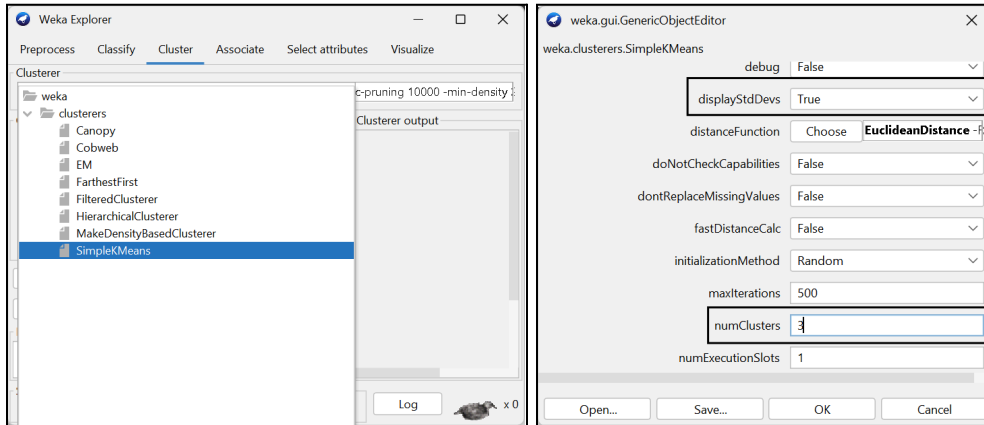
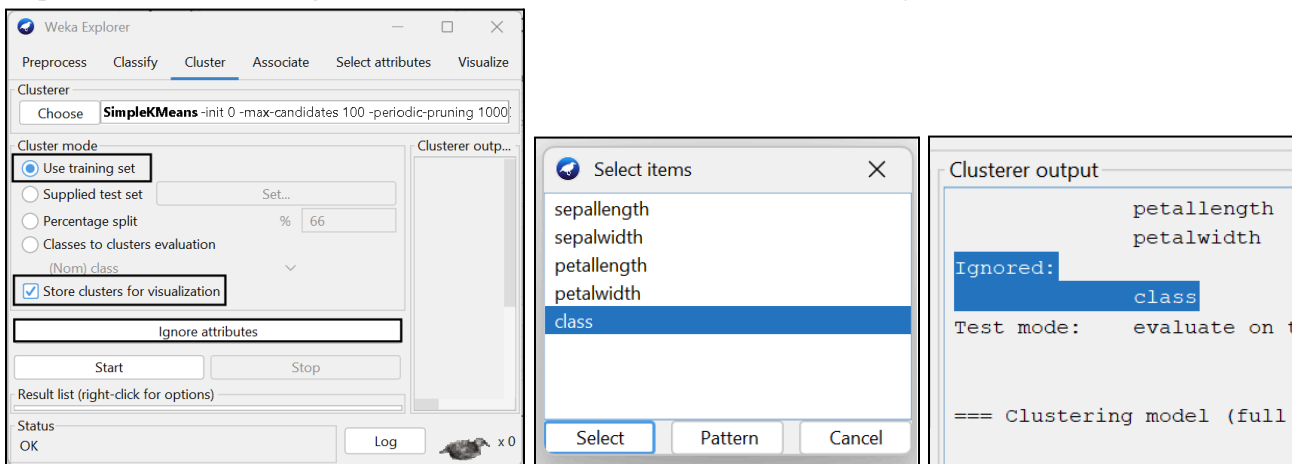


## Practical 10: Implementing Clustering with Weka

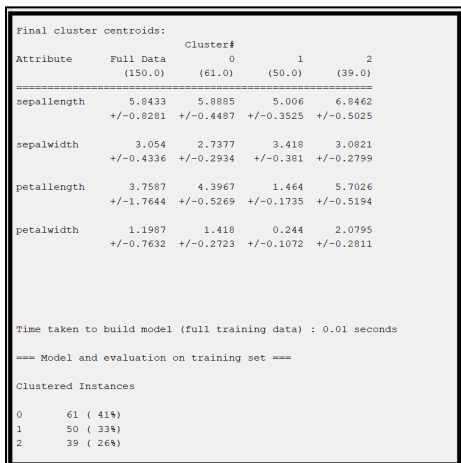
**Step 01:** Open Weka Explorer → Preprocess → load iris.arff file → move to cluster tab → choose simple k-means → click on simple kmeans and set ‘displayStdDevs as True and numClusters as 3’



