**Classification and Testing)**

Explanations:

The feature selection and classification system used was the same as question 3 however there was not multiple sheets the data was randomised by row, and split into 3 sections, training, Test1, Test2 it was then saved to individual csv files to save for Weka use. The classification model was creating using the training set and then tested on all 3. The details of the code and functions used are in question 3. The following code shuffles the rows within the Iris dataset.

Text

Description automatically generated

Features Selected:

Text

Description automatically generated

Boruta determined that all the features within the dataset are important for determining the CLASS a flower belongs to, so the original dataset will be used for classification.

However Wekas classification makes use of only PetalWidth, and Pedal Length, so the features selected can be narrowed down to PetalWidth and Pedal Length.

Classification:

Weka outputs the classification system below, with the pruned tree being the actual system that will be used to test on the different datasets.

Text

Description automatically generated

Testing:

Training Set:

Table

Description automatically generated with medium confidence

Test Set 1:

Table

Description automatically generated

Test Set 2:

Table

Description automatically generated

Results:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Training | Test AD | Test MCI |
| Sensitivity | 0.94 | 0.94 | 0.94 |
| Specificity | 0.97 | 0.98 | 0.97 |
| Accuracy | 94% | 94% | 94% |
| F1-Score | 0.94 | 0.94 | 0.94 |
| MCC | 0.913 | 0.914 | 0.910 |
| Youden’s J | 0.91 | 0.92 | 0.91 |

Evaluation:

The classification system created by the training set was proven to be highly effective when tested on the test sets.