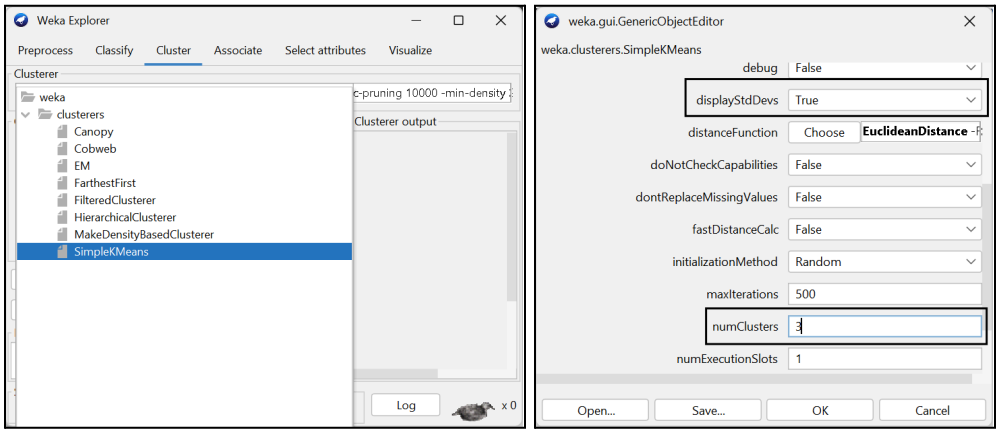
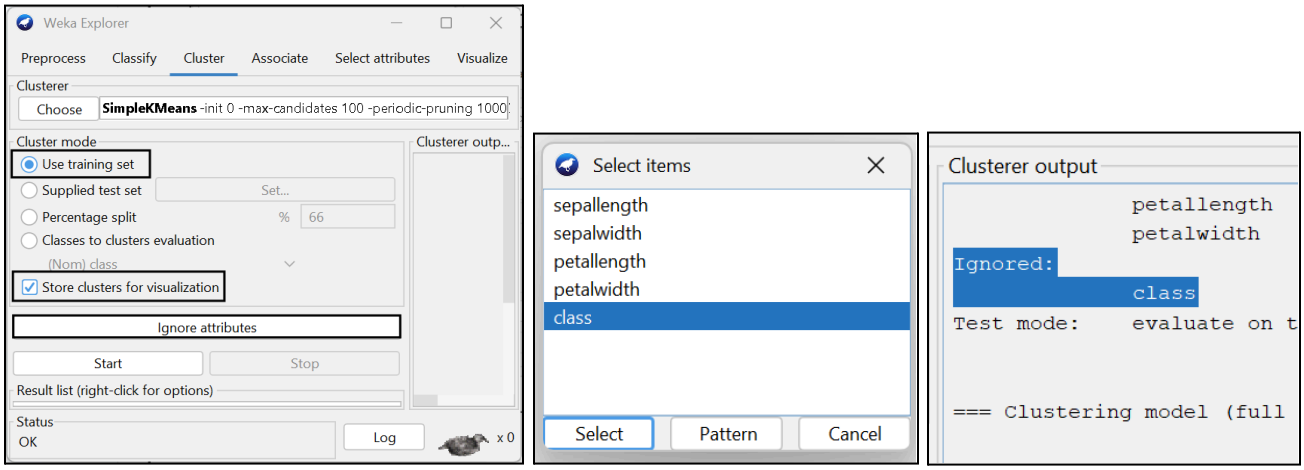
Practical 10: Implementing Clustering with Weka

