ARTIFICIAL INTELLIGENCE NEURAL NETWORKS WEKA PROJECT

Yash Thakor, Shaquel Seymour, Mo Izeiroski, Raymond Padilla

TABLE OF CONTENTS

01

Introduction

Introduction

Iris Dataset

It has four independent variable/predictors for each flower:

Sepal length
Sepal width
Petal Length
Petal Width

Researchers studied three different species of Iris flowers:

Iris setosa
Iris versicolor
Iris virginica

02

Preprocessing

Preprocessing

Removing Duplicates

It is important to remove all duplicates because they can skew the results of the experiment

Removing Null Instances

Missing values need to be removed or replaced because they can negatively skew the results of the experiment

Normalizing Numeric Attributes

This step can assist in eliminating any biases in the data.

03

Parameters

Parameters

Learning Rate

Default Value = 0.3

Our Values = 0.3, 0.6, and 0.9

The higher the number, the more the weight of the network will be adjusted during each trial

Number of Hidden Layers

Default Value = a

Our Values = 1, 2, 3

A higher number allows the network to learn complicated relationships between input and output

Momentum

Default Value = 0.2

Our Values = 0.3, 0.6, 0.9

A higher number determines how much of an impact previous weight updates have on current ones

04

Results

Results (Run 1)

```
    Node 0
Class Iris-versicolor
    Input
    Node 1
Class Iris-virginica
    Input
    Node 2


Time taken to build model: 0.07 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         141               95.9184 %
Incorrectly Classified Instances         6                4.0816 %
Kappa statistic                          0.9388
Mean absolute error                      0.1584
Root mean squared error                  0.2212
Relative absolute error                 35.6367 %
Root relative squared error             46.9133 %
Total Number of Instances              147

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Iris-setosa
                 0.880    0.000    1.000      0.880    0.936      0.910    0.996     0.993     Iris-versicolor
                 1.000    0.061    0.891      1.000    0.942      0.915    0.996     0.992     Iris-virginica
Weighted Avg.    0.959    0.020    0.964      0.959    0.959      0.941    0.997     0.995

=== Confusion Matrix ===

  a  b  c   <-- classified as
 48  0  0 |  a = Iris-setosa
  0 44  6 |  b = Iris-versicolor
  0  0 49 |  c = Iris-virginica
```

Results (Run 1)

```
=== Run information ===

Scheme:       weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.3 -N 500 -V 0 -S 0 -E 20 -H 1
Relation:     iris-weka.filters.unsupervised.instance.RemoveDuplicates-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribut
Instances:    147
Attributes:   5
              sepallength
              sepalwidth
              petallength
              petalwidth
              class
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

Sigmoid Node 0
    Inputs    Weights
    Threshold    -7.949495961787856
    Node 3    12.661062435341735
Sigmoid Node 1
    Inputs    Weights
    Threshold    -0.18820100537792367
    Node 3    -1.5308482980787899
Sigmoid Node 2
    Inputs    Weights
    Threshold    2.9299302079171228
    Node 3    -29.240039031265763
Sigmoid Node 3
    Inputs    Weights
    Threshold    -0.6817323117610186
    Attrib sepallength    0.6059804448903673
    Attrib sepalwidth    1.0484472191283536
    Attrib petallength    -2.675140105185179
    Attrib petalwidth    -2.1457856896821332
Class Iris-setosa
    Input
    Node 0
Class Iris-versicolor
    Input
```

Results (Run 2)

```
=== Run information ===

Scheme:       weka.classifiers.functions.MultilayerPerceptron -L 0.6 -M 0.6 -N 500 -V 0 -S 0 -E 20 -H 2
Relation:     iris-weka.filters.unsupervised.instance.RemoveDuplicates-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.
Instances:    147
Attributes:   5
              sepallength
              sepalwidth
              petallength
              petalwidth
              class
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

Sigmoid Node 0
    Inputs    Weights
    Threshold    -5.84597360064875
    Node 3    0.5686562292465024
    Node 4    10.609732484096137
Sigmoid Node 1
    Inputs    Weights
    Threshold    -5.244144418027655
    Node 3    10.179808305446638
    Node 4    -10.261251046201334
Sigmoid Node 2
    Inputs    Weights
    Threshold    5.301203216624713
    Node 3    -10.289425109400467
    Node 4    -3.9463784179187464
Sigmoid Node 3
    Inputs    Weights
    Threshold    15.628060542954993
    Attrib sepallength    4.88667383627808
    Attrib sepalwidth    6.949972215611563
    Attrib petallength    -27.20239529191818
    Attrib petalwidth    -17.180354473433034
Sigmoid Node 4
    Inputs    Weights
```

