

Weka Explorer

Preprocess Classify **Cluster** Associate Select attributes Visualize

Clusterer: Choose **SimpleKMeans** -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Cluster mode

- ☒ Use training set
- ☐ Supplied test set
- ☐ Percentage split % 66
- ☐ Classes to clusters evaluation
- (Nom) class
- ☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)

- 15:04:20 - SimpleKMeans
- 15:07:55 - SimpleKMeans

Cluster output

```

=== Run information ===

Scheme:      weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -
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:   evaluate on training data
  
```

Status OK Log x 0

Weka Explorer

Preprocess Classify **Cluster** Associate Select attributes Visualize

Clusterer: Choose **SimpleKMeans** -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Cluster mode

- ☒ Use training set
- ☐ Supplied test set
- ☐ Percentage split % 66
- ☐ Classes to clusters evaluation
- (Nom) class
- ☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)

- 15:04:20 - SimpleKMeans
- 15:07:55 - SimpleKMeans

Cluster output

```

kMeans
=====

Number of iterations: 5
Within cluster sum of squared errors: 5665.9976202840735

Initial starting points (random):

Cluster 0: 'no checking',36,'critical/other existing credit','new car',7855,<100,1<=X<4,4,'female div/dep/mar',none,2,'real
Cluster 1: <0,24,'critical/other existing credit','used car',6615,<100,unemployed,2,'male single',none,4,'no known property

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute          Full Data          Cluster#
                   (1000.0)          0          1
                   (643.0)          (357.0)
=====
checking_status    no checking         no checking         <0
duration           20.903             19.9285             22.6583
credit_history      existing paid        existing paid        existing paid
purpose            radio/tv            radio/tv            new car
credit_amount       3271.258            2924.7869           3895.2941
savings_status      <100                <100                <100
employment          1<=X<4             1<=X<4             >=7
installment_commitment 2.973              2.9611              2.9944
personal_status     male single         male single         male single
other_parties       none                none                none
residence_since     2.845              2.5599              3.3585
property_magnitude  car                 car no known property
  
```

Status OK Log x 0

