Predicting decode latency for a phone model based on os version, imgheight, imgwidth, and megapixels

Set up and used [Xerial SQLite Java database connectivity library](https://bitbucket.org/xerial/sqlite-jdbc) to drive results from RespDroid database to Weka

Prepared data for ~10000 image decode operations

Used Weka ro run four machine learning algorithms on data:

Linear Regression, KNN, Decision Tree with REP, and Neural Network

Conclusion:

KNN works the best

Linear Regression

=== Run information ===

Scheme: weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8 -num-decimal-places 4

Relation: QueryResult

Instances: 9277

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

Linear Regression Model

delay =

-33.4185 \* os +

16.0306 \* megapixels +

0.0432 \* imgsizekb +

276.4307

Time taken to build model: 0.2 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9512

**Mean absolute error 11.1885**

Root mean squared error 20.3746

Relative absolute error 23.447 %

Root relative squared error 30.8512 %

Total Number of Instances 9277

KNN Model

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

Relation: QueryResult

Instances: 9277

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

IB1 instance-based classifier

using 1 nearest neighbour(s) for classification

Time taken to build model: 0.01 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9824 => average relative error < 2%

Mean absolute error 7.5187

Root mean squared error 12.3239

Relative absolute error 15.7564 %

Root relative squared error 18.6608 %

Total Number of Instances 9277

Decision Tree with Reduced Error Pruning

=== Run information ===

Scheme: weka.classifiers.trees.REPTree -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0

Relation: QueryResult

Instances: 9277

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

REP Tree

======

Size of the tree : 1005

Time taken to build model: 0.16 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9817

**Mean absolute error 7.7437**

Root mean squared error 12.5657

Relative absolute error 16.2279 %

Root relative squared error 19.0269 %

Total Number of Instances 9277

Neural Network

=== Run information ===

Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a

Relation: QueryResult

Instances: 9277

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

Linear Node 0

Inputs Weights

Threshold 0.7125724411635623

Node 1 -1.713805992696076

Node 2 -3.626952659593815

Sigmoid Node 1

Inputs Weights

Threshold -1.774211170735627

Attrib os 0.31970514686083507

Attrib megapixels -0.8368069217332395

Attrib imgsizekb -2.6232308573288594

Sigmoid Node 2

Inputs Weights

Threshold -4.654459603406584

Attrib os -0.06112985434041348

Attrib megapixels -2.9144688943074044

Attrib imgsizekb 1.2585286450190087

Class

Input

Node 0

Time taken to build model: 1.77 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.935

**Mean absolute error 14.9888**

Root mean squared error 23.7243

Relative absolute error 31.4108 %

Root relative squared error 35.9232 %

Total Number of Instances 9277

KNN, k=10

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 10 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

