

Using domesticity as the X axis and aquatic as the y axis, we can see there’s a single outlier where domestic = yes, and aquatic = yes.

animal-name : carp

hair : 0.0

feathers : 0.0

egg : 1.0

milk : 0.0

airborne : 0.0

aquatic : 1.0

predator : 0.0

toothed : 1.0

backbone : 1.0

breathes : 0.0

venomous : 0.0

fins : 1.0

legs : 0

tail : 1.0

domestic : 1.0

catsize : 0.0

type : 4

Cluster : cluster2

=== Run information ===

Scheme: weka.clusterers.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 -S 20

Relation: zoo-weka.filters.unsupervised.attribute.NumericToNominal-Rlast-weka.filters.unsupervised.attribute.NumericToNominal-R14

Instances: 101

Attributes: 18

animal-name

hair

feathers

egg

milk

airborne

aquatic

predator

toothed

backbone

breathes

venomous

fins

legs

tail

domestic

catsize

type

Test mode: evaluate on training data

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

kMeans

======

Number of iterations: 3

Within cluster sum of squared errors: 370.36711436711437

Initial starting points (random):

Cluster 0: bear,1,0,0,1,0,0,1,1,1,1,0,0,4,0,0,1,1

Cluster 1: hare,1,0,0,1,0,0,0,1,1,1,0,0,4,1,0,0,1

Cluster 2: slug,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,7

Missing values globally replaced with mean/mode

Final cluster centroids:

Cluster#

Attribute Full Data 0 1 2

(101.0) (26.0) (21.0) (54.0)

=========================================================

animal-name frog frog antelope bass

hair 0.4257 0.6923 1 0.0741

feathers 0.198 0 0 0.3704

egg 0.5842 0.2692 0 0.963

milk 0.4059 0.7692 1 0

airborne 0.2376 0 0.0952 0.4074

aquatic 0.3564 0.4615 0 0.4444

predator 0.5545 1 0.0952 0.5185

toothed 0.604 0.9615 1 0.2778

backbone 0.8218 1 1 0.6667

breathes 0.7921 0.8462 1 0.6852

venomous 0.0792 0.0769 0 0.1111

fins 0.1683 0.3077 0 0.1667

legs 4 4 4 2

tail 0.7426 0.7692 0.9048 0.6667

domestic 0.1287 0.0769 0.2857 0.0926

catsize 0.4356 0.9231 0.5714 0.1481

type 1 1 1 2

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

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

Clustered Instances

0 26 ( 26%)

1 21 ( 21%)

2 54 ( 53%)