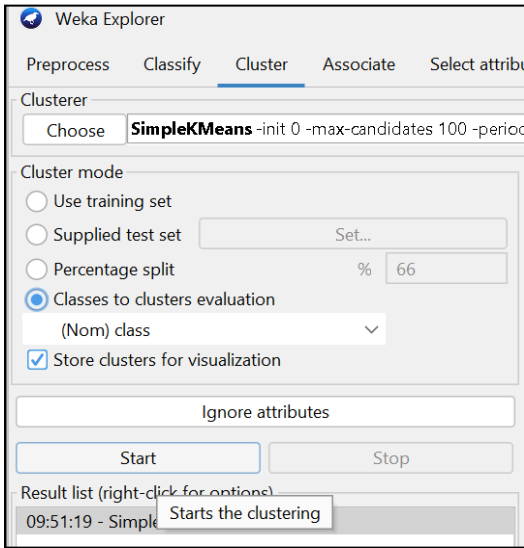
**Step 02:** select Use training set → tick store clusters for visualization → hit Ignore attributes → select Class → hit start.



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



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



Weka Explorer  
 Preprocess Classify **Cluster** Associate Select attrib...

Clusterer  
 Choose **SimpleKMeans** -init 0 -max-candidates 100 -period...

Cluster mode  
☐ Use training set  
☐ Supplied test set Set...  
☐ Percentage split % 66  
☒ **Classes to clusters evaluation**  
 (Nom) class  
☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)  
 09:51:19 - SimpleKMeans Starts the clustering

```

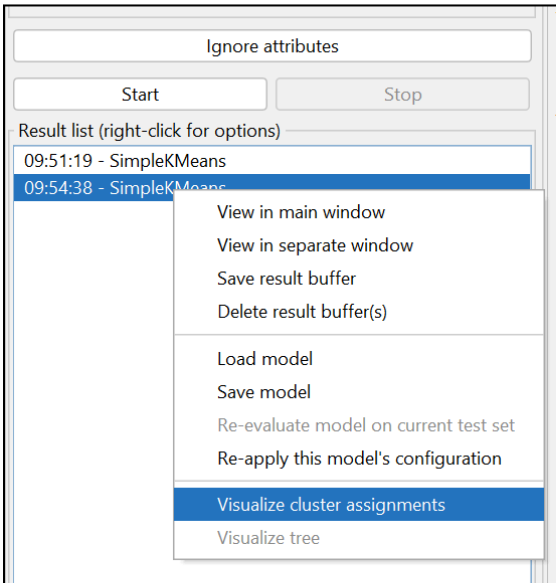
Class attribute: class
Classes to Clusters:

 0  1  2  <-- assigned to cluster
 0 50  0 | Iris-setosa
47  0  3 | Iris-versicolor
14  0 36 | Iris-virginica

Cluster 0 <-- Iris-versicolor
Cluster 1 <-- Iris-setosa
Cluster 2 <-- Iris-virginica

Incorrectly clustered instances :      17.
        
```

**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.



Ignore attributes

Start Stop

Result list (right-click for options)  
 09:51:19 - SimpleKMeans  
 09:54:38 - SimpleKMeans

- View in main window
- View in separate window
- Save result buffer
- Delete result buffer(s)
- Load model
- Save model
- Re-evaluate model on current test set
- Re-apply this model's configuration
- Visualize cluster assignments**
- Visualize tree



Weka Clusterer Visualize: 09:54:38 - SimpleKMeans (iris)

X: Instance\_number (Num) Y: sepalength (Num)  
 Colour: Cluster (Nom) Select Instance

Reset Clear Open Save Jitter

Plot: iris\_clustered

Class colour  
 cluster0 cluster1 cluster2

Weka: Instance info  
 Plot : 09:54:38 - SimpleKMeans (iris)  
 Instance: 62  
 Instance\_number : 61.0  
 sepalength : 5.9  
 sepalwidth : 3.0  
 petalength : 4.2  
 petalwidth : 1.5  
 class : Iris-versicolor  
 Cluster : cluster0