Relation: QueryResult

Instances: 9277

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

IB1 instance-based classifier

using 10 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9792

**Mean absolute error 8.4783**

Root mean squared error 13.4157

Relative absolute error 17.7673 %

Root relative squared error 20.3139 %

Total Number of Instances 9277

=== Sample predictions on test data ===

inst# actual predicted error

1 401 406.294 5.294

2 76 76.556 0.556

3 161 163.444 2.444

4 143 151.444 8.444

5 187 173.882 -13.118

6 167 170.188 3.188

7 181 171.25 -9.75

8 382 365.333 -16.667

9 65 76.5 11.5

10 169 169.938 0.938

11 161 105.714 -55.286

12 60 60.875 0.875

13 120 136.222 16.222

14 106 114.125 8.125

15 42 44.5 2.5

16 123 145 22

17 135 162.571 27.571

18 201 199.833 -1.167

19 77 87.111 10.111

20 145 167.25 22.25

21 95 78.333 -16.667

22 114 81.5 -32.5

23 158 153.368 -4.632

24 51 57.25 6.25

25 139 141.333 2.333

26 34 40 6

27 85 86.75 1.75

28 122 137.5 15.5

29 69 73.125 4.125

30 54 58.5 4.5

31 113 143 30

32 55 67.75 12.75

33 72 76 4

34 48 67.333 19.333

35 127 124.222 -2.778

36 46 55.333 9.333

37 160 157.875 -2.125

38 259 242.412 -16.588

39 126 130.778 4.778

40 67 71.75 4.75

41 113 111.421 -1.579

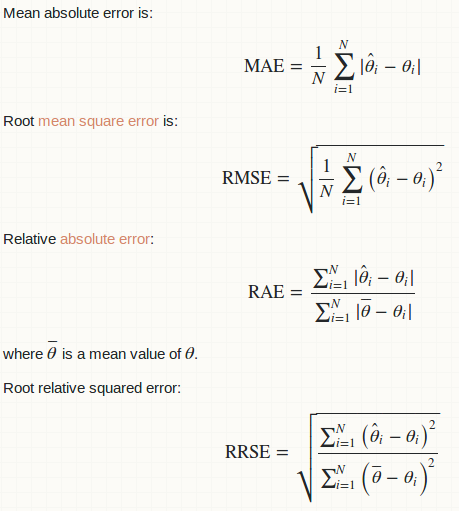
42 117 118.125 1.125

43 33 40.5 7.5

44 130 100.5 -29.5

45 96 105.778 9.778

Error Definitions:



Nexus 4: Only > 1MB instances

KNN:

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 100 -W 0 -X -I -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

Relation: QueryResult

Instances: 9469

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

IB1 instance-based classifier

using 6 inverse-distance-weighted nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9905

Mean absolute error 16.3709

Root mean squared error 21.988

Relative absolute error 14.1176 %

Root relative squared error 13.7671 %

Total Number of Instances 9469

Linear Regression:

25.2387 \* megapixels +

0.1013 \* imgsizekb +

49.3088

Nexus 6P (Only > 1MB instances)

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 10 -W 0 -X -I -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

Relation: QueryResult

Instances: 11936

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

IB1 instance-based classifier

using 1 inverse-distance-weighted nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9599

Mean absolute error 24.0608

Root mean squared error 30.0184

Relative absolute error 29.9969 %

Root relative squared error 28.0281 %

Total Number of Instances 11936

Linear Regression Model

delay =

22.1041 \* megapixels +

0.0508 \* imgsizekb +

89.4245

Support Vector Machine for Nexus 6P

=== Run information ===

Scheme: weka.classifiers.functions.SMOreg -C 1.0 -N 0 -I "weka.classifiers.functions.supportVector.RegSMOImproved -T 0.001 -V -P 1.0E-12 -L 0.001 -W 1" -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007"

Relation: QueryResult

Instances: 11936

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

SMOreg

weights (not support vectors):

+ 0 \* (normalized) os

+ 0.7335 \* (normalized) megapixels

+ 0.5122 \* (normalized) imgsizekb

+ 0.0757

Number of kernel evaluations: 1053321643 (5.288% cached)

Time taken to build model: 68.15 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9175

Mean absolute error 31.9333

Root mean squared error 42.6687

Relative absolute error 39.8117 %

Root relative squared error 39.8396 %

Total Number of Instances 11936

Prediction 4: Nexus 4, 5X, 6P

KNN

=== Run information ===

Scheme: weka.classifiers.lazy.IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""

Relation: QueryResult

Instances: 35208

Attributes: 4

os

megapixels

imgsizekb

delay

Test mode: 10-fold cross-validation

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

IB1 instance-based classifier

using 1 nearest neighbour(s) for classification

Time taken to build model: 0.01 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient 0.9568

Mean absolute error 28.3216

Root mean squared error 45.1109

Relative absolute error 24.6343 %

Root relative squared error 29.0636 %

Total Number of Instances 35208

Linear Regression Model

delay =

-72.6617 \* os +

21.0314 \* megapixels +

0.0643 \* imgsizekb +

536.4684