObjectEditor

X

weka.filters.unsupervised.attribute.ReplaceMissingValues

About

Replaces all missing values for nominal and numeric attributes in a dataset with the modes and means from the training data.

More

Capabilities

debugFalse

doNotCheckCapabilitiesFalse

ignoreClassFalse

Open...Save...OKCancel

Filter

ChooseReplaceMissingValuesApplyStop

Current relation
Relation: customer_churn-weka.filters.unsupervised.insta...
Instances: 3384Attributes: 20
Sum of weights: 3384

Attributes

AllNoneInvertPattern

No.	Name
1	<input checked="" type="checkbox"/> i>=CustomerID
2	<input checked="" type="checkbox"/> Churn
3	<input checked="" type="checkbox"/> Tenure
4	<input checked="" type="checkbox"/> PreferredLoginDevice
5	<input checked="" type="checkbox"/> CityTier
6	<input checked="" type="checkbox"/> WarehouseToHome
7	<input checked="" type="checkbox"/> PreferredPaymentMethod
8	<input checked="" type="checkbox"/> Gender
9	<input checked="" type="checkbox"/> HourSpendOnApp
10	<input checked="" type="checkbox"/> NumberOfDeviceRegistered
11	<input checked="" type="checkbox"/> PreferredOrderCat
12	<input checked="" type="checkbox"/> SatisfactionScore
13	<input checked="" type="checkbox"/> MaritalStatus
14	<input checked="" type="checkbox"/> NumberOfAddress
15	<input checked="" type="checkbox"/> Complain
16	<input checked="" type="checkbox"/> OrderAmountHikeFromlastYear
17	<input checked="" type="checkbox"/> CouponUsed
18	<input checked="" type="checkbox"/> OrderCount

Remove

Selected attribute

Name: i>=CustomerID
Missing: 0 (0%)Distinct: 2547Type: Numeric
Unique: 1856 (55%)

Statistic	Value
Minimum	50006
Maximum	55630
Mean	52826.453
StdDev	1633.54

Class: CashbackAmount (Num)Visualize All

Class Value	Frequency
251	218
218	210
210	221
221	235
235	206
206	242
242	192
192	219
219	248
248	229
229	231
231	229
229	227
227	226
226	

Choose

ReplaceMissingValues

Apply

Stop

Current relation

Relation: customer_churn-weka.filters.unsupervised.insta...

Attributes: 20

Instances: 3384

Sum of weights: 3384

Attributes

All

None

Invert

Pattern

No.	Name
1	<input type="checkbox"/> ?CustomerID
2	<input type="checkbox"/> Churn
3	<input checked="" type="checkbox"/> Tenure
4	<input type="checkbox"/> PreferredLoginDevice
5	<input type="checkbox"/> CityTier
6	<input type="checkbox"/> WarehouseToHome
7	<input type="checkbox"/> PreferredPaymentMode
8	<input type="checkbox"/> Gender
9	<input type="checkbox"/> HourSpendOnApp
10	<input type="checkbox"/> NumberOfDeviceRegistered
11	<input type="checkbox"/> PreferredOrderCat
12	<input type="checkbox"/> SatisfactionScore
13	<input type="checkbox"/> MaritalStatus
14	<input type="checkbox"/> NumberOfAddress
15	<input type="checkbox"/> Complain
16	<input type="checkbox"/> OrderAmountHikeFromlastYear
17	<input type="checkbox"/> CouponUsed
18	<input type="checkbox"/> OrderCount
19	<input type="checkbox"/> ...

Remove

Selected attribute

Name: Tenure

Missing: 0 (0%)

Distinct: 35

Type: Numeric

Unique: 2 (0%)

Statistic	Value
Minimum	0
Maximum	51
Mean	10.251
StdDev	8.386

Class: CashbackAmount (Num)

Visualize All

Tenure Value	Frequency
0	708
1	228
2	223
3	259
4	339
5	332
6	202
7	186
8	150
9	166
10	118
11	102
12	90
13	66
14	75
15	80
16	0
17	0
18	0
19	0
20	0
21	0
22	0
23	0
24	0
25	0
26	0
27	0
28	0
29	0
30	0

No.	Name
1	<input checked="" type="checkbox"/> CustomerID
2	<input type="checkbox"/> Churn
3	<input type="checkbox"/> Tenure
4	<input type="checkbox"/> PreferredLoginDevice
5	<input type="checkbox"/> CityTier
6	<input type="checkbox"/> WarehouseToHome
7	<input type="checkbox"/> PreferredPaymentMode
8	<input type="checkbox"/> Gender
9	<input type="checkbox"/> HourSpendOnApp
10	<input type="checkbox"/> NumberOfDeviceRegistered
11	<input type="checkbox"/> PreferredOrderCat
12	<input type="checkbox"/> SatisfactionScore
13	<input type="checkbox"/> MaritalStatus
14	<input type="checkbox"/> NumberOfAddress
15	<input type="checkbox"/> Complain
16	<input type="checkbox"/> OrderAmountHikeFromlastYear
17	<input type="checkbox"/> CouponUsed
18	<input type="checkbox"/> OrderCount

