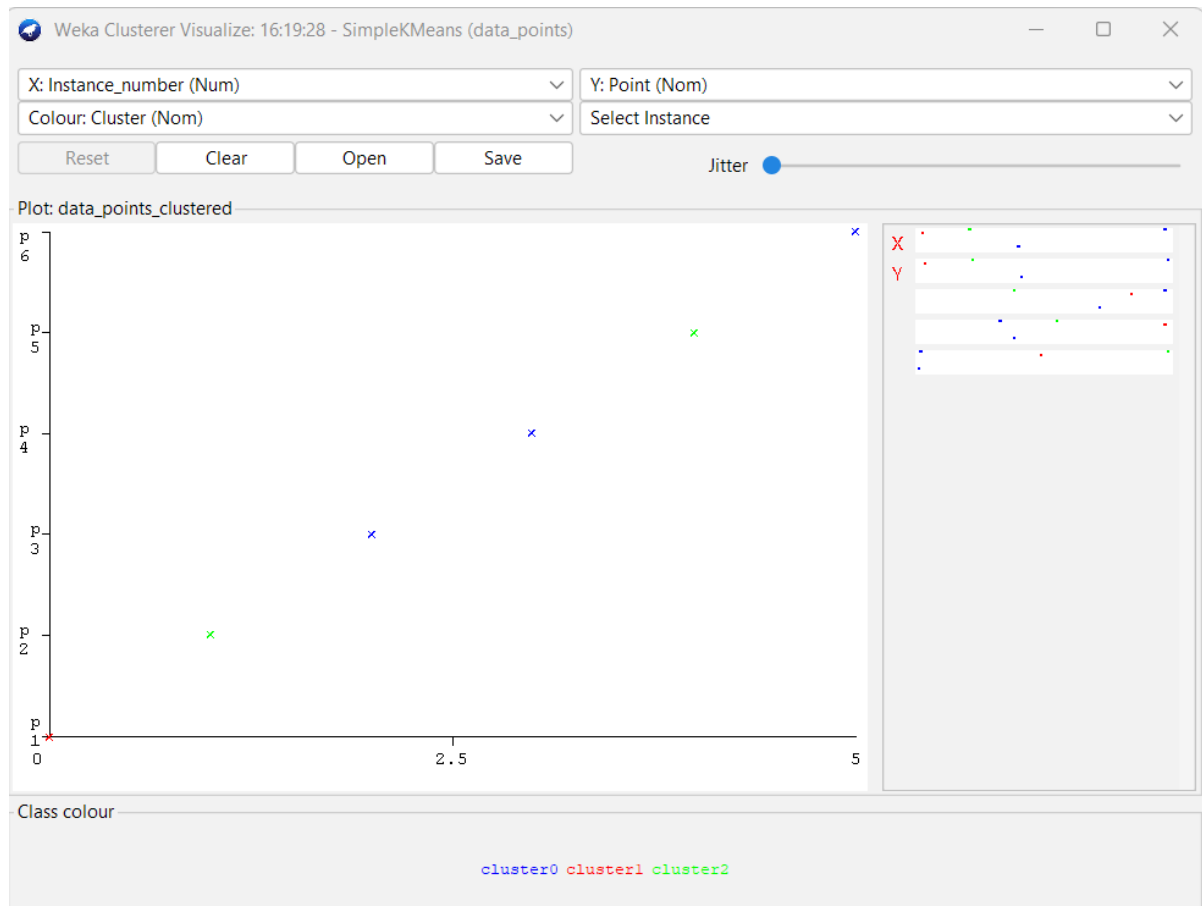


Mark Lopes

9913

Experiment-7 (k means clustering)

4) k-means Weka



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 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1

Cluster mode

☒ Use training set

☐ Supplied test set Set...

☐ Percentage split % 66

☐ Classes to clusters evaluation (Num) Spending Score

☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)

16:19:28 - SimpleKMeans

Clusterer output

```

=== Run information ===

Scheme:      weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning
Relation:    data_points
Instances:    6
Attributes:   3
              Point
              Income
              Spending Score
Test mode:    evaluate on training data

=== Clustering model (full training set) ===

kMeans
=====

Number of iterations: 3
Within cluster sum of squared errors: 3.2922227263268127

Initial starting points (random):

Cluster 0: p4,0.26,0.19
Cluster 1: p1,0.4,0.53
Cluster 2: p2,0.22,0.38

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute      Full Data      Cluster#
              (6.0)      (3.0)      (1.0)      (2.0)

```

Status OK Log

Start Stop

Result list (right-click for options)

16:19:28 - SimpleKMeans

Final cluster centroids:

| Attribute | Full Data | Cluster# | 0 | 1 | 2 |
|----------------|-----------|----------|-------|-------|---|
| | (6.0) | (3.0) | (1.0) | (2.0) | |
| Point | p1 | p3 | p1 | p2 | |
| Income | 0.2933 | 0.3533 | 0.4 | 0.15 | |
| Spending Score | 0.355 | 0.27 | 0.53 | 0.395 | |

Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

| | |
|---|----------|
| 0 | 3 (50%) |
| 1 | 1 (17%) |
| 2 | 2 (33%) |

Status OK Log