Latexia

Kuzmin Andriy

16-12-2017

Part I

Esercizio 1.0

Chiede di prendere come centro 0,0 e distinguere i punti entro il raggio 1 e quadrato di lato 2.5 e centro 0,0 con esclusione della regione 1 Usiamo User-classifier di Weka specificando con strumenti visuali Dati: circletrain.arff training set (100 esempi) circletest.arff test set (100 campioni diversi)

0.1 Rettangolo sotto-misura:

=== Run information ===

Scheme: weka.classifiers.trees.UserClassifier Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

Split on x AND y (In Set): N1 c(39.0) Split on x AND y (Not in Set): N2 q(61.0/14.0)

Time taken to build model: 246.78 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances 73 73 Incorrectly Classified Instances 27 27 Kappa statistic 0.4906 Mean absolute error 0.2998 Root mean squared error 0.43 Relative absolute error 60.6811 Root relative squared error 86.8747 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.975 0.433 0.600 0.975 0.743 0.556 0.771 0.595 q 0.567 0.025 0.971 0.567 0.716 0.556 0.771 0.810 c Weighted Avg. 0.730 0.188 0.823 0.730 0.727 0.556 0.771 0.724

=== Confusion Matrix ===

a b ¡– classified as 39 1 — a = q 26 34 — b = c

0.2 Rettangolo oltre-misura:

=== Run information ===

Scheme: weka.classifiers.trees.UserClassifier Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

Split on x AND y (In Set): N1 c(61.0/8.0) Split on x AND y (Not in Set): N2 q(39.0)

Time taken to build model: 19.86 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0 seconds

=== Summary ===

Correctly Classified Instances 87 87 Incorrectly Classified Instances 13 13 Kappa statistic 0.7303 Mean absolute error 0.1916 Root mean squared error 0.3527 Relative absolute error 38.7842 Root relative squared error 71.2623 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.850 0.117 0.829 0.850 0.840 0.730 0.867 0.765 q 0.883 0.150 0.898 0.883 0.891 0.730 0.867 0.864 c Weighted Avg. 0.870 0.137 0.871 0.870 0.870 0.730 0.867 0.824

=== Confusion Matrix ===

a b ¡– classified as 34 6 — a = q 7 53 — b = c

0.3 Con 5 lati:

=== Run information ===

Scheme: weka.classifiers.trees.UserClassifier Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

Split on x AND y (In Set): N1 c(52.0) Split on x AND y (Not in Set): N2 q(48.0/1.0)

Time taken to build model: 345.24 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0 seconds

=== Summary ===

Correctly Classified Instances 91 91 Incorrectly Classified Instances 9 9 Kappa statistic 0.8133 Mean absolute error 0.0965 Root mean squared error 0.2968 Relative absolute error 19.5213 Root relative squared error 59.9663 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.900 0.083 0.878 0.900 0.889 0.813 0.908 0.830 q 0.917 0.100 0.932 0.917 0.924 0.813 0.908 0.905 c Weighted Avg. 0.910 0.093 0.911 0.910 0.910 0.813 0.908 0.875

=== Confusion Matrix ===

a b ¡– classified as 36 4 — a = q 5 55 — b = c

0.4 Con 6 lati:

=== Run information ===

Scheme: weka.classifiers.trees.UserClassifier Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

Split on x AND y (In Set): N1 c(52.0) Split on x AND y (Not in Set): N2 q(48.0/1.0)

Time taken to build model: 28.94 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0 seconds

=== Summary ===

Correctly Classified Instances 96 96 Incorrectly Classified Instances 4 4 Kappa statistic 0.918 Mean absolute error 0.0475 Root mean squared error 0.1963 Relative absolute error 9.6131 Root relative squared error 39.6552 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 1.000 0.067 0.909 1.000 0.952 0.921 0.967 0.909 q 0.933 0.000 1.000 0.933 0.966 0.921 0.967 0.973 c Weighted Avg. 0.960 0.027 0.964 0.960 0.960 0.921 0.967 0.948

