

Data Mining and Warehousing



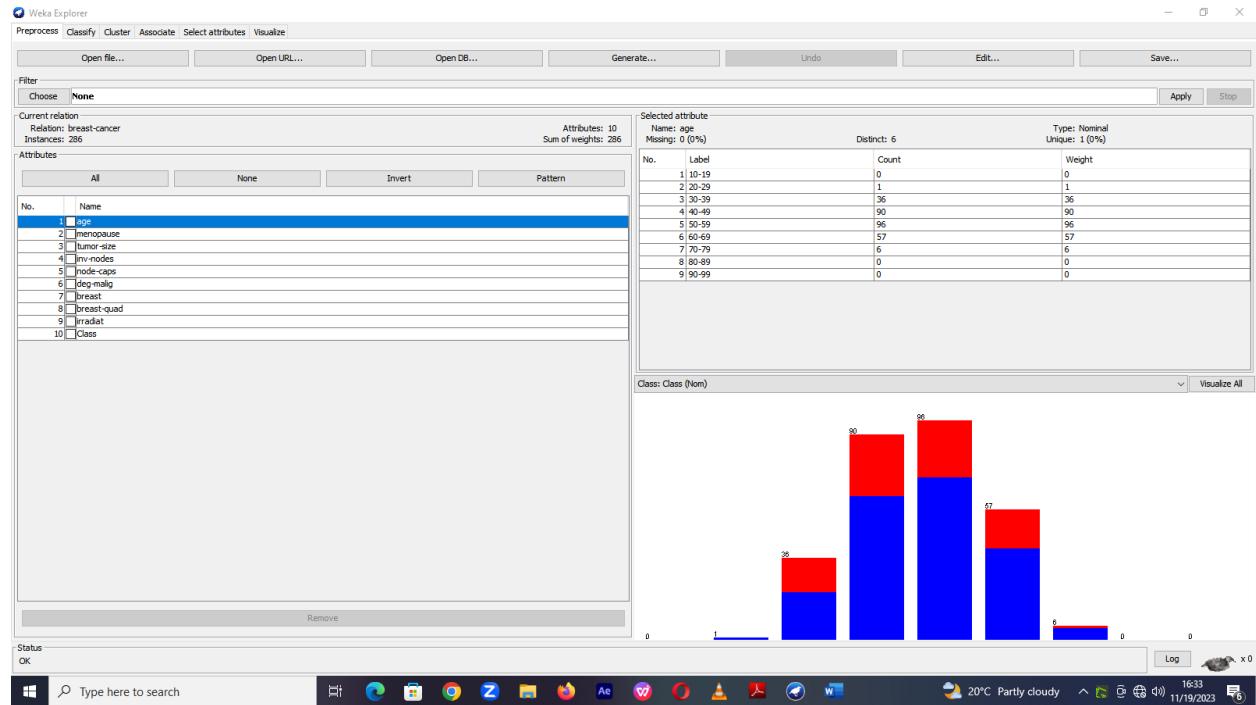
WEKA Tool Practical

By

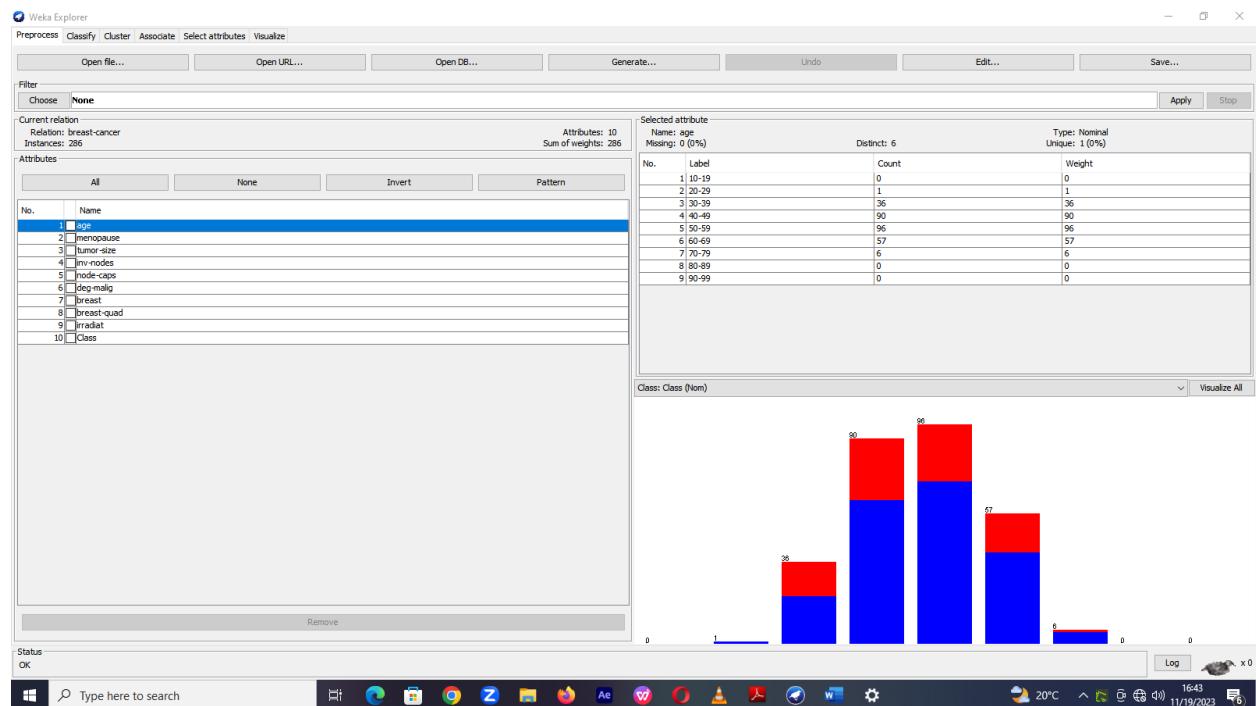


DEWTON KIPROP

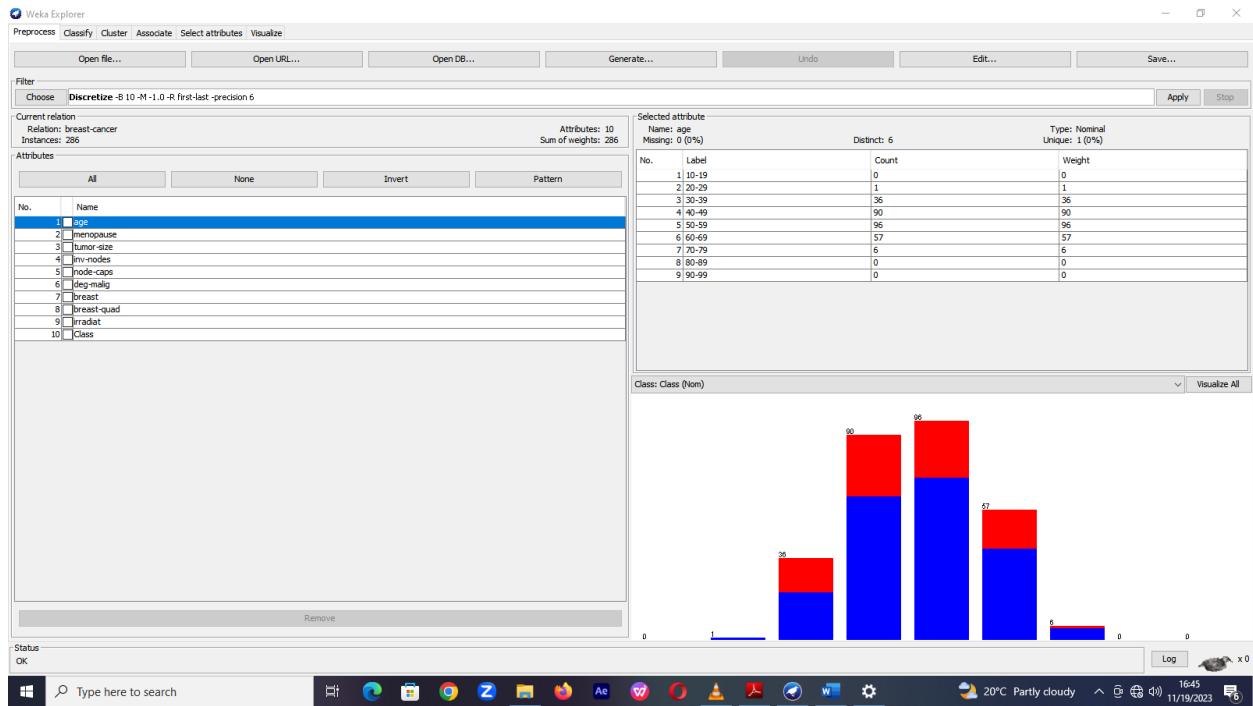
Step 1 I opened weka platform



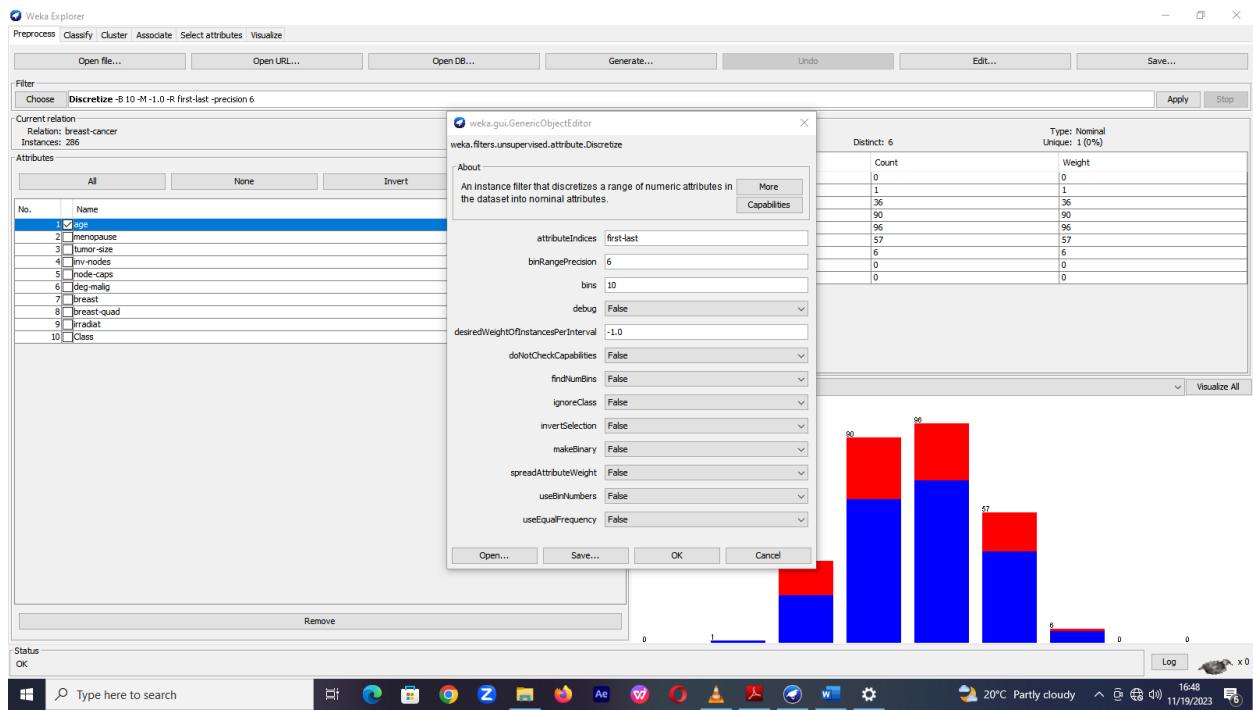
Step 2 selecting the button chose



