K-Means Cluster

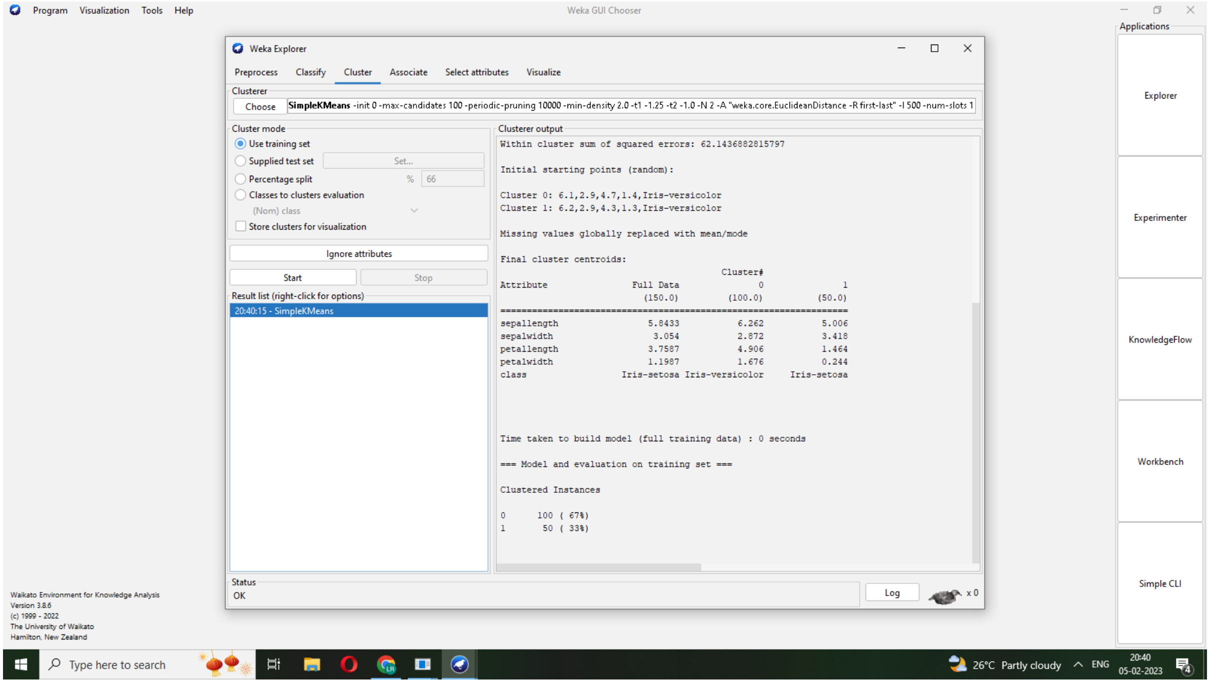
Aim:

To implement and design kmeans cluter algorithm using weka.

Algorithm:

- Determine root node
- Calculate entropy for classes
- Calculate entropy after split for each attributes.
- Calculate information gain
- Perform split
- Perform further split
- Compute kmeans cluster algorithm

Output:



Clusterer: 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

Cluster mode:
● Use training set
○ Supplied test set
○ Percentage split % 66
○ Classes to clusters evaluation
(Nom) class
□ Store clusters for visualization

Result list (right-click for options)
20:40:15 - SimpleKMeans

Clusterer output

```
Within cluster sum of squared errors: 62.1436882815797

Initial starting points (random):

Cluster 0: 6.1,2.9,4.7,1.4,Iris-versicolor
Cluster 1: 6.2,2.9,4.3,1.3,Iris-versicolor

Missing values globally replaced with mean/mode

Final cluster centroids:
                                         Cluster#
Attribute                 Full Data             0             1
                            (150.0)       (100.0)        (50.0)
==================================================================
sepallength                  5.8433         6.262         5.006
sepalwidth                    3.054         2.872         3.418
petallength                  3.7587         4.906         1.464
petalwidth                   1.1987         1.676         0.244
class                   Iris-setosa Iris-versicolor     Iris-setosa




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

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

Clustered Instances

0      100 ( 67%)
1       50 ( 33%)
```

Status
OK

Waikato Environment for Knowledge Analysis
Version 3.8.6
(c) 1999 - 2022
The University of Waikato
Hamilton, New Zealand