

# Prediction of categorical data using decision tree

reprocess Classify Cluster Associate Select attributes Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

ter

Choose **None** Apply Stop

rent relation: german\_credit  
instances: 1000  
Attributes: 21  
Sum of weights: 1000

tributes

All None Invert Pattern

No. Name

- ☒ 1 checking\_status
- ☐ 2 duration
- ☐ 3 credit\_history
- ☐ 4 purpose
- ☐ 5 credit\_amount
- ☐ 6 savings\_status
- ☐ 7 employment
- ☐ 8 installment\_commitment
- ☐ 9 personal\_status
- ☐ 10 other\_parties
- ☐ 11 residence\_since
- ☐ 12 property\_magnitude
- ☐ 13 age
- ☐ 14 other\_payment\_plans
- ☐ 15 housing
- ☐ 16 existing\_credits
- ☐ 17 job
- ☐ 18 num\_dependents
- ☐ 19 own\_telephone
- ☐ 20 foreign\_worker
- ☐ 21 class

Remove

Selected attribute

Name: checking\_status  
Missing: 0 (0%)  
Distinct: 4  
Type: Nominal  
Unique: 0 (0%)

No.	Label	Count	Weight
1	<0	274	274
2	0<=X<200	269	269
3	>=200	63	63
4	no checking	394	394

Class: class (Nom) Visualize All

Log x 0

reprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N 5"**

Test options

- ☐ Use training set
- ☐ Supplied test set Set...
- ☒ Cross-validation Folds 10
- ☐ Percentage split % 66

More options...

Nom) class

Start Stop

Result list (right-click for options)

- 1:12-41 - rules.DecisionTable

Classifier output

```
=== Run information ===

#Scheme: weka.classifiers.rules.DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N 5"
Relation: german_credit
Instances: 1000
Attributes: 21
checking_status
duration
credit_history
purpose
credit_amount
savings_status
employment
installment_commitment
personal_status
other_parties
residence_since
property_magnitude
age
other_payment_plans
housing
existing_credits
job
num_dependents
own_telephone
foreign_worker
class

Test mode: 10-fold cross-validation

=== Classifier model (full training set) ===

Decision Table:

Number of training instances: 1000
Number of rules : 68
Non matches covered by Majority class.
Best first.
Start set: no attributes
```

Preprocess

Classifier

Cluster

Associate

Select attributes

Visualize

Classifier

ChooseDecisionTable - X 1 - 5 "weka.attributeSelection.BestFirst-D 1 - N 5"

est options

Use training set

Set...

Supplied test set

Set...

Cross-validation

Folds

10

Percentage split

%

66

More options...

Nom) class

Start

Stop

result list (right-click for options)

1:12:41 - rules.DecisionTable

Classifier output

Non matches covered by Majority class.

Best first.

Start set: no attributes

Search direction: forward

Stale search after 5 node expansions

Total number of subsets evaluated: 148

Merit of best subset found: 75.1

Evaluation (for feature selection): CV (leave one out)

Feature set: 1,2,3,20,21

Time taken to build model: 0.22 seconds

==== Stratified cross-validation ====

==== Summary ====

Correctly Classified Instances	710	71	%
Incorrectly Classified Instances	290	29	%
Kappa statistic	0.2033		
Mean absolute error	0.3677		
Root mean squared error	0.4321		
Relative absolute error	87.505 %		
Root relative squared error	94.2815 %		
Total Number of Instances	1000		

==== Detailed Accuracy By Class ====

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.893	0.717	0.744	0.893	0.812	0.220	0.723	0.854	good
	0.283	0.107	0.531	0.283	0.370	0.220	0.723	0.478	bad
Weighted Avg.	0.710	0.534	0.680	0.710	0.679	0.220	0.723	0.741	

==== Confusion Matrix ====

a	b	<-- classified as
625	75	a = good
215	85	b = bad