Results (Run 2)

```
Sigmoid Node 4
    Inputs    Weights
    Threshold    -3.7760085331694175
    Attrib sepallength    -1.1488402062960867
    Attrib sepalwidth    3.2308965046306817
    Attrib petallength    -4.379968759731345
    Attrib petalwidth    -4.44952936055677
Class Iris-setosa
    Input
    Node 0
Class Iris-versicolor
    Input
    Node 1
Class Iris-virginica
    Input
    Node 2


Time taken to build model: 0.06 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         141               95.9184 %
Incorrectly Classified Instances         6                4.0816 %
Kappa statistic                          0.9388
Mean absolute error                      0.0347
Root mean squared error                  0.1553
Relative absolute error                  7.818  %
Root relative squared error             32.9326 %
Total Number of Instances              147

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Iris-setosa
                 0.920    0.021    0.958      0.920    0.939      0.909    0.952     0.927     Iris-versicolor
                 0.959    0.041    0.922      0.959    0.940      0.910    0.973     0.907     Iris-virginica
Weighted Avg.    0.959    0.021    0.960      0.959    0.959      0.939    0.975     0.944
```

```
=== Confusion Matrix ===

  a  b  c   <-- classified as
 48  0  0 |  a = Iris-setosa
  0 46  4 |  b = Iris-versicolor
  0  2 47 |  c = Iris-virginica
```

Results (Run 3)

```
=== Run information ===

Scheme:       weka.classifiers.functions.MultilayerPerceptron -L 0.9 -M 0.9 -N 500 -V 0 -S 0 -E 20 -H 3
Relation:     iris-weka.filters.unsupervised.instance.RemoveDuplicates-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.N
Instances:    147
Attributes:   5
              sepallength
              sepalwidth
              petallength
              petalwidth
              class
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

Sigmoid Node 0
    Inputs    Weights
    Threshold    -6.623056654182139
    Node 3    0.15385685409647334
    Node 4    6.716877755436658
    Node 5    6.842560559211109
Sigmoid Node 1
    Inputs    Weights
    Threshold    -4.384022378913526
    Node 3    8.745871821505322
    Node 4    -5.2352046765575535
    Node 5    -6.771851425747993
Sigmoid Node 2
    Inputs    Weights
    Threshold    4.329334186793814
    Node 3    -8.496975990180806
    Node 4    -6.678786749015226
    Node 5    -5.93593939381201
Sigmoid Node 3
    Inputs    Weights
    Threshold    36.14333075050978
    Attrib sepallength    8.806855559086614
    Attrib sepalwidth    17.031558565788682
    Attrib petallength    -60.39053828727485
```

Results (Run 3)

