

Experiment-3

The screenshot shows a software interface for clustering. It has a 'Cluster mode' section with four radio buttons: 'Use training set', 'Supplied test set' (with a 'Set...' button), 'Percentage split' (selected, with a '%' sign and a text box containing '80'), and 'Classes to clusters evaluation' (with a '(Nom) class' dropdown). Below these is a checked checkbox for 'Store clusters for visualization'. There is an 'Ignore attributes' text box. Two buttons, 'Start' and 'Stop', are present. At the bottom is a 'Result list (right-click for options)' with one entry: '17:26:43 - EM'.

method

CLUTTERING TASK

The task of grouping data points based on their similarity with each other is called Clustering or Cluster Analysis.

The various types of clustering are:

Hierarchical clustering

Partitioning clustering

Hierarchical clustering is further subdivided into:

Agglomerative clustering

Divisive clustering

Partitioning clustering is further subdivided into:

K-Means clustering

Fuzzy C-Means clustering

W There is no labeled data for this clustering, unlike in supervised learning. K-Means performs the division of objects into clusters that share similarities and are dissimilar to the objects belonging to another cluster.

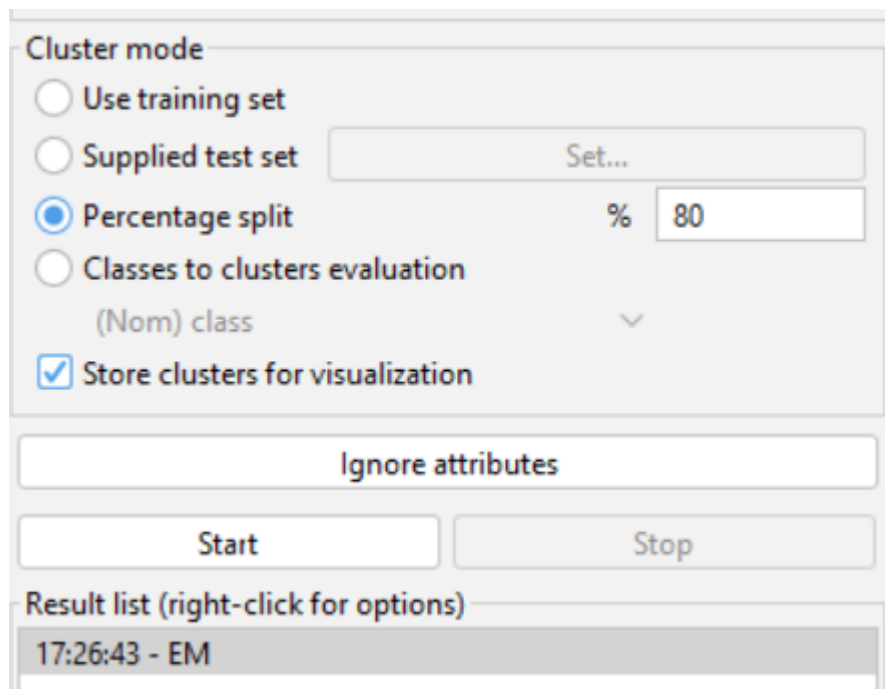
STEPS INVOLVED

dataset used :- iris.arff

NORMALISE THE DATASET

Go to cluster menu and use percentage split 80%

CLUSTER METHOD



The screenshot shows the 'Cluster' dialog box in WEKA. Under the 'Cluster mode' section, the 'Percentage split' option is selected with a radio button. To its right, there is a percentage sign and a text box containing the value '80'. Below this, the 'Classes to clusters evaluation' option is unselected, with '(Nom) class' and a dropdown arrow next to it. The 'Store clusters for visualization' checkbox is checked. Below the 'Cluster mode' section is an 'Ignore attributes' text box. At the bottom of the dialog are 'Start' and 'Stop' buttons. Below the buttons is a 'Result list (right-click for options)' section, which contains a single entry: '17:26:43 - EM'.