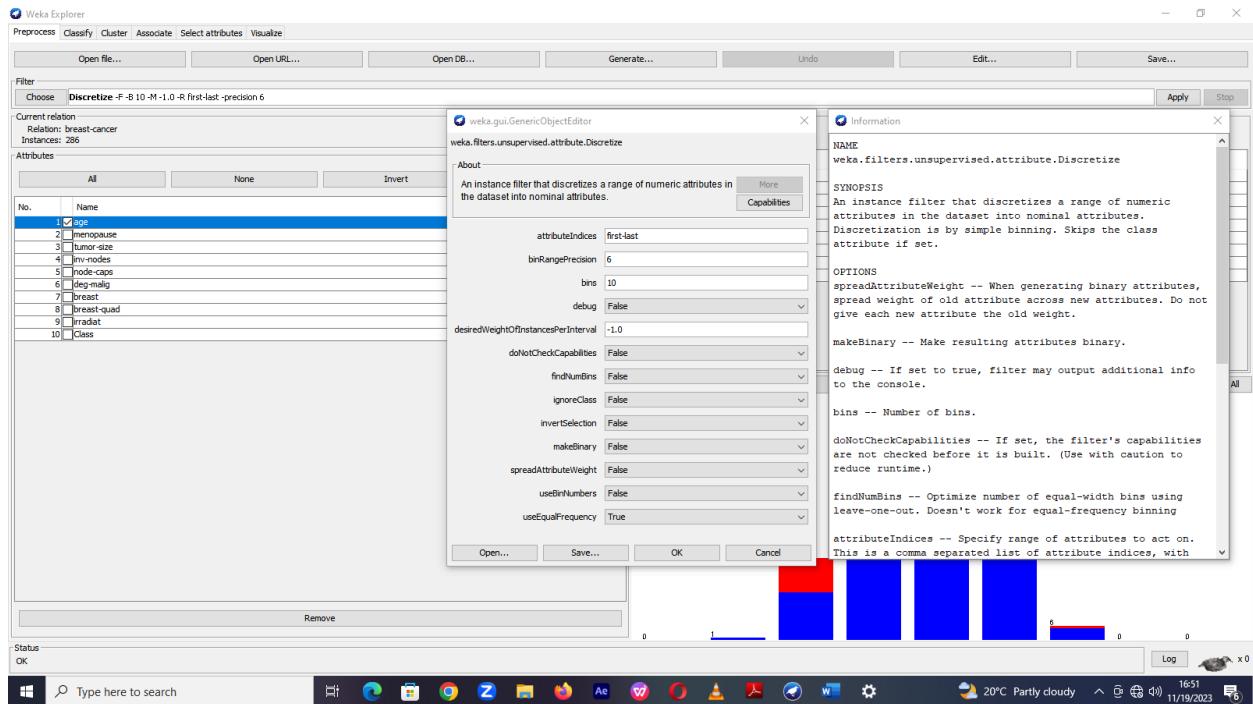
Part 3 navigating to discretize



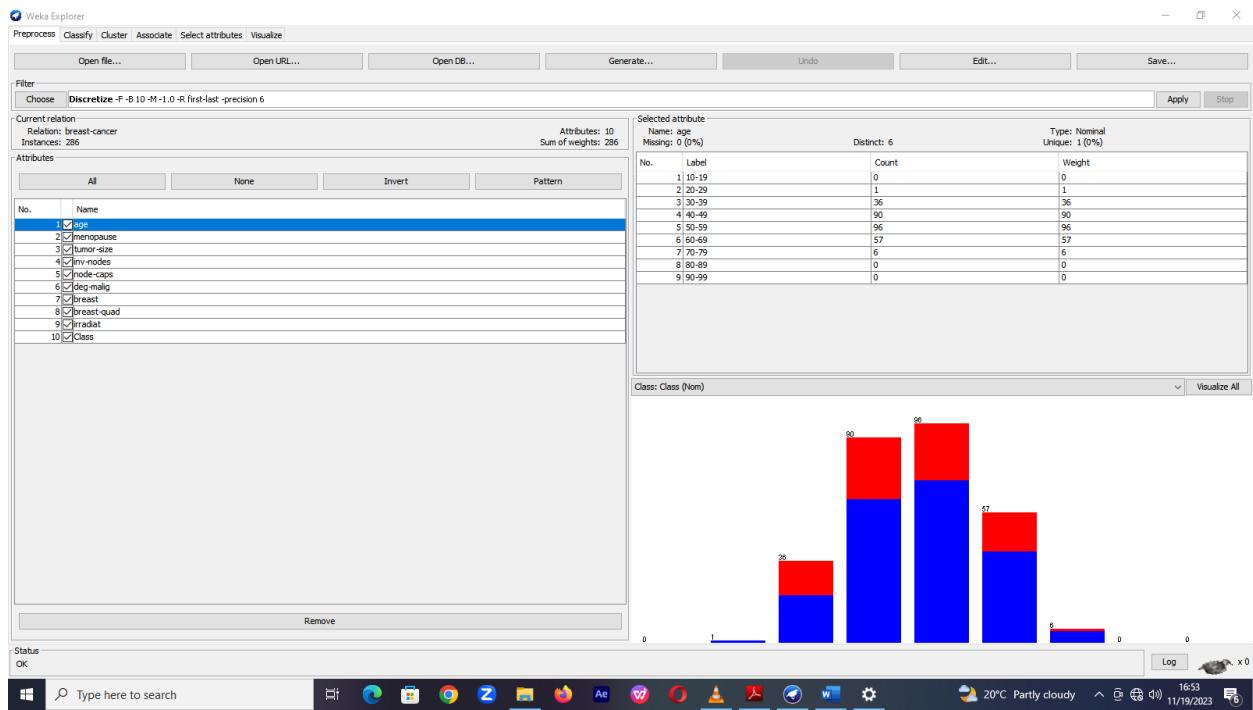
Finding the more option by clicking on the B10



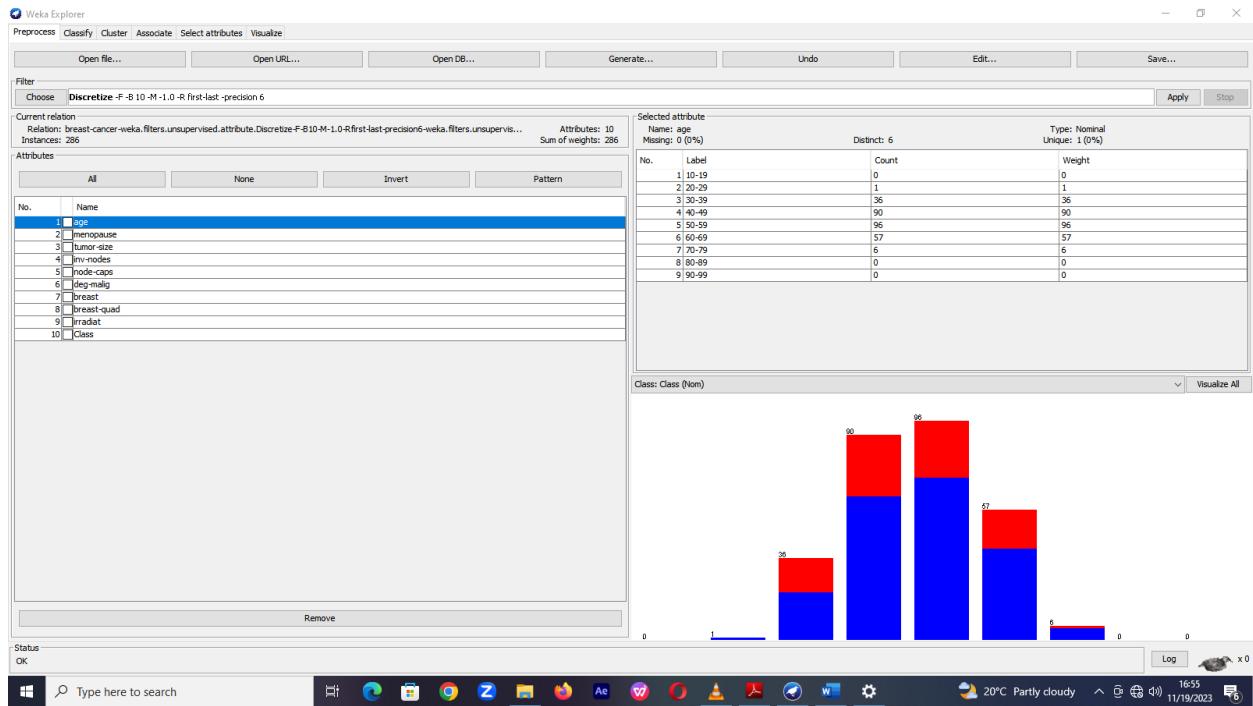
Clicking more, creating 10 bins and enabling true in UseEqualFrequency