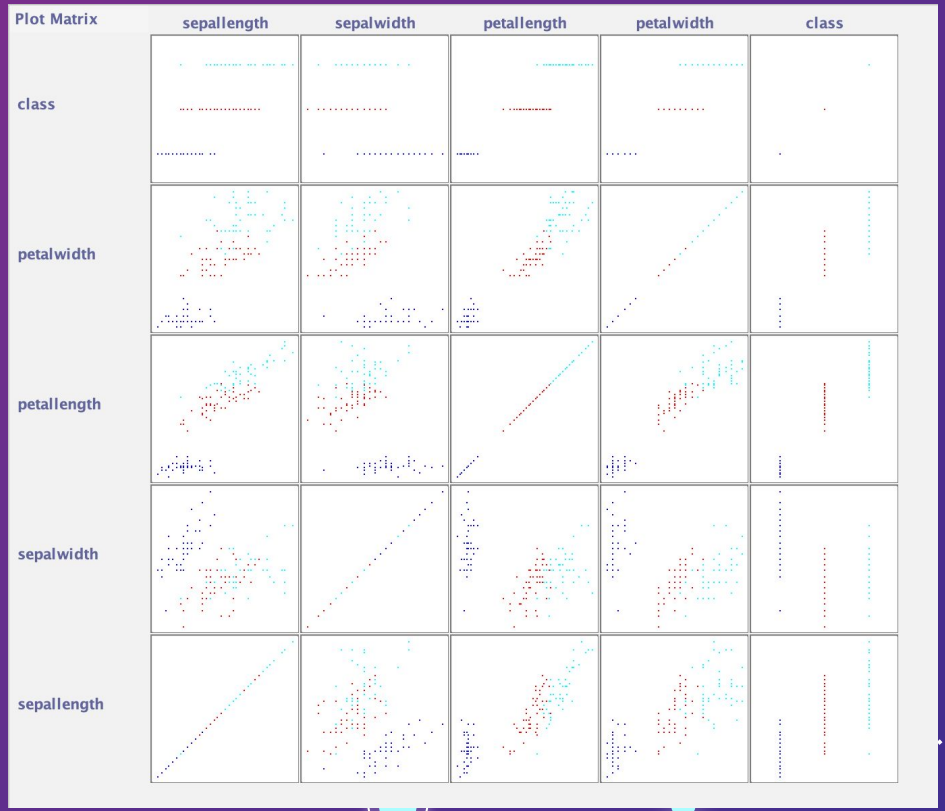
```
    Attrib petallength    -60.39053828727485
    Attrib petalwidth    -40.2391014121553
Sigmoid Node 4
    Inputs    Weights
    Threshold    -3.060842290250805
    Attrib sepallength    -0.3451554903138964
    Attrib sepalwidth    3.015163439352126
    Attrib petallength    -3.54424341148039
    Attrib petalwidth    -3.9609614981359473
Sigmoid Node 5
    Inputs    Weights
    Threshold    -4.001392493104583
    Attrib sepallength    -0.5260090547337303
    Attrib sepalwidth    3.142412857495734
    Attrib petallength    -4.161320438022374
    Attrib petalwidth    -4.417078771764026
Class Iris-setosa
    Input
    Node 0
Class Iris-versicolor
    Input
    Node 1
Class Iris-virginica
    Input
    Node 2


Time taken to build model: 0.07 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         143               97.2789 %
Incorrectly Classified Instances         4                2.7211 %
Kappa statistic                          0.9592
Mean absolute error                      0.0248
Root mean squared error                  0.1309
Relative absolute error                  5.5898 %
Root relative squared error             27.7557 %
Total Number of Instances              147
```

```
=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Iris-setosa
                 0.960    0.021    0.960      0.960    0.960      0.939    0.979     0.972     Iris-versicolor
                 0.959    0.020    0.959      0.959    0.959      0.939    0.984     0.967     Iris-virginica
Weighted Avg.    0.973    0.014    0.973      0.973    0.973      0.959    0.988     0.979

=== Confusion Matrix ===

  a  b  c   <-- classified as
 48  0  0 |  a = Iris-setosa
  0 48  2 |  b = Iris-versicolor
  0  2 47 |  c = Iris-virginica
```

