***Id is removed***

Weka output for INFO Gain Attribute Eval

=== Run information ===

Evaluator: weka.attributeSelection.InfoGainAttributeEval

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: merge\_data-weka.filters.unsupervised.attribute.Remove-R1

Instances: 160057

Attributes: 11

chain

offer

market

repeattrips

repeater

offerdate

category

quantity

company

offervalue

brand

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:

Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 repeater):

Information Gain Ranking Filter

Ranked attributes:

0.8435 4 repeattrips

0.0575 2 offer

0.0568 11 brand

0.0566 7 category

0.0504 9 company

0.0348 10 offervalue

0.0313 6 offerdate

0.0251 1 chain

0.0108 3 market

0 8 quantity

Selected attributes: 4,2,11,7,9,10,6,1,3,8 : 10

Weka Output gain ratio attribute eval

=== Run information ===

Evaluator: weka.attributeSelection.GainRatioAttributeEval

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: merge\_data-weka.filters.unsupervised.attribute.Remove-R1

Instances: 160057

Attributes: 11

chain

offer

market

repeattrips

repeater

offerdate

category

quantity

company

offervalue

brand

Evaluation mode: evaluate on all training data

=== Attribute Selection on all input data ===

Search Method:

Attribute ranking.

Attribute Evaluator (supervised, Class (nominal): 5 repeater):

Gain Ratio feature evaluator

Ranked attributes:

1 4 repeattrips

0.01852 7 category

0.01833 2 offer

0.01788 11 brand

0.01786 9 company

0.01568 10 offervalue

0.00602 6 offerdate

0.00524 1 chain

0.00301 3 market

0 8 quantity

Selected attributes: 4,7,2,11,9,10,6,1,3,8 : 10

CLASSIFY

J48 –c 0.25 –M 2

Cross validation folds **10**

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: merge\_data-weka.filters.unsupervised.attribute.Remove-R1

Instances: 160057

Attributes: 11

chain

offer

market

repeattrips

repeater

offerdate

category

quantity

company

offervalue

brand

Test mode: 10-fold cross-validation

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

J48 pruned tree

------------------

repeattrips <= 0: f (116619.0)

repeattrips > 0: t (43438.0)

Number of Leaves : 2

Size of the tree : 3

Time taken to build model: 0.64 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 160057 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0 %

Root relative squared error 0 %

Total Number of Instances 160057

=== 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 t

1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 f

Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000

=== Confusion Matrix ===

a b <-- classified as

43438 0 | a = t

0 116619 | b = f

J48 –c 0.25 –M 2

Cross validation folds **20**

=== Run information ===

Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2

Relation: merge\_data-weka.filters.unsupervised.attribute.Remove-R1

Instances: 160057

Attributes: 11

chain

offer

market

repeattrips

repeater

offerdate

category

quantity

company

offervalue

brand

Test mode: 20-fold cross-validation

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

J48 pruned tree

------------------

repeattrips <= 0: f (116619.0)

repeattrips > 0: t (43438.0)

Number of Leaves : 2

Size of the tree : 3

Time taken to build model: 0.45 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 160057 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0

Root mean squared error 0

Relative absolute error 0 %

Root relative squared error 0 %

Total Number of Instances 160057

=== 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 t

1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 f

Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000

=== Confusion Matrix ===

a b <-- classified as

43438 0 | a = t

0 116619 | b = f

**Random Forest**

Cross validation folds **10**

=== Run information ===

Scheme: weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Relation: merge\_data-weka.filters.unsupervised.attribute.Remove-R1

Instances: 160057

Attributes: 11

chain

offer

market

repeattrips

repeater

offerdate

category

quantity

company

offervalue

brand

Test mode: 10-fold cross-validation

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

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 22.98 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 160057 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0.0004

Root mean squared error 0.0034

Relative absolute error 0.0888 %

Root relative squared error 0.7609 %

Total Number of Instances 160057

=== 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 t

1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 f

Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000

=== Confusion Matrix ===

a b <-- classified as

43438 0 | a = t

0 116619 | b = f

**Random Forest**

Cross validation folds **20**

=== Run information ===

Scheme: weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1

Relation: merge\_data-weka.filters.unsupervised.attribute.Remove-R1

Instances: 160057

Attributes: 11

chain

offer

market

repeattrips

repeater

offerdate

category

quantity

company

offervalue

brand

Test mode: 20-fold cross-validation

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

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 20.26 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 160057 100 %

Incorrectly Classified Instances 0 0 %

Kappa statistic 1

Mean absolute error 0.0004

Root mean squared error 0.0034

Relative absolute error 0.0896 %

Root relative squared error 0.7664 %

Total Number of Instances 160057

=== 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 t

1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 f

Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000

=== Confusion Matrix ===

a b <-- classified as

43438 0 | a = t

0 116619 | b = f

Multilayer Perceptron

Cross validation folds **10**

Amaan