Applying OK



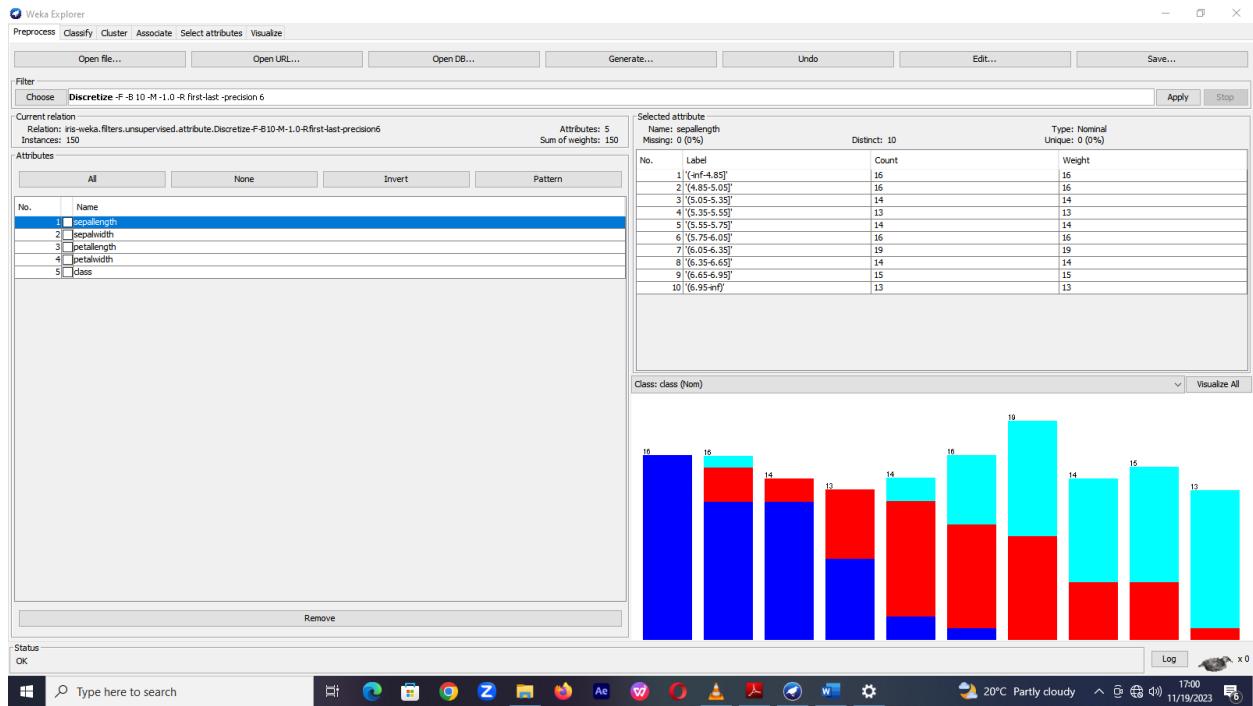
Clicking Apply button



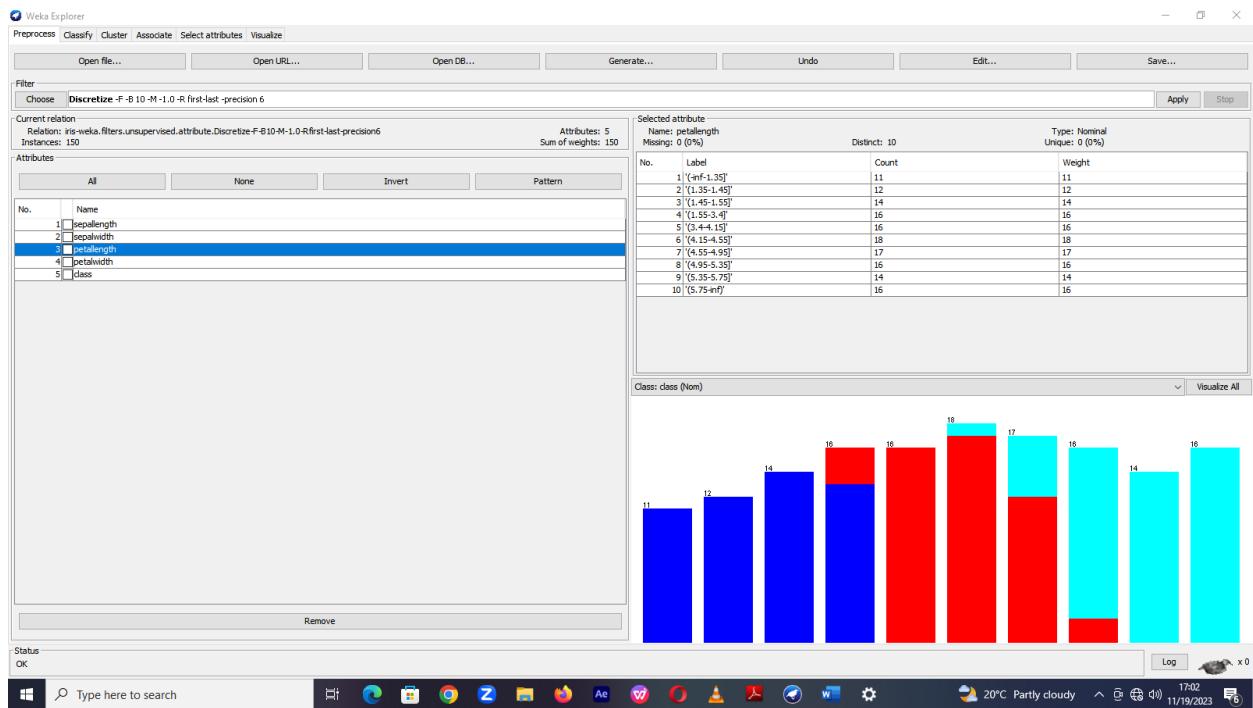
Visualizing all



Opening another file



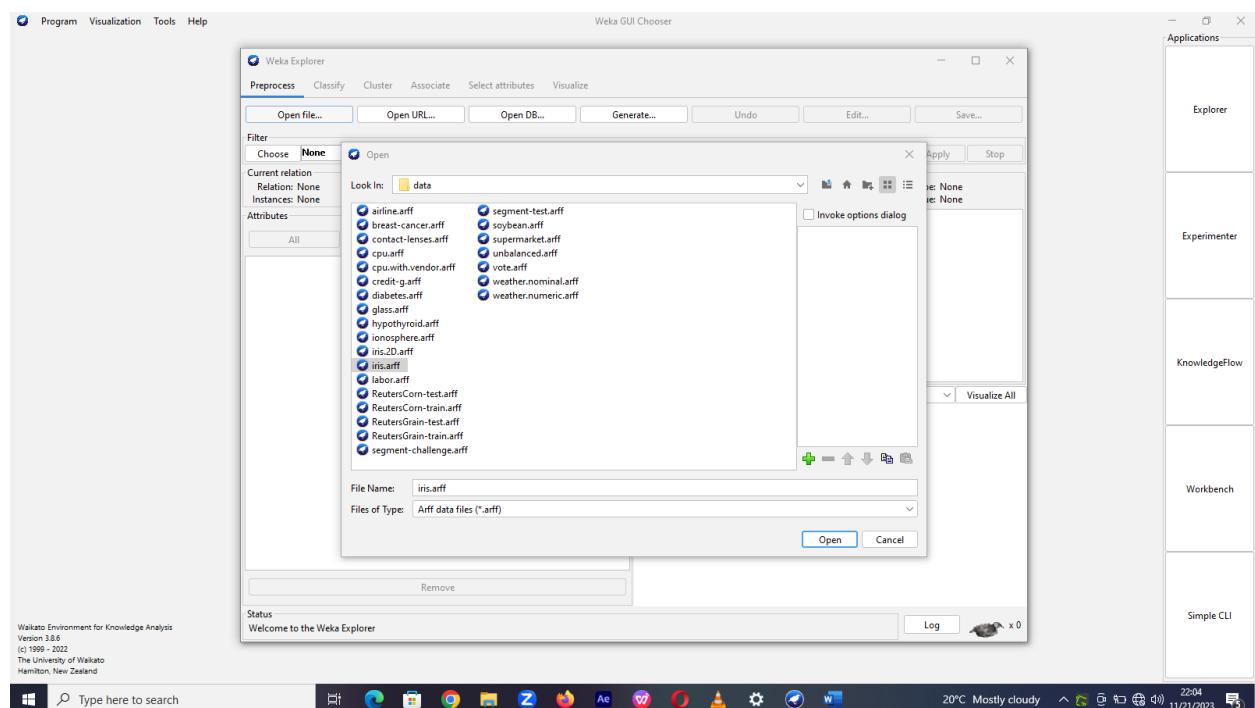
Opening petallength



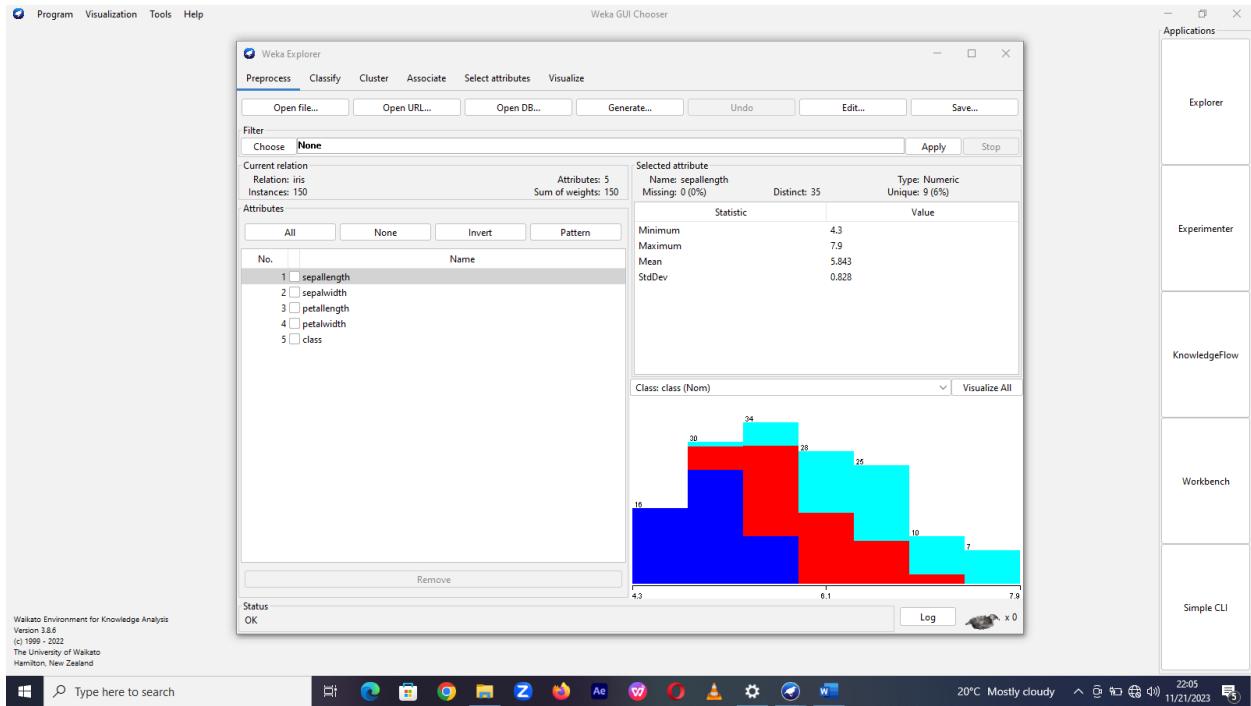
Visualization for all