=== Confusion Matrix ===

a b ¡– classified as 40 0 — a = q 4 56 — b = c

0.5 Con 7 lati:

=== Run information ===

Scheme: weka.classifiers.trees.UserClassifier Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

Split on x AND y (In Set): N1 c(53.0) Split on x AND y (Not in Set): N2 q(47.0)

Time taken to build model: 37.67 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0 seconds

=== Summary ===

Correctly Classified Instances 92 92 Incorrectly Classified Instances 8 8 Kappa statistic 0.8361 Mean absolute error 0.08 Root mean squared error 0.2828 Relative absolute error 16.1905 Root relative squared error 57.1449 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.950 0.100 0.864 0.950 0.905 0.839 0.925 0.840 q 0.900 0.050 0.964 0.900 0.931 0.839 0.925 0.928 c Weighted Avg. 0.920 0.070 0.924 0.920 0.921 0.839 0.925 0.893

=== Confusion Matrix ===

a b ¡– classified as 38 2 — a = q 6 54 — b = c

0.6 Con J48:

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2 Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

J48 pruned tree ——————

y ¡= 0.843849 — y ¡= -0.980568: q (15.0) — y ¿ -0.980568 — — x ¡= 0.918846 — — — x ¡= -0.937157: q (4.0) — — — x ¿ -0.937157: c (55.0/3.0) — — x ¿ 0.918846: q (9.0/1.0) y ¿ 0.843849: q (17.0)

Number of Leaves : 5

Size of the tree : 9

Time taken to build model: 0.02 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.01 seconds

=== Summary ===

Correctly Classified Instances 2362 90.8112 Incorrectly Classified Instances 239 9.1888 Kappa statistic 0.8164 Mean absolute error 0.1228 Root mean squared error 0.2946 Relative absolute error 24.4981 Root relative squared error 58.6865 Total Number of Instances 2601

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.881 0.063 0.938 0.881 0.909 0.818 0.908 0.886 q 0.937 0.119 0.880 0.937 0.907 0.818 0.908 0.857 c Weighted Avg. 0.908 0.090 0.910 0.908 0.908 0.818 0.908 0.872

=== Confusion Matrix ===

a b ¡– classified as 1190 160 — a = q 79 1172 — b = c

0.7 KNN

Select lazy, then select IBk (the IB stands for Instance-Based, and the k allows us to specify the number of neighbors to examine). lazy, then select IBk (the IB stands for Instance-Based, and the k allows us to specify the number of neighbors to examine).

0.7.1 1-NN

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A ẅeka.core.EuclideanDistance -R first-last¨" Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

IB1 instance-based classifier using 1 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.01 seconds

=== Summary ===

Correctly Classified Instances 91 91 Incorrectly Classified Instances 9 9 Kappa statistic 0.8117 Mean absolute error 0.098 Root mean squared error 0.2972 Relative absolute error 19.8413 Root relative squared error 60.0468 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.875 0.067 0.897 0.875 0.886 0.812 0.904 0.835 q 0.933 0.125 0.918 0.933 0.926 0.812 0.904 0.897 c Weighted Avg. 0.910 0.102 0.910 0.910 0.910 0.812 0.904 0.872

=== Confusion Matrix ===

a b ¡– classified as 35 5 — a = q 4 56 — b = c

0.7.2 3-NN

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 3 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A ẅeka.core.EuclideanDistance -R first-last¨ Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

0.7.3 IB1 instance-based classifier

using 3 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.01 seconds

=== Summary ===

Correctly Classified Instances 88 88 Incorrectly Classified Instances 12 12 Kappa statistic 0.7436 Mean absolute error 0.1689 Root mean squared error 0.2943 Relative absolute error 34.1769 Root relative squared error 59.4657 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.775 0.050 0.912 0.775 0.838 0.750 0.950 0.885 q 0.950 0.225 0.864 0.950 0.905 0.750 0.950 0.955 c Weighted Avg. 0.880 0.155 0.883 0.880 0.878 0.750 0.950 0.927