Remove

轉 Churn(target 欄位)nominal ，原為 numeric

The screenshot shows the Weka GUI with the 'NumericToNominal' filter selected in the 'Filter' list. The 'Current relation' is 'customer_churn-weka.filters.unsupervised.insta...'. The 'Selected attribute' is 'Churn', which is of type 'Nominal' with 2 distinct values. A table shows the distribution of the 'Churn' attribute:

No.	Label	Count	Weight
1	0	2816	2816
2	1	568	568

The 'Churn' attribute is highlighted in the 'Attributes' list. A blue circle highlights the 'Churn' attribute in the list and the 'NumericToNominal' filter in the 'Filter' list.

最後將 Churn 欄位移到最後一欄(因為 experimenter 預設最後一欄為 target)

The screenshot shows the Weka GUI with the 'Reorder' filter selected in the 'Filter' list. The 'Current relation' is 'customer_churn-weka.filters.unsupervised.insta...'. The 'Selected attribute' is 'Churn', which is of type 'Nominal' with 2 distinct values. A dialog box titled 'weka.gui.GenericObjectEditor' is open, showing the 'attributeIndices' field set to '2-19,1'. This indicates that the attributes are being reordered so that 'Churn' (attribute 1) is the last attribute.

weka.gui.GenericObjectEditor
weka.filters.unsupervised.attribute.Reorder

About
A filter that generates output with a new order of the attributes. [More](#)
[Capabilities](#)

attributeIndices

debug

doNotCheckCapabilities

[Open...](#) [Save...](#) [OK](#) [Cancel](#)

Choose **Reorder - R 2-19,1** Apply Stop

Current relation
Relation: customer_churn-weka.filters.unsupervised.insta...
Instances: 3384
Attributes: 19
Sum of weights: 3384

Attributes

No.	Name
2	<input type="checkbox"/> PreferredLoginDevice
3	<input type="checkbox"/> CityTier
4	<input type="checkbox"/> WarehouseToHome
5	<input type="checkbox"/> PreferredPaymentMode
6	<input type="checkbox"/> Gender
7	<input type="checkbox"/> HourSpendOnApp
8	<input type="checkbox"/> NumberOfDeviceRegistered
9	<input type="checkbox"/> PreferredOrderCat
10	<input type="checkbox"/> SatisfactionScore
11	<input type="checkbox"/> MaritalStatus
12	<input type="checkbox"/> NumberOfAddress
13	<input type="checkbox"/> Complain
14	<input type="checkbox"/> OrderAmountHikeFromlastYear
15	<input type="checkbox"/> CouponUsed
16	<input type="checkbox"/> OrderCount
17	<input type="checkbox"/> DaySinceLastOrder
18	<input type="checkbox"/> CashbackAmount
19	<input type="checkbox"/> Churn

Status
OK Log

Selected attribute
Name: Churn
Missing: 0 (0%)
Distinct: 2
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	0	2816	2816
2	1	568	568

Class: Churn (Nom) Visualize All

第四題

Weka Experiment Environment

Setup Run Analyse

Experiment Configuration Mode Simple

Results Destination
ARFF file

Experiment Type
Cross-validation
Number of folds: 10
☒ Classification ☐ Regression

Iteration Control
Number of repetitions: 10
☒ Data sets first ☐ Algorithms first

Datasets

☐ Use relative paths
D:\電商技\customer_churn1.arff

Algorithms

Logistic -R 1.0E-8 -M -1 -num-decimal-places 4
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVec

第五題

Source

Got 200 results

File... Database... Experiment

Actions

Perform test Save output Open Explorer...

Configure test

Testing with: Paired T-Tester (corrected)

Select rows and cols: Rows Cols Swap

Comparison field: Percent_correct

Significance: 0.05

Sorting (asc) by: <default>

Test base: Select

Displayed Columns: Select

Show std. deviations: ☒

Output Format: Select

Test output

Tester: weka.experiment.PairedCorrectedTTester -G 4,5,6 -D 1 -R 2 -S 0.05 -V -result

Analysing: Percent_correct

Datasets: 1

Resultsets: 2

Confidence: 0.05 (two tailed)

Sorted by: -

Date: 2023/5/2 下午9:08

Dataset (1) functions.Logi | (2) functions.S

'customer_churn-weka.filt(100)	88.75 (1.26)		88.61 (1.33)
	(v/ /*)		(0/1/0)

Key:

(1) functions.Logistic '-R 1.0E-8 -M -1 -num-decimal-places 4' 3932117032546553727

(2) functions.SMO '-C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -R \"functions.supportVect

Result list

21:07:55 - Available resultsets

21:08:03 - Percent_correct - functions.Logistic '-R 1.0E-8 -M -1

結論:最後依照 weka 得出的輸出結果，在 95%信心水準下，兩者並無顯著差異，故支持虛無假設，兩者平均為相同