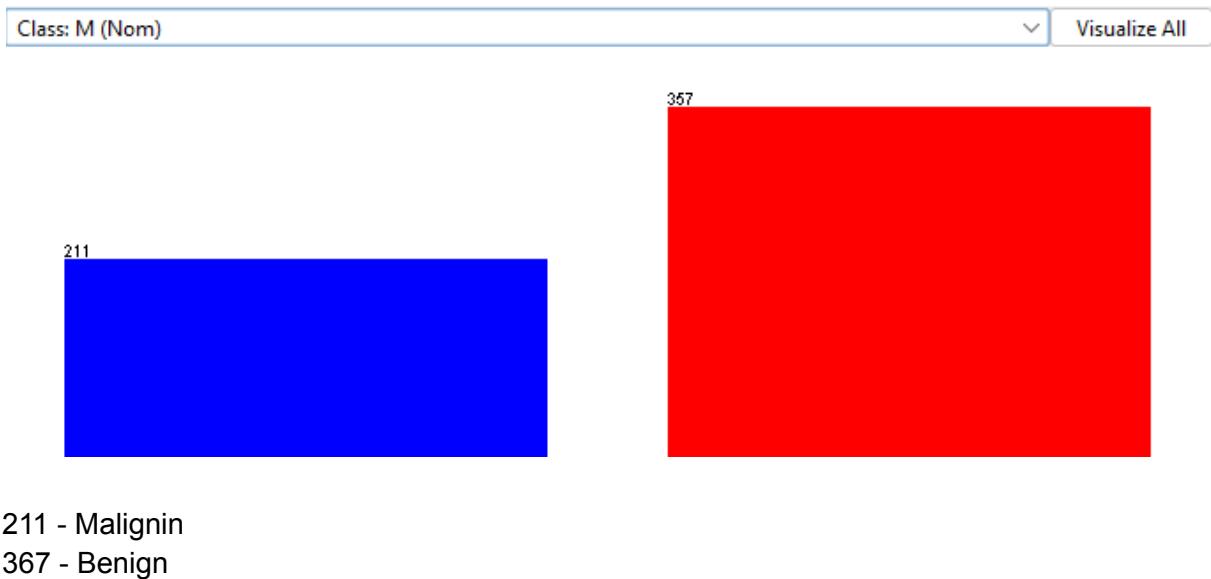


Exp 3 : Breast Cancer Wisconsin Dataset

Dataset Distribution and Classes:

- 2 Classes ‘B’ for benign and ‘M’ for malignin



Classifiers

Logistic -

Datasplit - 80 % => train , 20% => test

Learning rate => 1.0E-8

```
==== Evaluation on test split ====
Time taken to test model on test split: 0 seconds
==== Summary ====
Correctly Classified Instances      102          89.4737 %
Incorrectly Classified Instances   12           10.5263 %
Total Number of Instances         114
==== Detailed Accuracy By Class ====
          TP Rate  FP Rate  Precision  Recall   F-Measure  ROC Area  PRC Area  Class
          0.902    0.110    0.822     0.902    0.860     0.944    0.855     M
          0.890    0.098    0.942     0.890    0.915     0.933    0.935     B
Weighted Avg.      0.895    0.102    0.899     0.895    0.896     0.937    0.906
==== Confusion Matrix ====
      a   b   <- classified as
37   4 |  a = M
 8  65 |  b = B
```

MultiLayer Perceptron

Hyperparameters

weka.gui.GenericObjectEditor X

weka.classifiers.functions.MultilayerPerceptron

About

A classifier that uses backpropagation to learn a multi-layer perceptron to classify instances.

[More](#) [Capabilities](#)

GUI	False	
autoBuild	True	
batchSize	100	
debug	False	
decay	False	
doNotCheckCapabilities	False	
hiddenLayers	a	
learningRate	0.3	
momentum	0.2	
nominalToBinaryFilter	True	
normalizeAttributes	True	
normalizeNumericClass	True	
numDecimalPlaces	2	This will normalize the class if it is numeric
reset	True	
resume	False	
seed	0	
trainingTime	500	
validationSetSize	0	

Metrics

```
==== Evaluation on test split ====

Time taken to test model on test split: 0.01 seconds

==== Summary ====

Correctly Classified Instances      112                  98.2456 %
Incorrectly Classified Instances   2                   1.7544 %
Total Number of Instances         114

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

          TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   PRC Area   Class
          0.951     0.000     1.000      0.951     0.975      1.000     1.000      M
          1.000     0.049     0.973      1.000     0.986      1.000     1.000      B
Weighted Avg.      0.982     0.031     0.983      0.982     0.982      1.000     1.000

==== Confusion Matrix ====

    a   b   <-- classified as
39  2 |  a = M
 0 73 |  b = B
```

Random Forest

Hyperparameters

KValue	0
allowUnclassifiedInstances	False
batchSize	100
breakTiesRandomly	False
debug	False
doNotCheckCapabilities	False
maxDepth	0
minNum	1.0
minVarianceProp	0.001
numDecimalPlaces	2
numFolds	0
seed	1

Metrics

```
==== Evaluation on test split ====

Time taken to test model on test split: 0 seconds

==== Summary ====

Correctly Classified Instances          104                  91.2281 %
Incorrectly Classified Instances       10                   8.7719 %
Total Number of Instances              114

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

      TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   PRC Area   Class
      0.829     0.041     0.919      0.829     0.872      0.894      0.823      M
      0.959     0.171     0.909      0.959     0.933      0.894      0.898      B
Weighted Avg.   0.912     0.124     0.913      0.912     0.911      0.894      0.871

==== Confusion Matrix ====

 a   b   <-- classified as
34   7 |  a = M
 3 70 |  b = B
```

Clustering

Simple K-means clustering

Metrics

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

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

Clustered Instances

0      188 ( 33%)
1      380 ( 67%)
```

Cluster visualization