=== Confusion Matrix ===

a b ¡– classified as 31 9 — a = q 3 57 — b = c

0.7.4 5-NN

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 5 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A ẅeka.core.EuclideanDistance -R first-last¨ Relation: circle Instances: 100 Attributes: 3 x y class Test mode: user supplied test set: size unknown (reading incrementally)

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

IB1 instance-based classifier using 5 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0 seconds

=== Summary ===

Correctly Classified Instances 87 87 Incorrectly Classified Instances 13 13 Kappa statistic 0.7257 Mean absolute error 0.2052 Root mean squared error 0.3137 Relative absolute error 41.5244 Root relative squared error 63.3867 Total Number of Instances 100

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.800 0.083 0.865 0.800 0.831 0.727 0.940 0.870 q 0.917 0.200 0.873 0.917 0.894 0.727 0.940 0.949 c Weighted Avg. 0.870 0.153 0.870 0.870 0.869 0.727 0.940 0.917

=== Confusion Matrix ===

a b ¡– classified as 32 8 — a = q 5 55 — b = c

0.8 Usando circleall come validation set

=== Summary ===

Correctly Classified Instances 2260 86.8897 Incorrectly Classified Instances 341 13.1103 Kappa statistic 0.7388 Mean absolute error 0.189 Root mean squared error 0.3052 Total Number of Instances 2601

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.806 0.063 0.932 0.806 0.865 0.746 0.956 0.943 q 0.937 0.194 0.817 0.937 0.873 0.746 0.956 0.940 c Weighted Avg. 0.869 0.126 0.877 0.869 0.869 0.746 0.956 0.942

=== Confusion Matrix ===

a b ¡– classified as 1088 262 — a = q 79 1172 — b = c

Part II

Esercizio 1.1

Random split a 66

0.9 J48 pruned tree

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

y ¡= 0.843849 — y ¡= -0.980568: q (15.0) — y ¿ -0.980568 — — x ¡= 0.918846 — — — x ¡= -0.937157: q (4.0) — — — x ¿ -0.937157: c (55.0/3.0) — — x ¿ 0.918846: q (9.0/1.0) y ¿ 0.843849: q (17.0)

Number of Leaves : 5

Size of the tree : 9

Time taken to build model: 0.02 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances 30 88.2353 Incorrectly Classified Instances 4 11.7647 Kappa statistic 0.7655 Mean absolute error 0.1544 Root mean squared error 0.3561 Relative absolute error 30.6701 Root relative squared error 70.2629 Total Number of Instances 34

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.833 0.063 0.938 0.833 0.882 0.771 0.877 0.854 q 0.938 0.167 0.833 0.938 0.882 0.771 0.877 0.811 c Weighted Avg. 0.882 0.112 0.888 0.882 0.882 0.771 0.877 0.833

=== Confusion Matrix ===

a b ¡– classified as 15 3 — a = q 1 15 — b = c

Part III

Esercizio 2.1

generare xor per backpropagation

Chapter 1

1a

1.1 Rete monostrato (cio senza strati nascosti)

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 3000 -V 0 -S 0 -E 20 -H 0 -G -B -C -I -R Relation: xor Instances: 4 Attributes: 3 x-val y-val class Test mode: evaluate on training data

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

Sigmoid Node 0 Inputs Weights Threshold -0.07208537408174595 Attrib x-val -0.0030674627268831647 Attrib y-val 0.03251510490495748 Sigmoid Node 1 Inputs Weights Threshold 0.07208537408174623 Attrib x-val 0.0030674627268826504 Attrib y-val -0.03251510490495738 Class 0 Input Node 0 Class 1 Input Node 1

Time taken to build model: 2.96 seconds

=== Predictions on training set ===

inst actual predicted error prediction 1 1:0 2:1 + 0.518 2 2:1 2:1 0.519 3 2:1 2:1 0.51 4 1:0 2:1 + 0.511

=== Evaluation on training set ===

Time taken to test model on training data: 0 seconds

=== Summary ===

Correctly Classified Instances 2 50 Incorrectly Classified Instances 2 50 Kappa statistic 0 Mean absolute error 0.5 Root mean squared error 0.5002 Relative absolute error 100 Root relative squared error 100.0445 Total Number of Instances 4