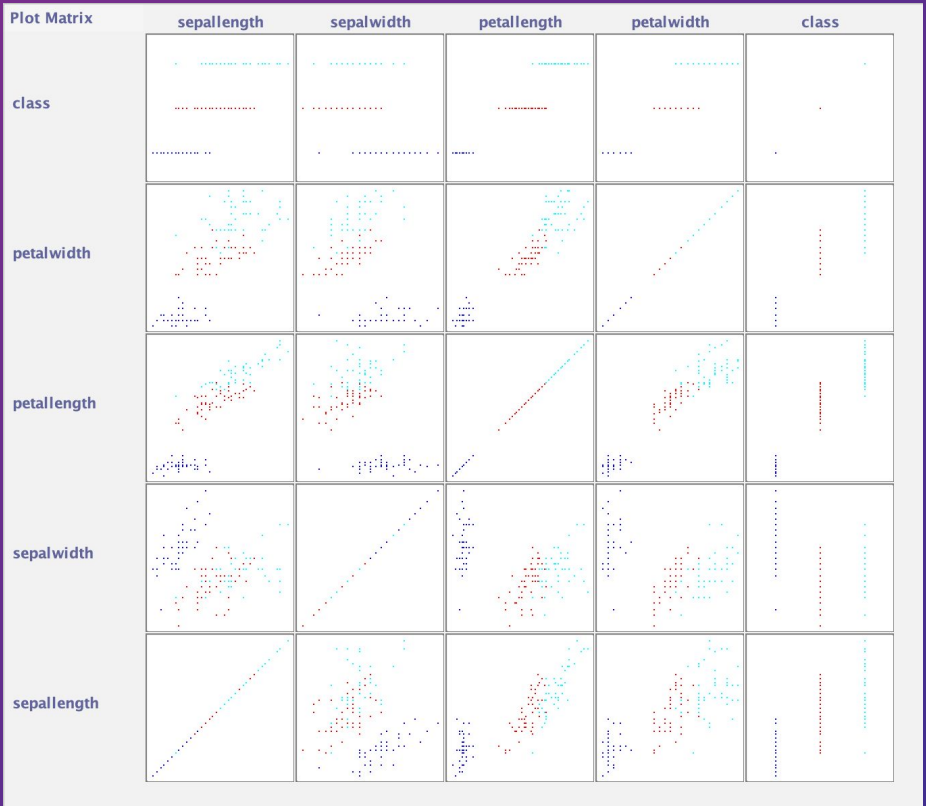
05

Visuals

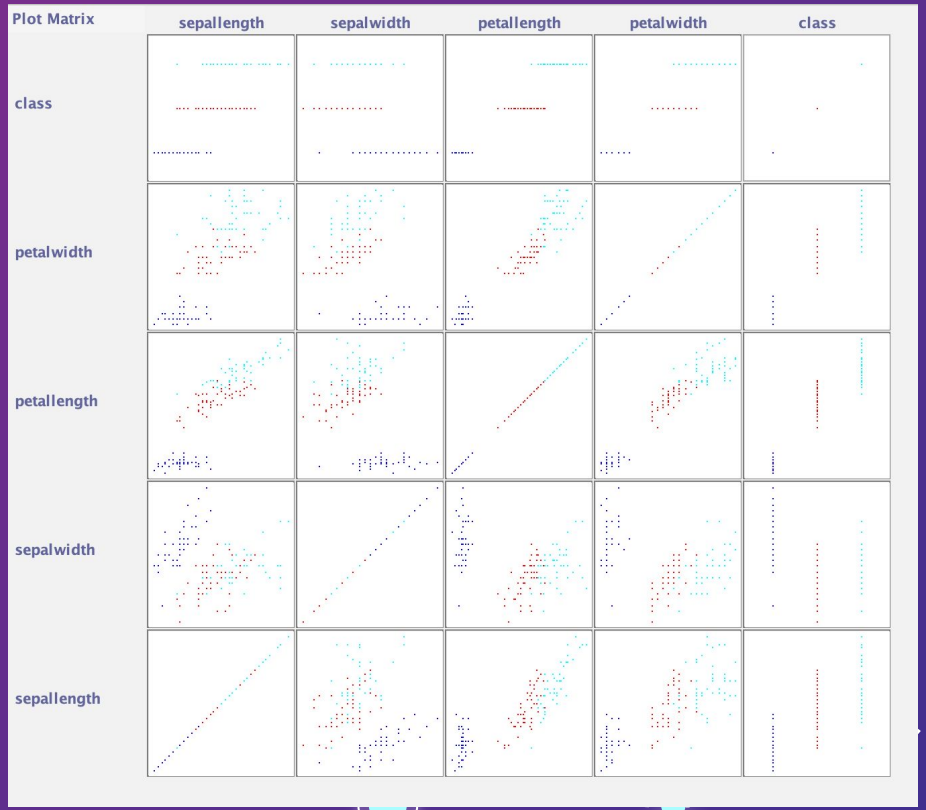
Visuals (Run 1)



Visuals (Run 2)



Visuals (Run 3)



06

Conclusion

Conclusion

It is clear from the examination of the Iris dataset that classification problems can be successfully handled by neural network learning. The data set underwent preprocessing, which enhanced the accuracy and dependability of the results. The studies conducted with different parameter settings produced variable degrees of accuracy, highlighting the significance of precise parameter selection in the neural network learning process. The provided slides allow you to view the experiment's findings. It is crucial to remember that the quality and completeness of the input data as well as the accessibility of computational resources for model training and testing are constraints on the efficiency of the neural network learning approach.