Cluster mode

☐ Use training set

☐ Supplied test set

☒ Percentage split %

☐ Classes to clusters evaluation
(Nom) class

☒ Store clusters for visualization

Result list (right-click for options)

17:26:43 - EM

```

kMeans
=====

Number of iterations: 2
Within cluster sum of squared errors: 56.76649282223416

Initial starting points (random):

Cluster 0: 0.5,0.333333,0.508475,0.5,Iris-versicolor
Cluster 1: 0.805556,0.666667,0.864407,1,Iris-virginica

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute          Full Data          Cluster#
                   (120.0)          0          1
                   (79.0)          (41.0)
=====
sepalength          0.4347          0.3326          0.6314
sepalwidth          0.4347          0.4509          0.4035
petallength         0.4736          0.3169          0.7755
petalwidth          0.4635          0.289          0.7998
class               Iris-virginica Iris-versicolor Iris-virginica

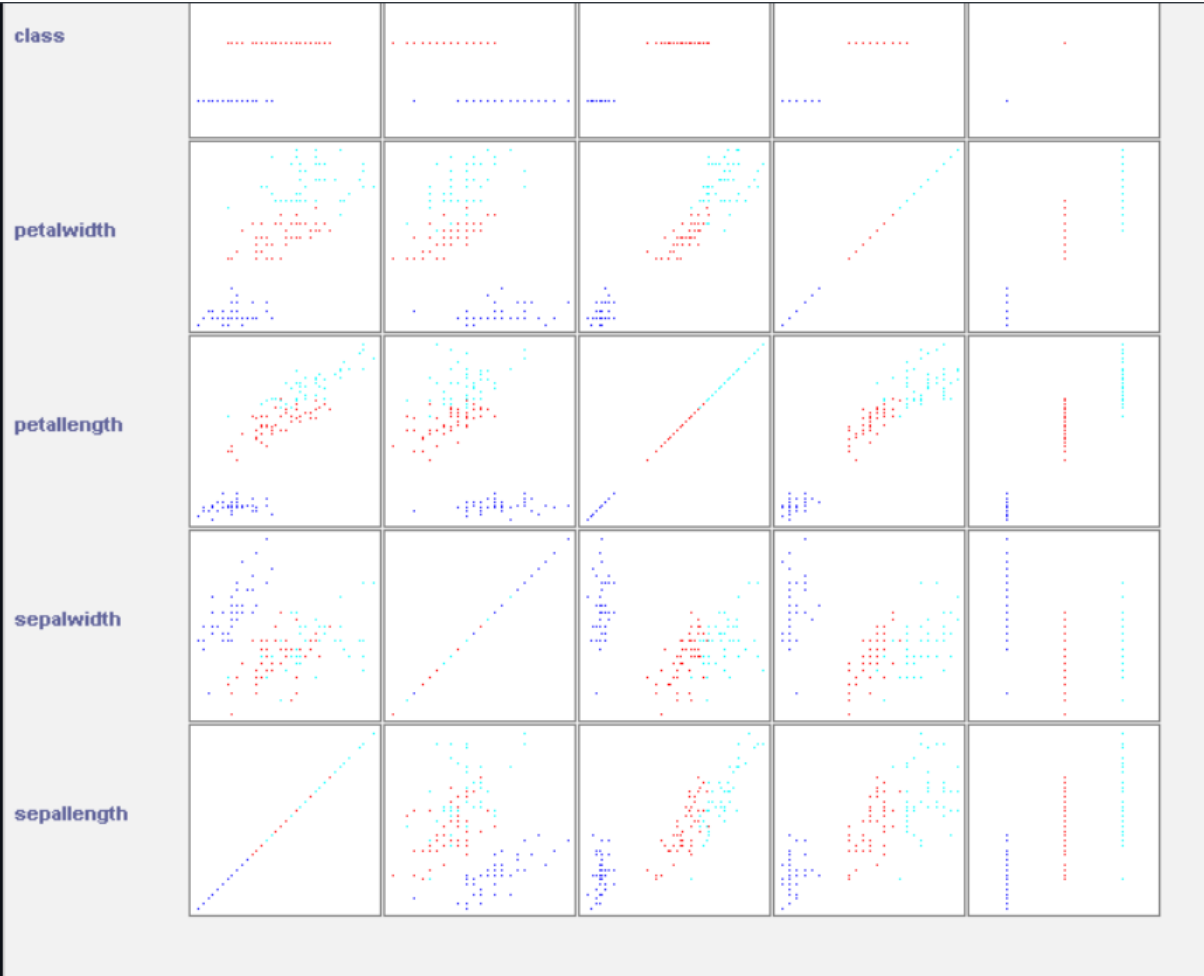
Time taken to build model (percentage split) : 0 seconds

Clustered Instances

0      21 ( 70%)
1       9 ( 30%)