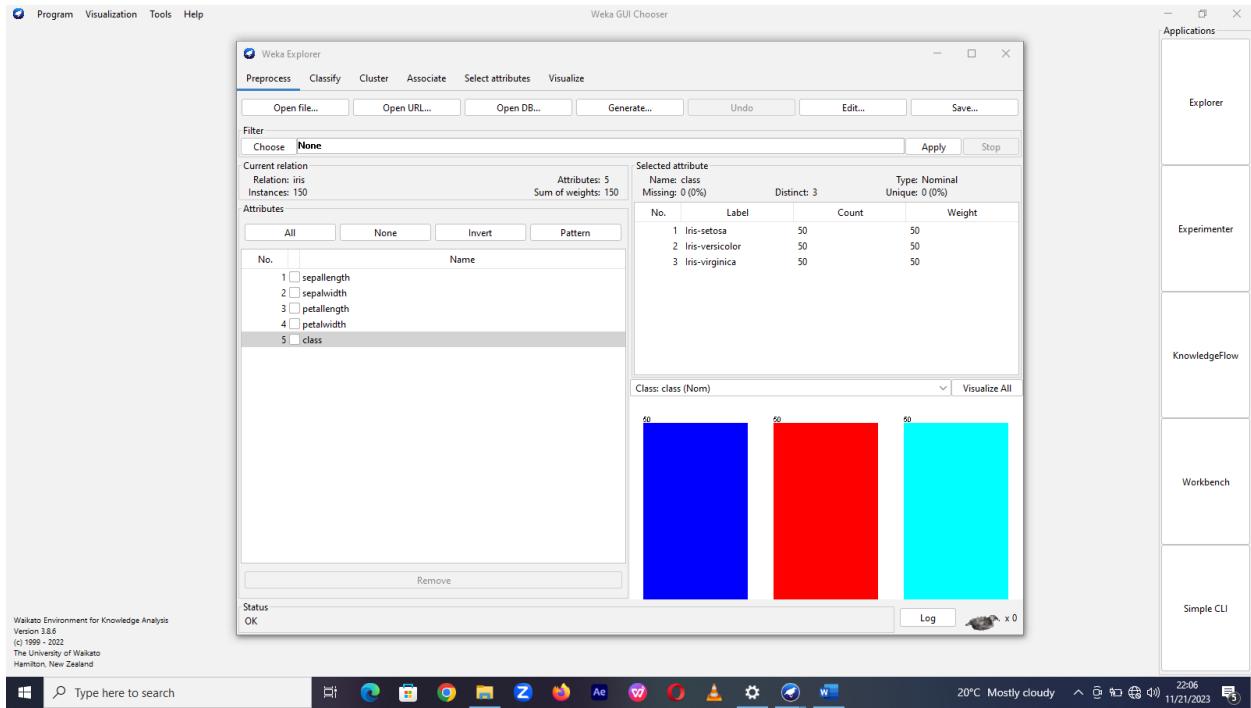
Reading in the Iris dataset



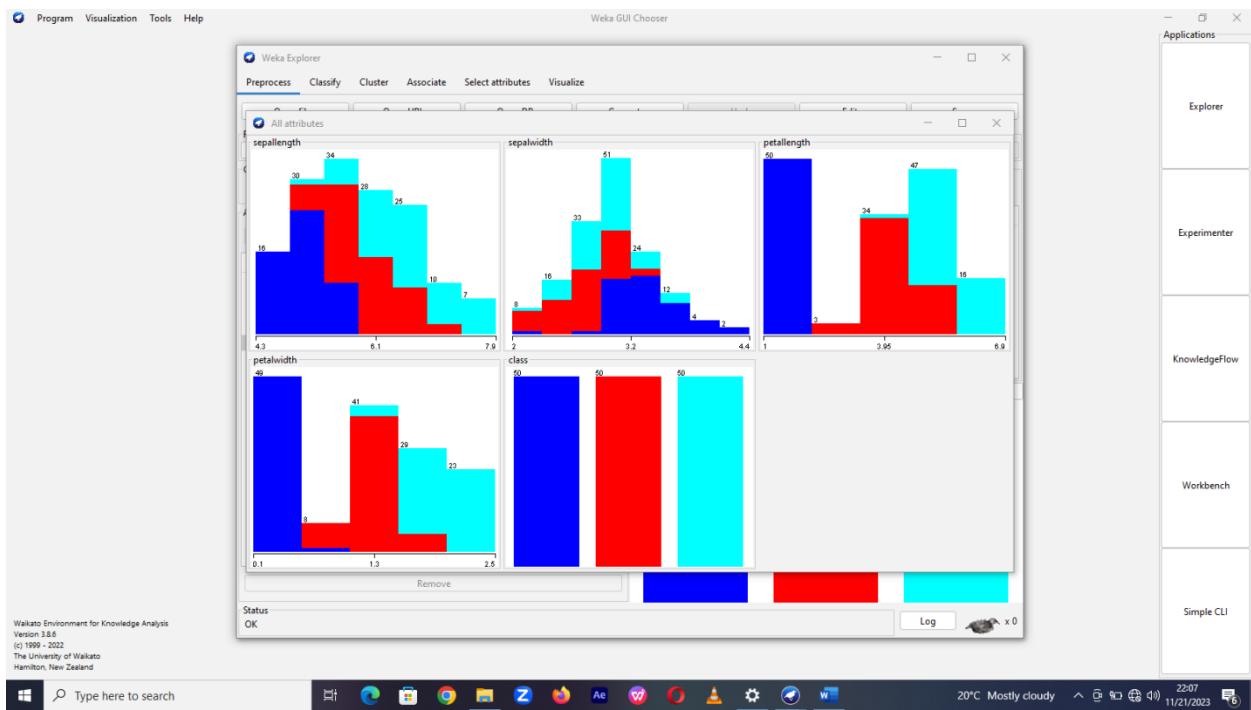
Reading the data in Iris



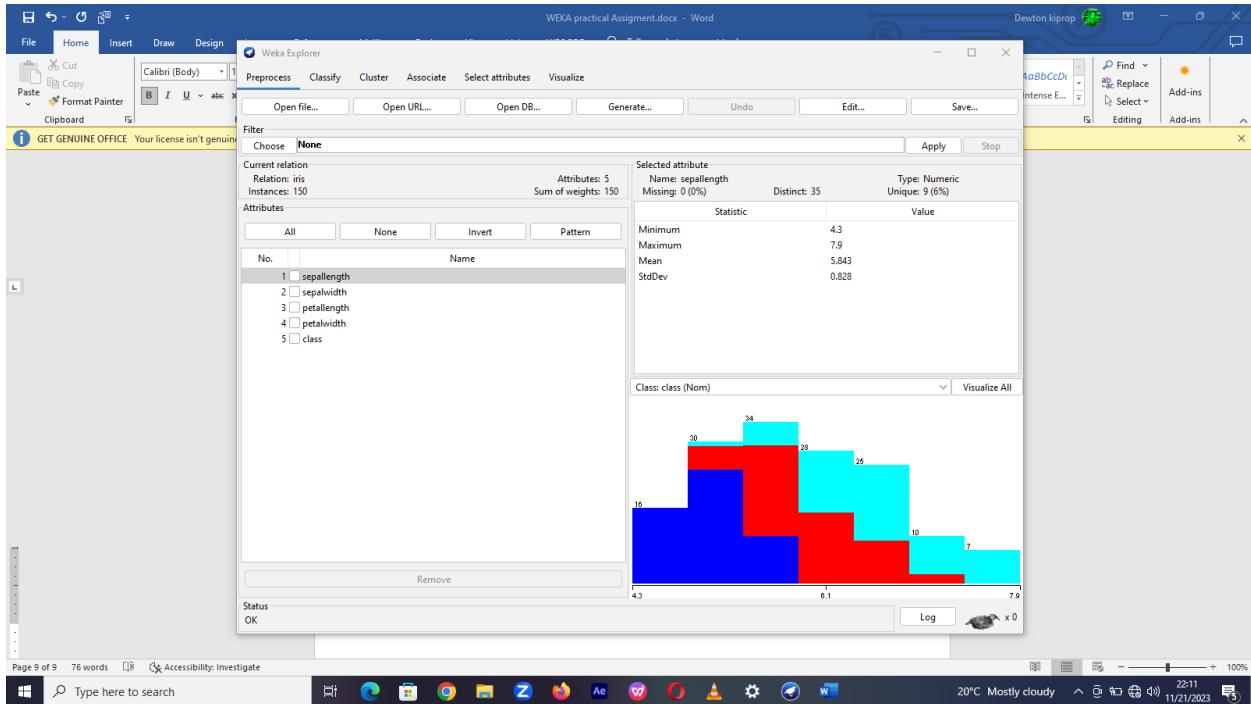
Class data reading



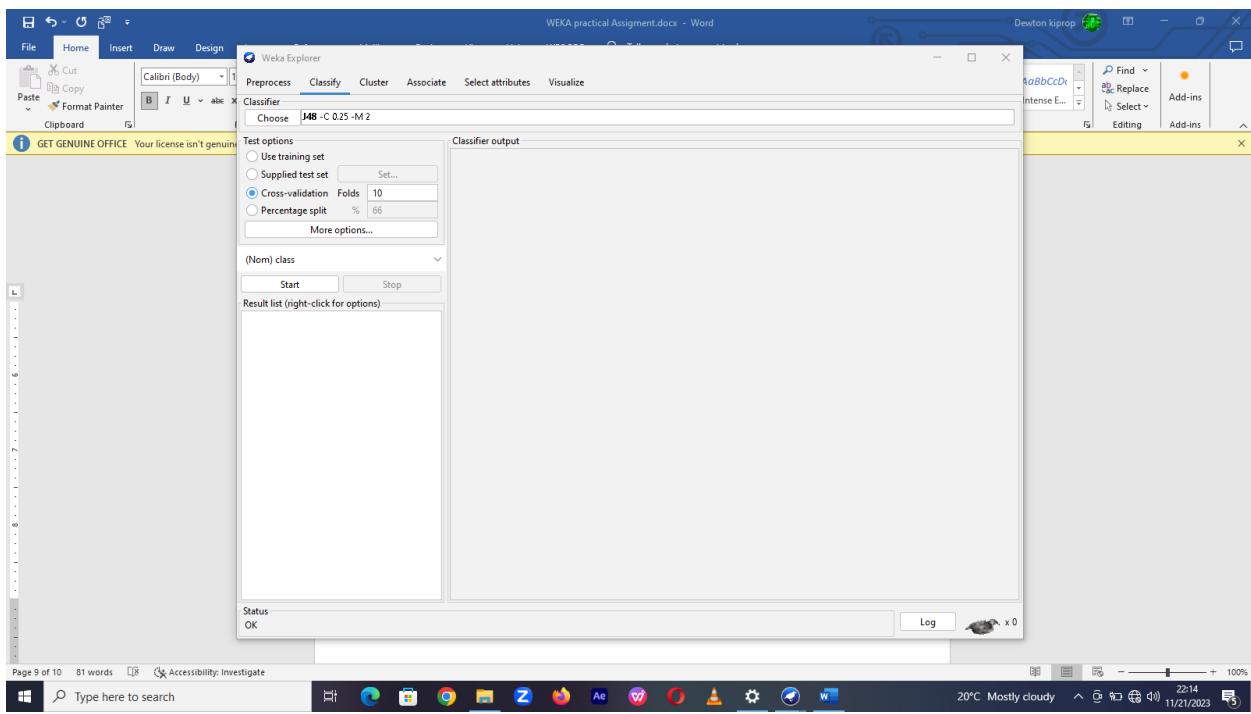
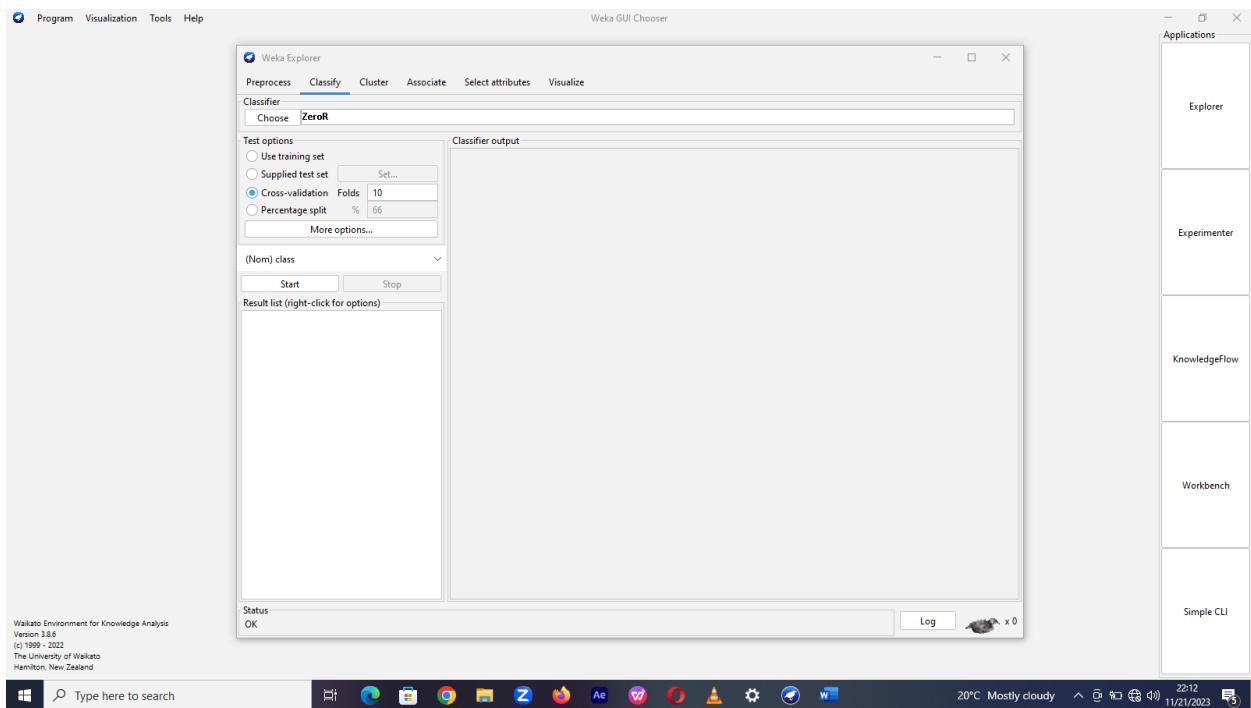
Visualization