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.000 0.000 0.000 0.000 0.000 0.000 0.500 0.583 0 1.000 1.000 0.500 1.000 0.667 0.000 0.500 0.750 1 Weighted Avg. 0.500 0.500 0.250 0.500 0.333 0.000 0.500 0.667

=== Confusion Matrix ===

a b ¡– classified as 0 2 — a = 0 0 2 — b = 1

1.2 Strato nascosto con due unit (-H2)

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 3000 -V 0 -S 0 -E 20 -H 2 -G -B -C -I -R Relation: xor Instances: 4 Attributes: 3 x-val y-val class Test mode: evaluate on training data

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

Sigmoid Node 0 Inputs Weights Threshold -0.729303312647522 Node 2 0.015319018150725288 Node 3 4.910509240814651 Sigmoid Node 1 Inputs Weights Threshold 0.7297457517697482 Node 2 -0.05695966069767247 Node 3 -4.901615563545365 Sigmoid Node 2 Inputs Weights Threshold -1.8552784499595782 Attrib x-val -2.572438840390672 Attrib y-val -2.5374560121166523 Sigmoid Node 3 Inputs Weights Threshold 1.0396525478269123 Attrib x-val -6.875242413123358 Attrib y-val -6.853941969110357 Class 0 Input Node 0 Class 1 Input Node 1

Time taken to build model: 6.42 seconds

=== Predictions on training set ===

inst actual predicted error prediction 1 1:0 1:0 0.948 2 2:1 2:1 0.671 3 2:1 2:1 0.671 4 1:0 2:1 + 0.675

=== Evaluation on training set ===

Time taken to test model on training data: 0 seconds

=== Summary ===

Correctly Classified Instances 3 75 Incorrectly Classified Instances 1 25 Kappa statistic 0.5 Mean absolute error 0.346 Root mean squared error 0.4104 Relative absolute error 69.2005 Root relative squared error 82.0871 Total Number of Instances 4

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class 0.500 0.000 1.000 0.500 0.667 0.577 0.500 0.750 0 1.000 0.500 0.667 1.000 0.800 0.577 0.500 0.583 1 Weighted Avg. 0.750 0.250 0.833 0.750 0.733 0.577 0.500 0.667

=== Confusion Matrix ===

a b ¡– classified as 1 1 — a = 0 0 2 — b = 1

1.3 xorall

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 3000 -V 0 -S 0 -E 20 -H 2 -G -B -C -I -R Relation: xor Instances: 10201 Attributes: 3 x-val y-val class Test mode: evaluate on training data

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

Sigmoid Node 0 Inputs Weights Threshold 4.089387590132767 Node 2 2.721145158478633 Node 3 2.73825579705843 Sigmoid Node 1 Inputs Weights Threshold -4.056546722562926 Node 2 -2.6971960865848574 Node 3 -2.7991341332139714 Sigmoid Node 2 Inputs Weights Threshold 1.539113620853852 Attrib x-val 0.7537586384064179 Attrib y-val 0.7078101500800653 Sigmoid Node 3 Inputs Weights Threshold

1.5900863645039536 Attrib x-val 0.7307338769046867 Attrib y-val 0.7083982637469124 Class 0 Input Node 0 Class 1 Input Node 1

Time taken to build model: 18.57 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 3.71 seconds

=== Summary ===

Correctly Classified Instances 10201 100 Incorrectly Classified Instances 0 0 Kappa statistic 1 Mean absolute error 0.0001 Root mean squared error 0.0001 Relative absolute error 123.3406 Root relative squared error 124.6553 Total Number of Instances 10201

=== 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 0.000 ? 1.000 0 0.000 0.000 0.000 0.000 0.000 0.000 ? ? 1 Weighted Avg. 1.000 0.000 1.000 1.000 1.000 0.000 0.000 1.000

=== Confusion Matrix ===

a b ¡– classified as 10201 0 — a = 0 0 0 — b = 1

Chapter 2

1b

Part IV

Esercizio 3.1

Analizziamo esempio 3 di Galib