```

VISUALISATION



```

KMeans
=====

Number of iterations: 2
Within cluster sum of squared errors: 56.76649282223416

Initial starting points (random):

Cluster 0: 0.5,0.333333,0.508475,0.5,Iris-versicolor
Cluster 1: 0.805556,0.666667,0.864407,1,Iris-virginica

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute          Full Data          Cluster#
                   (120.0)          (79.0)          (41.0)
=====
sepalength          0.4347          0.3326          0.6314
sepalwidth          0.4347          0.4509          0.4035
petallength         0.4736          0.3169          0.7755
petalwidth          0.4635          0.289          0.7998
class               Iris-virginica Iris-versicolor Iris-virginica

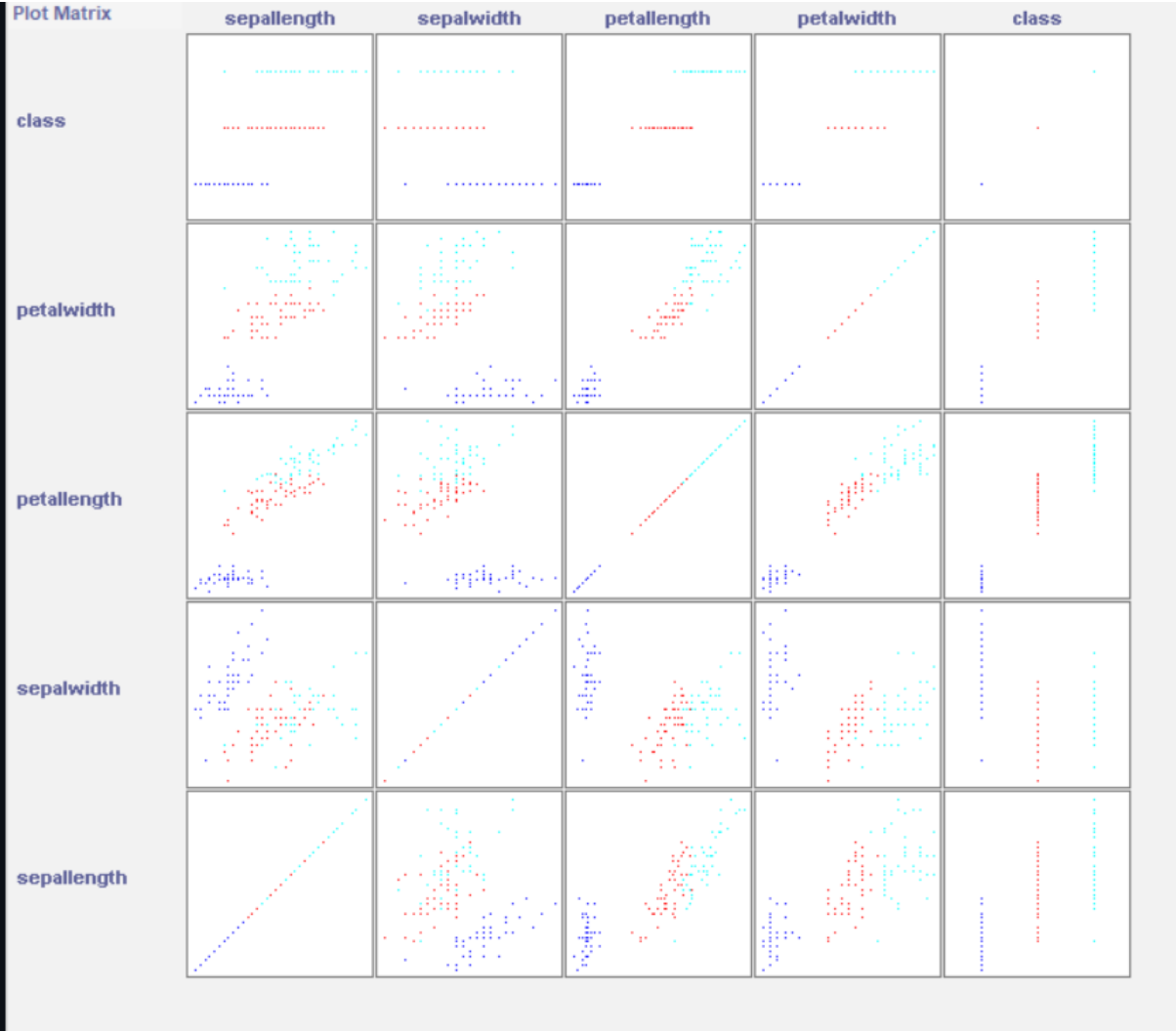
Time taken to build model (percentage split) : 0 seconds

Clustered Instances

0      21 ( 70%)
1       9 ( 30%)

```

Kmeans



VISUALISE