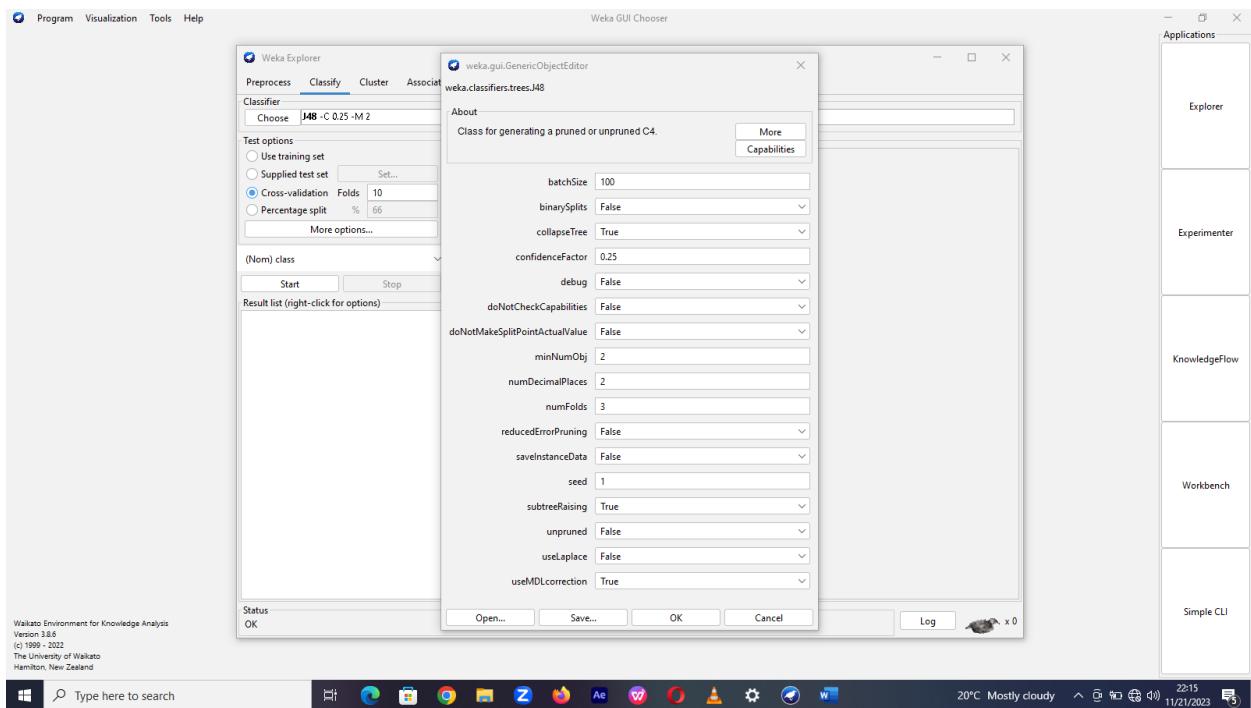
Predicting algorithms



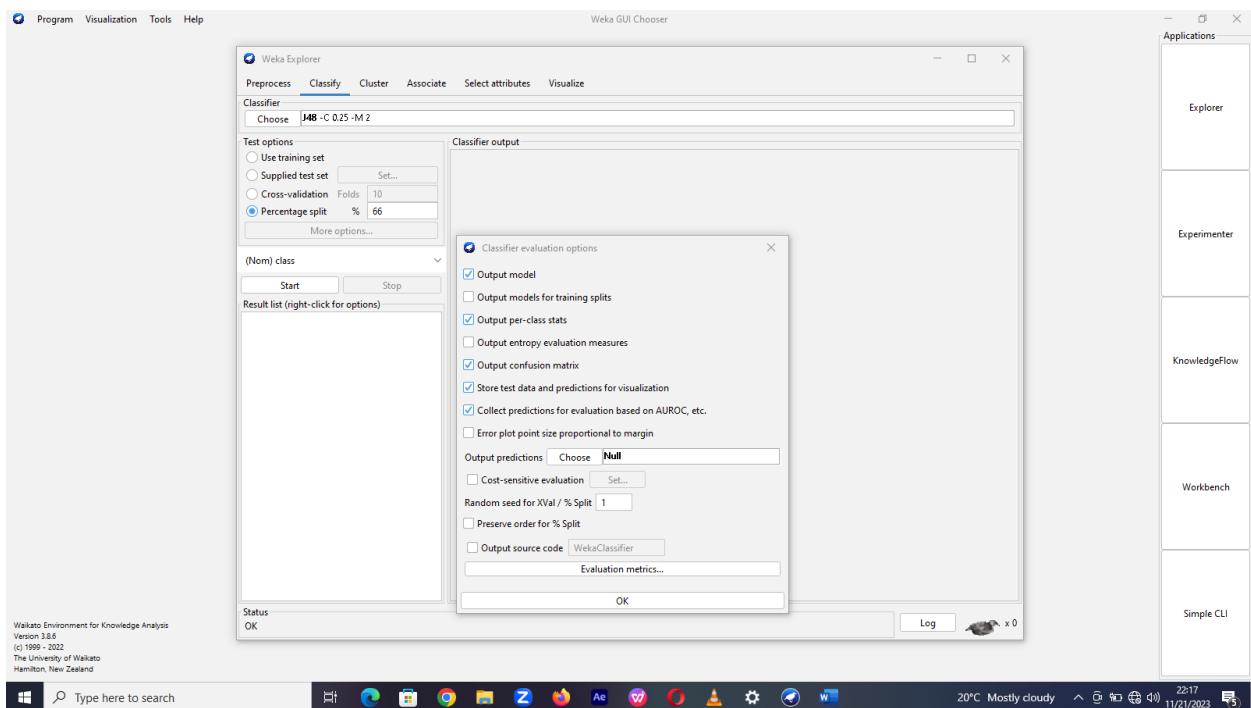
Step 2 choosing classify tab



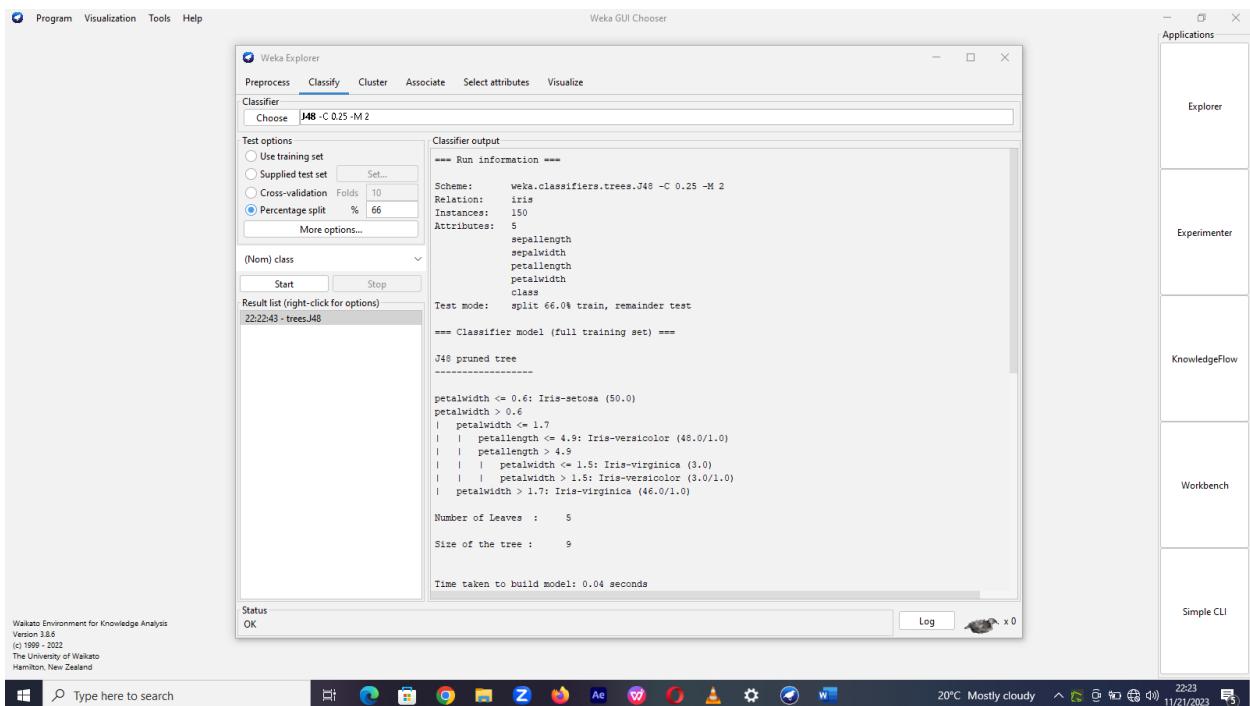
Clicking the J48 -C 0.25 -M 2



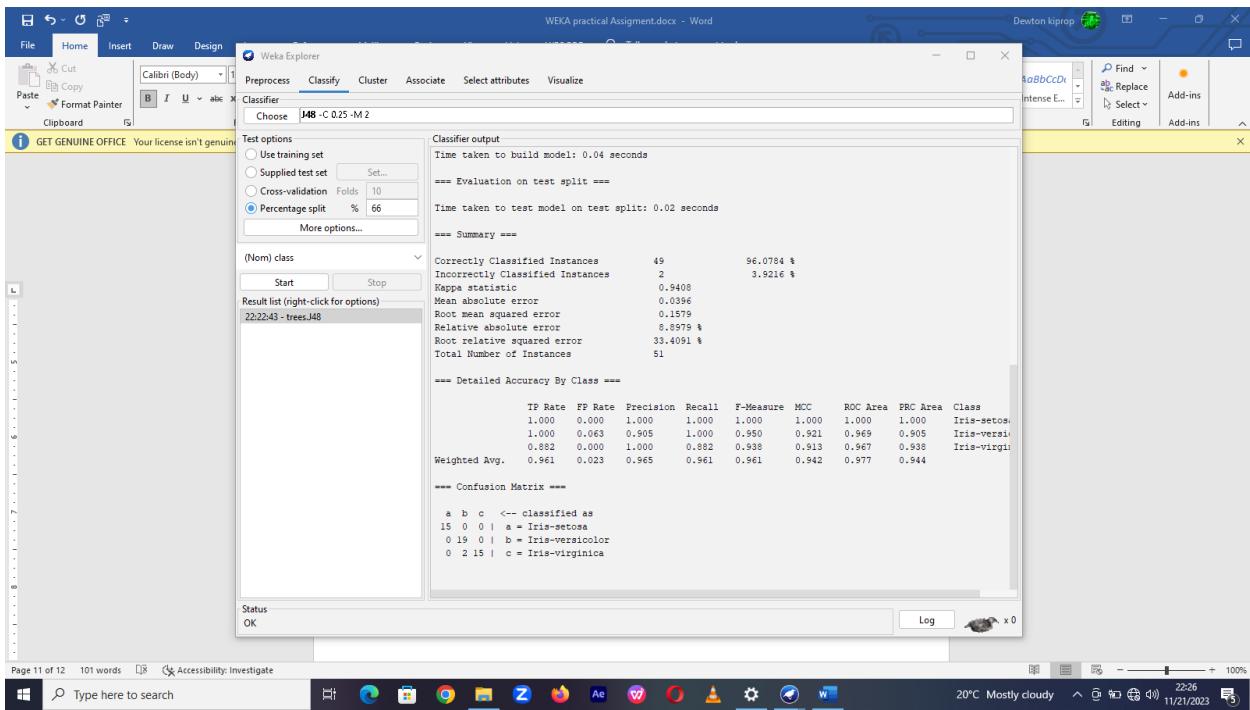
After accepting results and clicking more options



