**KNN**

Record % correctly classified and confusion matrix (ALL positive class). Derive the confusion matrix when AML is the positive class. Calculate the class-dependent TP and FP rates. Show your work and compare to Weka.

Correctly Classified Instances 68 94.4444 %

TP Rate FP Rate Class

1.000 0.143 ALL

0.857 0.000 AML

|  |  |
| --- | --- |
|  |  |

=== Confusion Matrix ===

ALL as the positive class (Weka):

a b <-- classified as

44 0 | a = ALL

4 24 | b = AML

AML as the positive class (Derived):

b a <-- classified as

24 4 | b = AML

0 44 | a = ALL

My calculated values match the Weka results.

**KNN:** Capture the ROC curves for ALL positive and AML positive.

|  |
| --- |
|  |
|  |

**ZeroR**

Record % correctly classified and confusion matrix (ALL positive class). Derive the confusion matrix when AML is the positive class. Calculate the class-dependent TP and FP rates. Show your work and compare to Weka.

Correctly Classified Instances 44 61.1111 %

TP Rate FP Rate Class

1.000 1.000 ALL

0.000 0.000 AML

|  |  |
| --- | --- |
|  |  |

=== Confusion Matrix ===

ALL as the positive class (Weka):

a b <-- classified as

44 0 | a = ALL

28 0 | b = AML

AML as the positive class (Derived):

b a <-- classified as

0 28 | b = AML

0 44 | a = ALL

My calculated values match the Weka results.

**ZeroR:** Capture the ROC curves for ALL positive and AML positive.

|  |
| --- |
|  |
|  |

**KNN**

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 5 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

Relation: leukemia gene expression data names changed

Instances: 72

Attributes: 151

[list of attributes omitted]

Test mode: 5-fold cross-validation

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

IB1 instance-based classifier

using 5 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 68 94.4444 %

Incorrectly Classified Instances 4 5.5556 %

Kappa statistic 0.88

Mean absolute error 0.0917

Root mean squared error 0.2235

Relative absolute error 19.2509 %

Root relative squared error 45.7979 %

Total Number of Instances 72

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

1.000 0.143 0.917 1.000 0.957 0.886 0.976 0.978 ALL

0.857 0.000 1.000 0.857 0.923 0.886 0.976 0.967 AML

Weighted Avg. 0.944 0.087 0.949 0.944 0.944 0.886 0.976 0.973

=== Confusion Matrix ===

a b <-- classified as

44 0 | a = ALL

4 24 | b = AML

**ZeroR**

=== Run information ===

Scheme: weka.classifiers.rules.ZeroR

Relation: leukemia gene expression data names changed

Instances: 72

Attributes: 151

[list of attributes omitted]

Test mode: 5-fold cross-validation

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

ZeroR predicts class value: ALL

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 44 61.1111 %

Incorrectly Classified Instances 28 38.8889 %

Kappa statistic 0

Mean absolute error 0.4765

Root mean squared error 0.4879

Relative absolute error 100 %

Root relative squared error 100 %

Total Number of Instances 72

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

1.000 1.000 0.611 1.000 0.759 ? 0.469 0.596 ALL

0.000 0.000 ? 0.000 ? ? 0.469 0.373 AML

Weighted Avg. 0.611 0.611 ? 0.611 ? ? 0.469 0.509

=== Confusion Matrix ===

a b <-- classified as

44 0 | a = ALL

28 0 | b = AML