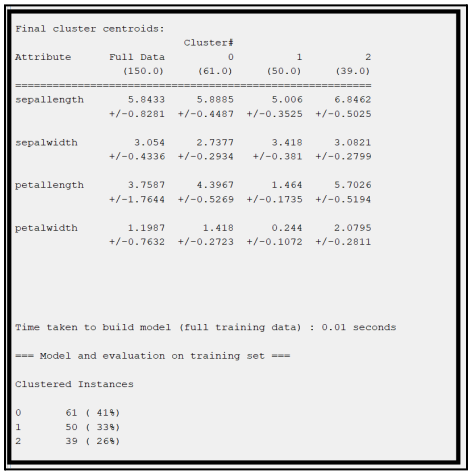
Step 01: Open Weka Explorer → Preprocess → load iris.arff file → move to cluster tab → choose simple k-means → click on simple kmeans and set '**displaystdDevs as True and numClusters as 3'**



Step 02: select **Use training set** → tick **store clusters for visualization** → hit Ignore attributes → select **Class** → hit **start.**

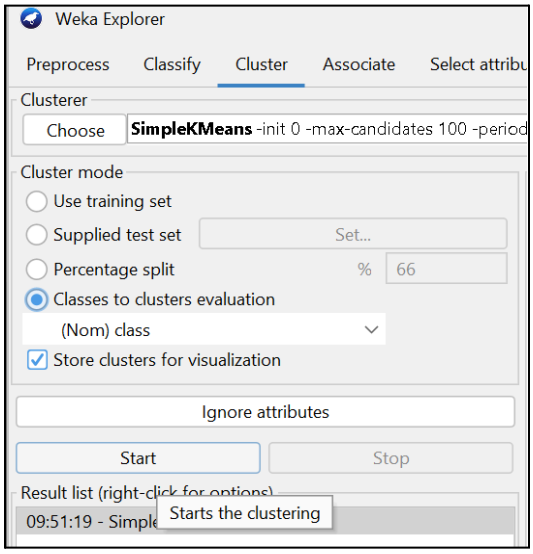


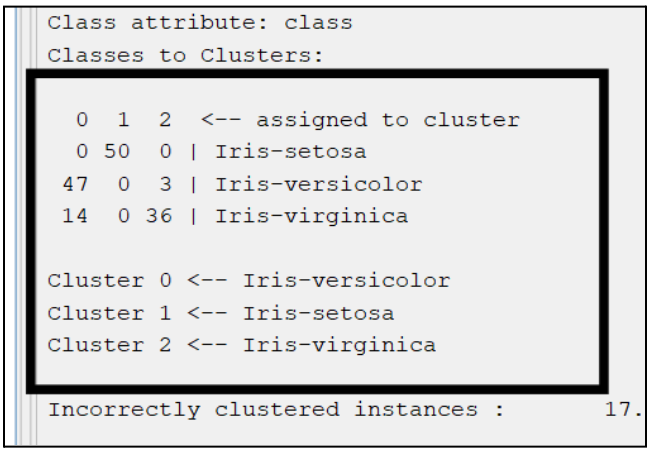
Step 03: In the Cluster output we can see that the K-Means algorithm forms **three clusters** with **61, 50, and 39 instances.**



```
Final cluster centroids:
                        Cluster#
Attribute       Full Data           0          1          2
                  (150.0)      (61.0)     (50.0)     (39.0)
============================================================
sepallength        5.8433      5.8885      5.006     6.8462
                 +/-0.8281  +/-0.4487  +/-0.3525  +/-0.5025

sepalwidth         3.054      2.7377      3.418     3.0821
                 +/-0.4336  +/-0.2934   +/-0.381  +/-0.2799

petallength        3.7587      4.3967      1.464     5.7026
                 +/-1.7644  +/-0.5269  +/-0.1735  +/-0.5194

petalwidth         1.1987       1.418      0.244     2.0795
                 +/-0.7632  +/-0.2723  +/-0.1072  +/-0.2811




Time taken to build model (full training data) : 0.01 seconds

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

Clustered Instances

0       61 ( 41%)
1       50 ( 33%)
2       39 ( 26%)
```

Step 04: To compare results with actual clusters, select Classes to Cluster Evaluation in Cluster mode and re-run the algorithm.





Step 05: You can also visualize the cluster. Clusters can be visualized using any input attribute. Clusters plotted with **Petal Length** and **Petal Width**. Increase **Jitter** to view all samples.