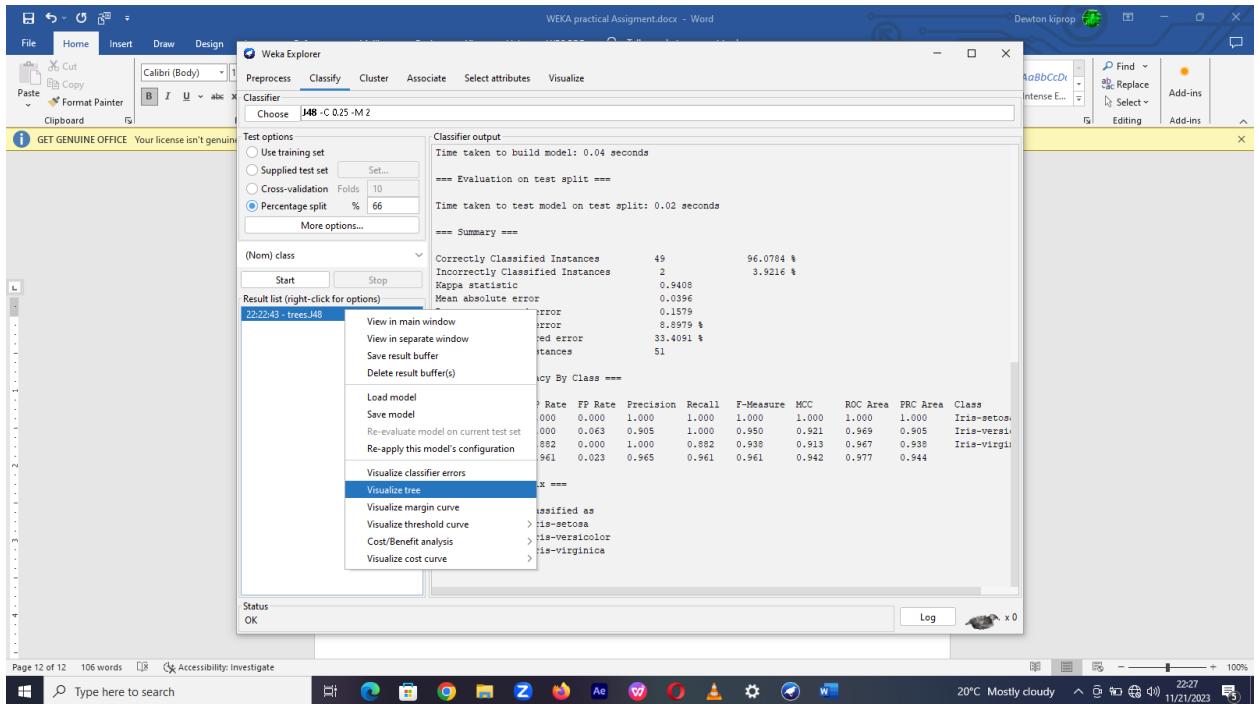
After clicking start button



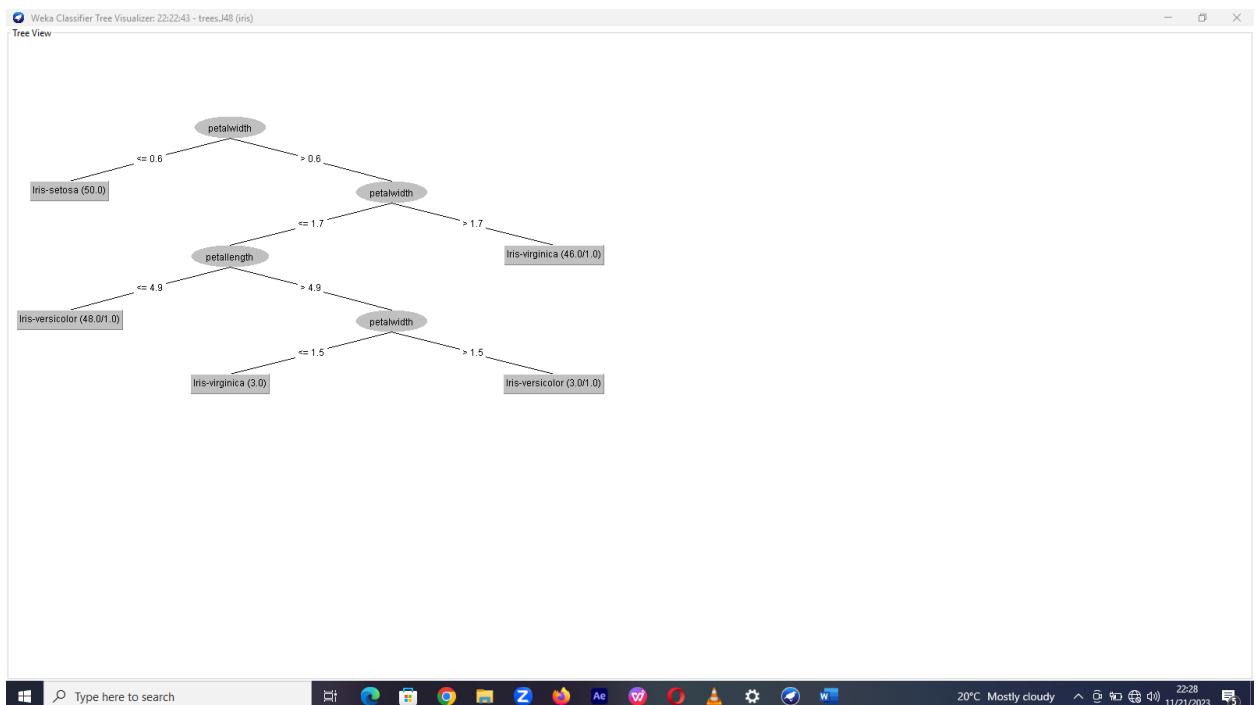
Scrolling down



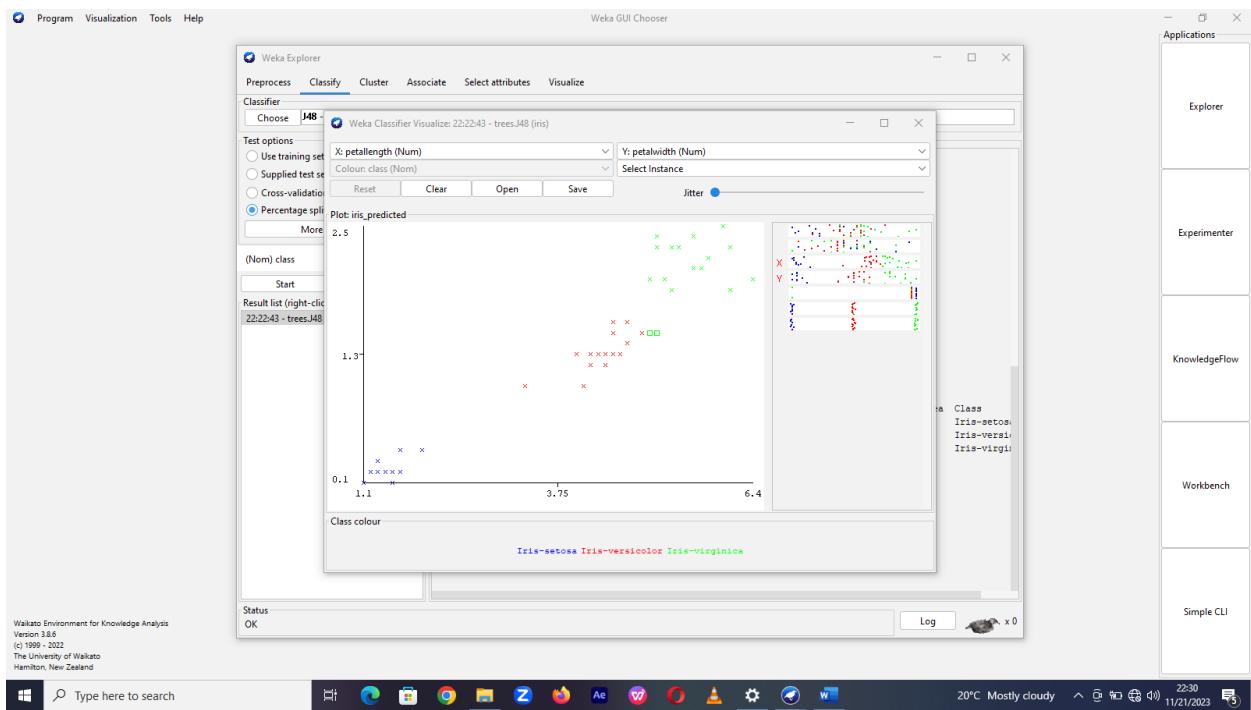
Right clicking on the model



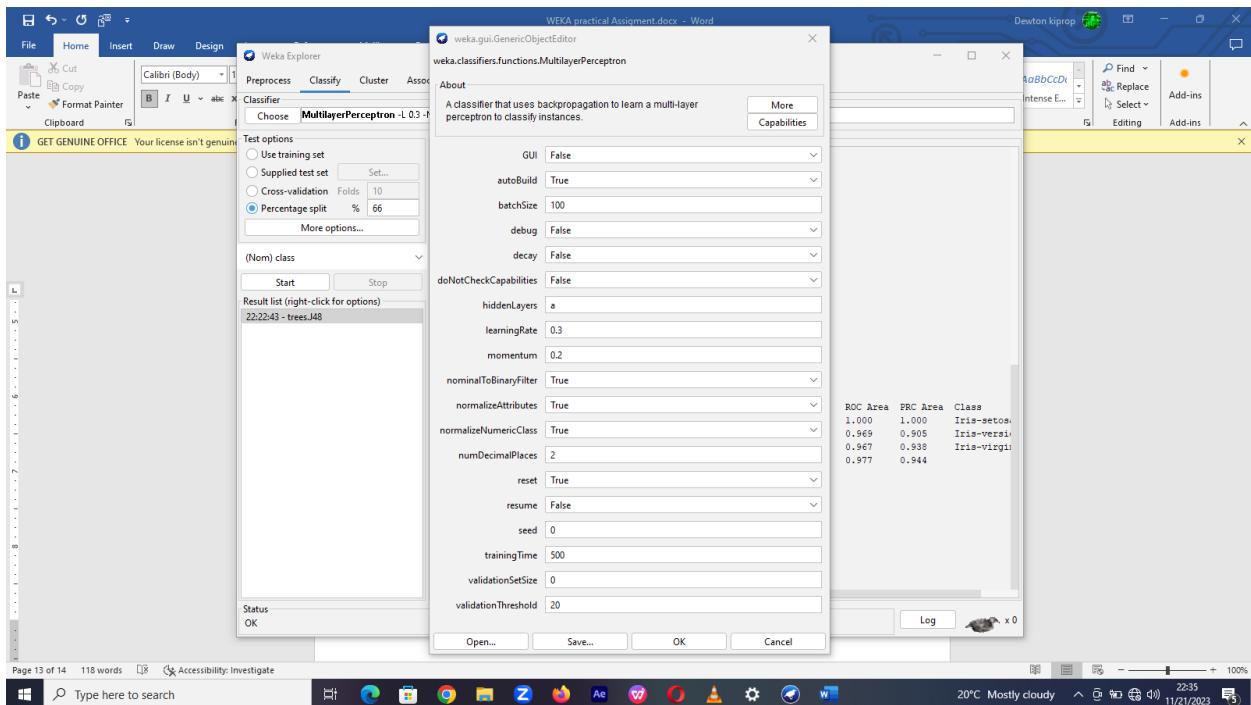
The results

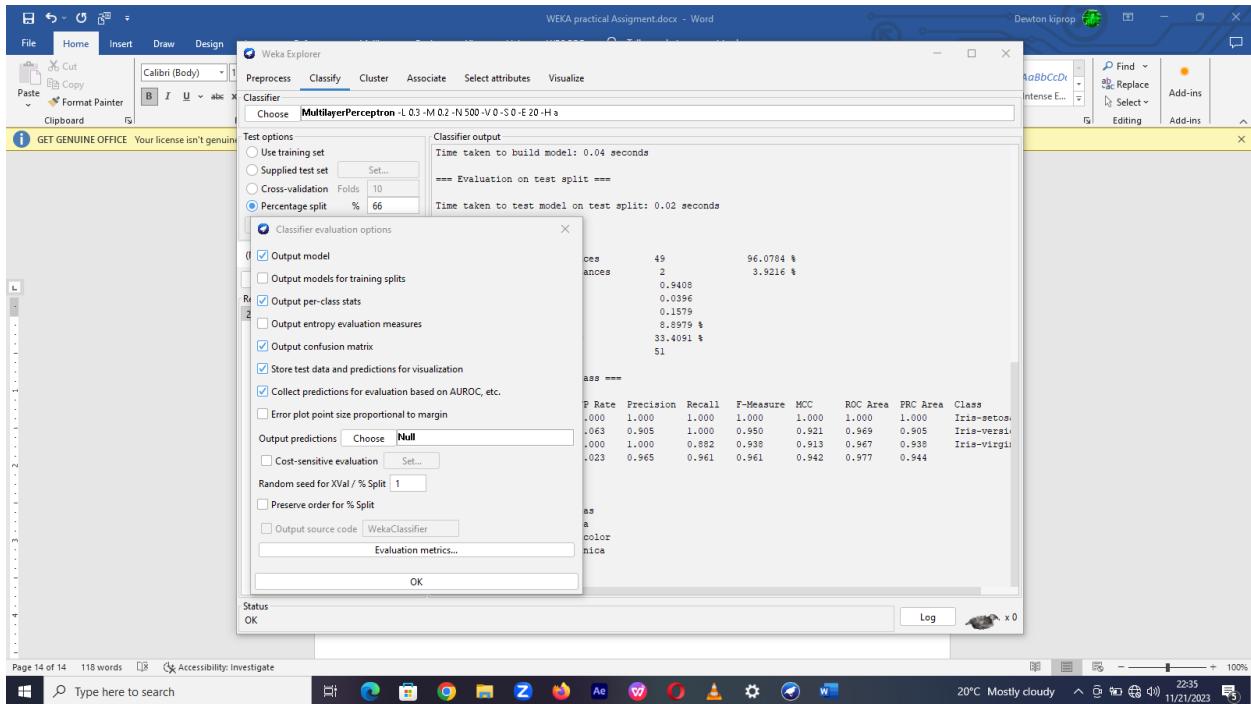


Right clicking and selecting visualize errors

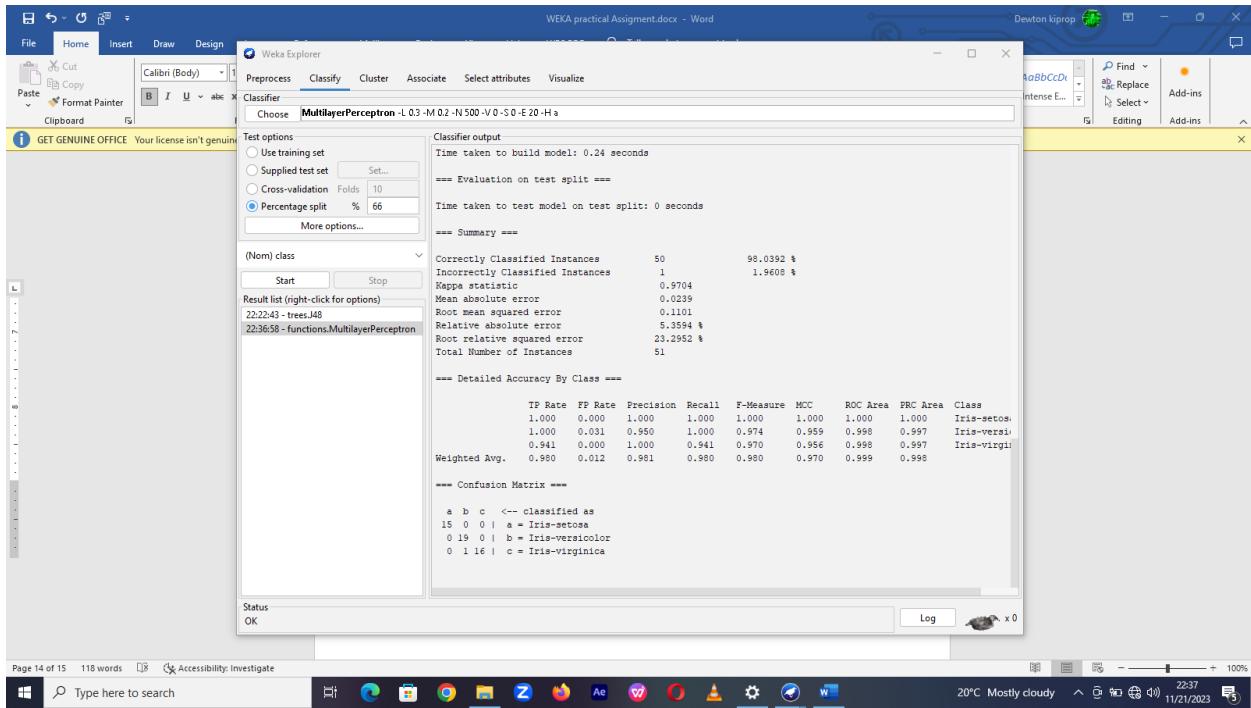


Creating ANN for Iris

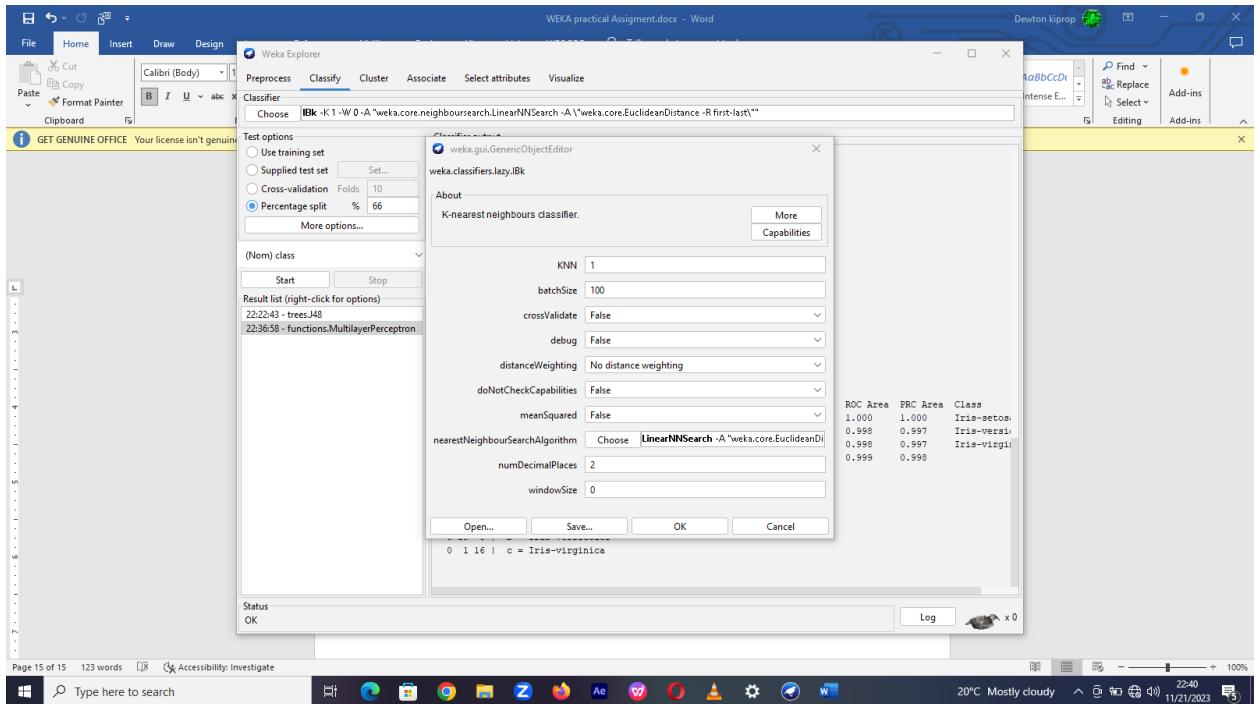




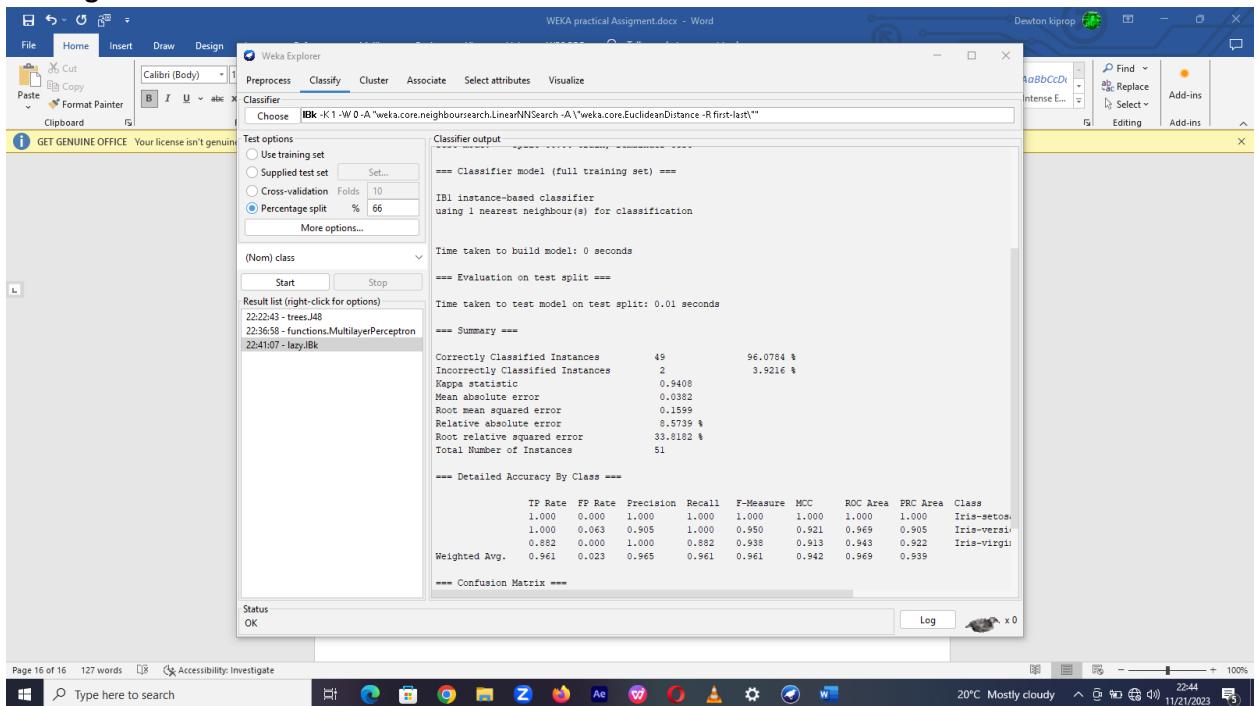
Clicking Start



Creating IBk model



Clicking start for K=1



WEKA practical Assignment.docx - Word

Dewton kiprop

Weka Explorer

Classifier Choose: IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"

Test options

- Use training set
- Supplied test set Set...
- Cross-validation Folds 10
- Percentage split % 66**
- More options...

(Nom) class

Start Stop

Result list (right-click for options)

- 22:243 - treesJ48
- 22:368 - functions.MultilayerPerceptron
- 22:4107 - lazy.IBK

Classifier output

```

Time taken to build model: 0 seconds
===
Evaluation on test split ===
Time taken to test model on test split: 0.01 seconds
===
Summary ===

Correctly Classified Instances      49      96.0784 %
Incorrectly Classified Instances   2       3.9216 %
Kappa statistic                   0.9408
Mean absolute error               0.0382
Root mean squared error          0.1599
Relative absolute error           8.5739 %
Root relative squared error      33.8182 %
Total Number of Instances        51

===
Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   FPR Area   Class
1.000     0.000    1.000     1.000    1.000     1.000   1.000     1.000   Iris-setos
1.000     0.063    0.905     1.000    0.950     0.921     0.969     0.905   Iris-versi
0.882     0.000    1.000     0.882    0.938     0.913     0.943     0.922   Iris-virgi
0.961     0.023    0.965     0.961    0.961     0.942     0.969     0.939

===
Confusion Matrix ===

a b c   <-- classified as
15 0 0 | a = Iris-setosa
0 19 0 | b = Iris-versicolor
0 2 15 | c = Iris-virginica

```

Status OK

Log x 0

Page 15 of 16 123 words Accessibility: Investigate

20°C Mostly cloudy 22:41 11/21/2023

FOR k=5

WEKA practical Assignment.docx - Word

Dewton kiprop

Weka Explorer

Classifier Choose: IBk -K 5 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"

Test options

- Use training set
- Supplied test set Set...
- Cross-validation Folds 10
- Percentage split % 66**
- More options...

(Nom) class

Start Stop

Result list (right-click for options)

- 22:243 - treesJ48
- 22:368 - functions.MultilayerPerceptron
- 22:4107 - lazy.IBK
- 22:4618 - lazy.IBK

Classifier output

```

petalwidth
class
Test mode: split 66.0% train, remainder test
===
Classifier model (full training set) ===
IBI instance-based classifier
using 5 nearest neighbour(s) for classification

Time taken to build model: 0 seconds
===
Evaluation on test split ===
Time taken to test model on test split: 0.01 seconds
===
Summary ===

Correctly Classified Instances      50      98.0392 %
Incorrectly Classified Instances   1       1.9608 %
Kappa statistic                   0.9704
Mean absolute error               0.0261
Root mean squared error          0.0998
Relative absolute error           5.8511 %
Root relative squared error      21.1145 %
Total Number of Instances        51

===
Detailed Accuracy By Class ===

      TP Rate   FP Rate   Precision   Recall   F-Measure   MCC   ROC Area   FPR Area   Class
1.000     0.000    1.000     1.000    1.000     1.000   1.000     1.000   Iris-setos
1.000     0.031    0.950     1.000    0.974     0.959     0.999     0.997   Iris-versi
0.941     0.000    1.000     0.941    0.970     0.956     0.999     0.997   Iris-virgi

```

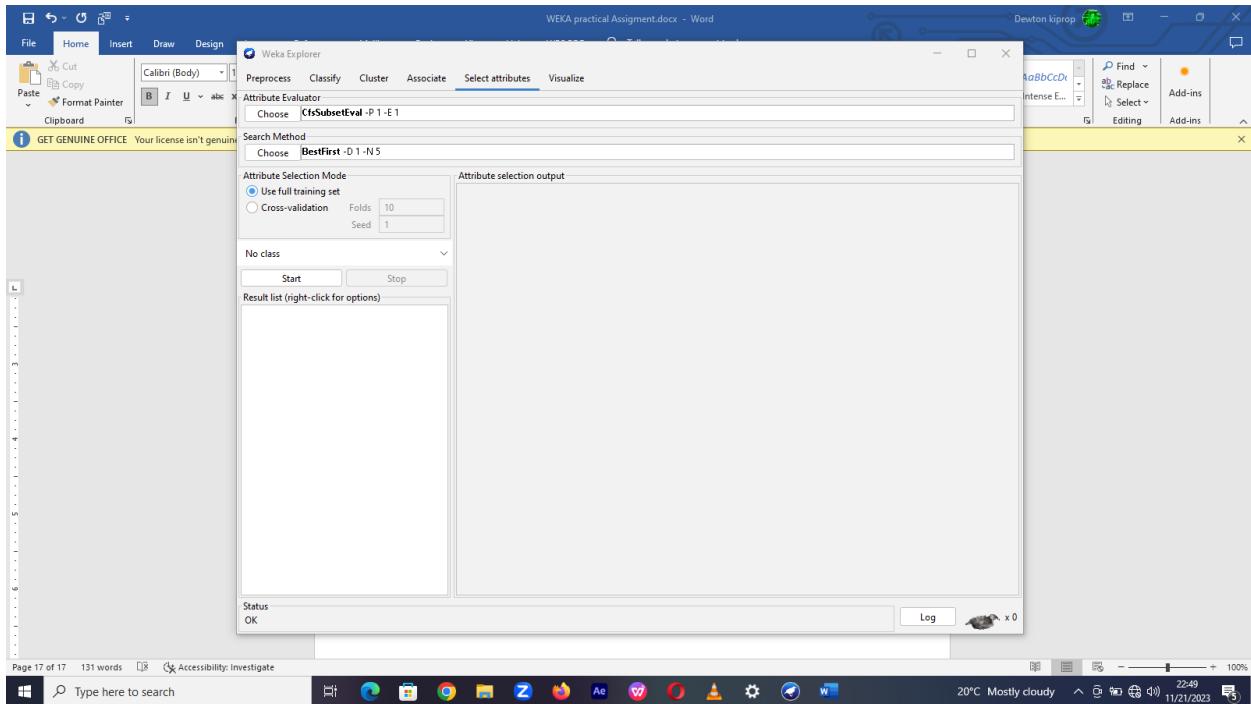
Status OK

Log x 0

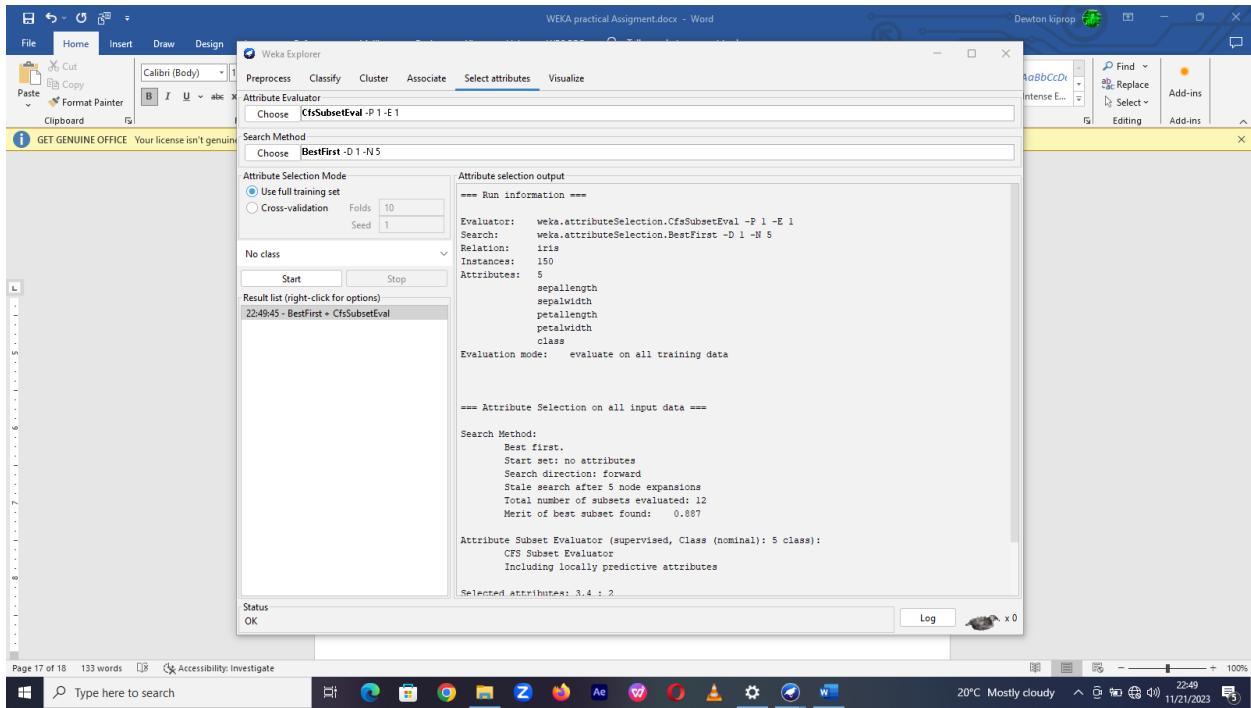
Page 16 of 17 129 words Accessibility: Investigate

20°C Mostly cloudy 22:46 11/21/2023

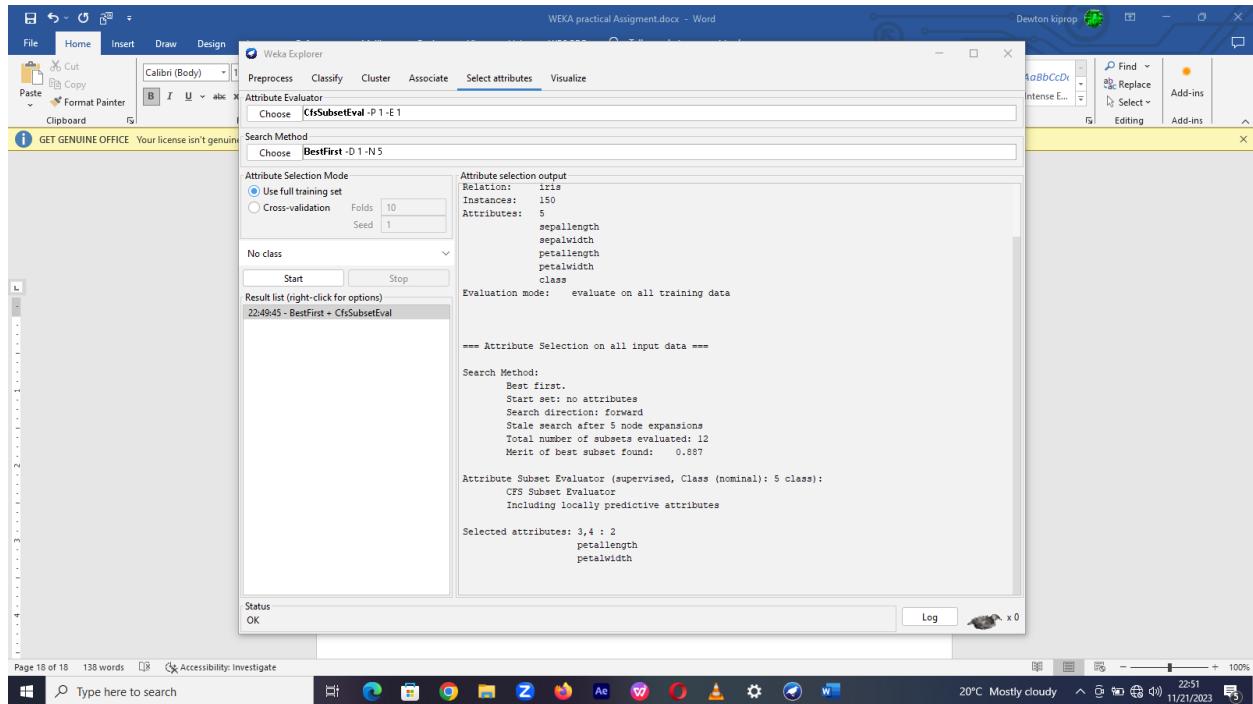
Features selection



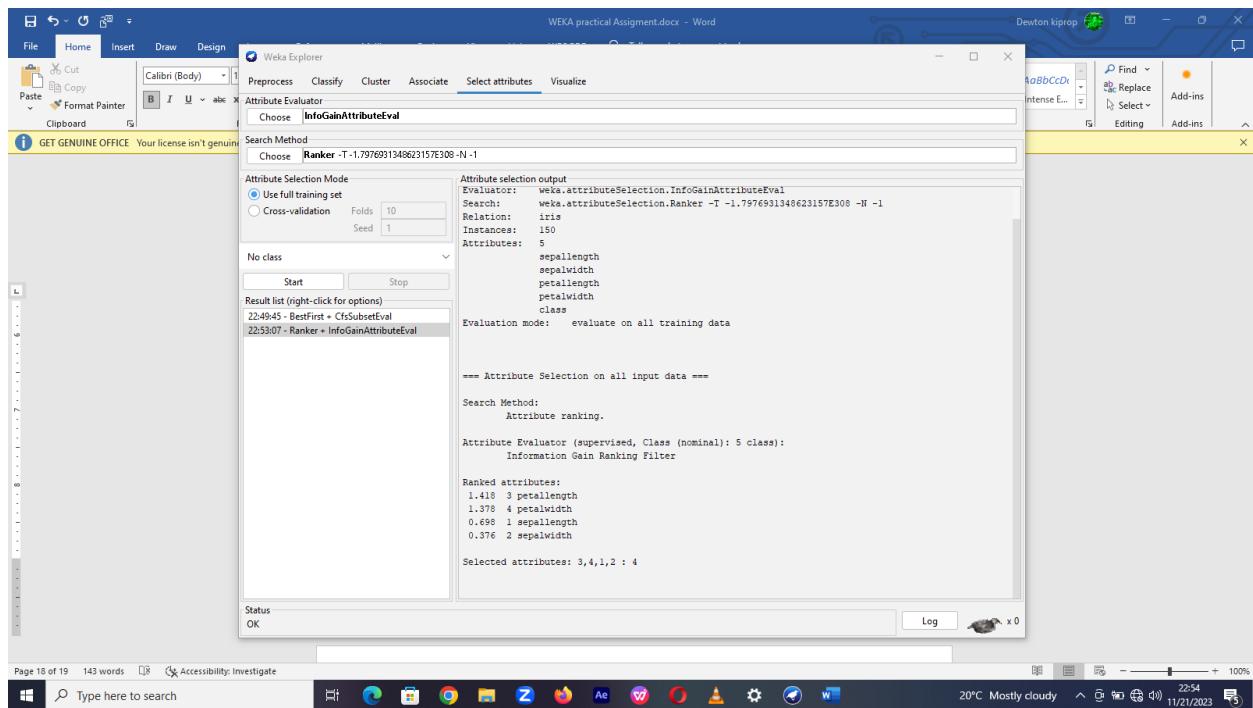
Clicking start



Best first search subset selection



Rank features by info gain



Using different data

