**1R:**

=== Run information ===

Scheme: weka.classifiers.rules.OneR -B 6

Relation: csb-ch19a

Instances: 187109

Attributes: 10

Age

Gender

Education

Smoker

Cigarettes/day

Years\_smoked

Years\_quit

Followup\_Time

Death\_codes

Freq

Test mode: 10-fold cross-validation

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

Followup\_Time:

< 0.5 -> 0

< 5.5 -> 1

>= 5.5 -> 0

(177920/187109 instances correct)

Time taken to build model: 0.43 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 177920 95.089 %

Incorrectly Classified Instances 9189 4.911 %

Kappa statistic 0.7849

Mean absolute error 0.0327

Root mean squared error 0.1809

Relative absolute error 19.3706 %

Root relative squared error 62.2433 %

Total Number of Instances 187109

=== Detailed Accuracy By Class ===

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

1.000 0.217 0.964 1.000 0.982 0.868 0.891 0.964 0

0.785 0.020 0.845 0.785 0.814 0.790 0.882 0.690 1

0.000 0.000 ? 0.000 ? ? 0.500 0.023 2

Weighted Avg. 0.951 0.188 ? 0.951 ? ? 0.881 0.910

=== Confusion Matrix ===

a b c <-- classified as

159936 36 0 | a = 0

4931 17984 0 | b = 1

965 3257 0 | c = 2

**NaiveBayes:**

=== Run information ===

Scheme: weka.classifiers.bayes.NaiveBayes

Relation: csb-ch19a

Instances: 187109

Attributes: 10

Age

Gender

Education

Smoker

Cigarettes/day

Years\_smoked

Years\_quit

Followup\_Time

Death\_codes

Freq

Test mode: 10-fold cross-validation

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

Naive Bayes Classifier

Class

Attribute 0 1 2

(0.85) (0.12) (0.02)

============================================

Age

mean 57.7062 63.7189 62.9292

std. dev. 9.5033 8.9509 8.2489

weight sum 159972 22915 4222

precision 1 1 1

Gender

0 92474.0 16105.0 2865.0

1 67500.0 6812.0 1359.0

[total] 159974.0 22917.0 4224.0

Education

0 70048.0 12169.0 2438.0

1 89926.0 10748.0 1786.0

[total] 159974.0 22917.0 4224.0

Smoker

1 207.0 981.0 298.0

2 129681.0 11959.0 1340.0

3 30087.0 9978.0 2587.0

[total] 159975.0 22918.0 4225.0

Cigarettes/day

mean 25.0533 23.9455 25.8941

std. dev. 15.4378 14.4502 14.9276

weight sum 159972 22915 4222

precision 5 5 5

Years\_smoked

mean 26.6937 34.2428 37.4514

std. dev. 11.4754 13.9715 14.2878

weight sum 159972 22915 4222

precision 1.0635 1.0635 1.0635

Years\_quit

mean 10.3881 7.9352 3.554

std. dev. 9.6092 10.3166 6.863

weight sum 159972 22915 4222

precision 1 1 1

Followup\_Time

mean 5.8986 3.765 3.8314

std. dev. 0.7725 1.6929 1.6841

weight sum 159972 22915 4222

precision 1 1 1

Freq

mean 3.4813 0.7386 0

std. dev. 99.5061 4.8214 3.4356

weight sum 159972 22915 4222

precision 20.6139 20.6139 20.6139

Time taken to build model: 0.33 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 161198 86.1519 %

Incorrectly Classified Instances 25911 13.8481 %

Kappa statistic 0.5342

Mean absolute error 0.1679

Root mean squared error 0.2733

Relative absolute error 99.346 %

Root relative squared error 94.007 %

Total Number of Instances 187109

=== Detailed Accuracy By Class ===

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

0.894 0.194 0.964 0.894 0.928 0.607 0.913 0.978 0

0.794 0.124 0.472 0.794 0.592 0.543 0.904 0.577 1

0.006 0.001 0.093 0.006 0.011 0.018 0.864 0.123 2

Weighted Avg. 0.862 0.181 0.884 0.862 0.866 0.586 0.911 0.910

=== Confusion Matrix ===

a b c <-- classified as

142973 16800 199 | a = 0

4670 18200 45 | b = 1

601 3596 25 | c = 2

**J48:**

=== Run information ===

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

Relation: csb-ch19a

Instances: 187109

Attributes: 10

Age

Gender

Education

Smoker

Cigarettes/day

Years\_smoked

Years\_quit

Followup\_Time

Death\_codes

Freq

Test mode: 10-fold cross-validation

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

J48 pruned tree

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

Followup\_Time <= 5

| Followup\_Time <= 0: 0 (2741.0/45.0)

| Followup\_Time > 0

| | Freq <= 3

| | | Cigarettes/day <= 0

| | | | Age <= 53

| | | | | Age <= 44: 1 (90.0/5.0)

| | | | | Age > 44

| | | | | | Freq <= 2

| | | | | | | Gender = 0

| | | | | | | | Education = 0: 1 (15.0/2.0)

| | | | | | | | Education = 1: 2 (10.0/2.0)

| | | | | | | Gender = 1: 2 (23.0/3.0)

| | | | | | Freq > 2: 1 (16.0)

| | | | Age > 53: 2 (221.0/19.0)

| | | Cigarettes/day > 0: 1 (20177.0/3037.0)

| | Freq > 3: 1 (725.0)

Followup\_Time > 5

| Smoker = 1

| | Freq <= 105

| | | Freq <= 4

| | | | Age <= 51

| | | | | Gender = 0: 1 (18.0/1.0)

| | | | | Gender = 1

| | | | | | Age <= 42: 1 (10.0)

| | | | | | Age > 42

| | | | | | | Freq <= 1: 2 (4.0)

| | | | | | | Freq > 1: 1 (3.0/1.0)

| | | | Age > 51: 2 (57.0)

| | | Freq > 4: 1 (139.0)

| | Freq > 105: 0 (184.0)

| Smoker = 2: 0 (129967.0/2986.0)

| Smoker = 3

| | Freq <= 1

| | | Years\_smoked <= 29

| | | | Years\_smoked <= 19: 0 (2834.0/48.0)

| | | | Years\_smoked > 19

| | | | | Cigarettes/day <= 40

| | | | | | Age <= 40: 0 (303.0/12.0)

| | | | | | Age > 40

| | | | | | | Age <= 50

| | | | | | | | Age <= 44

| | | | | | | | | Cigarettes/day <= 15: 0 (76.0/3.0)

| | | | | | | | | Cigarettes/day > 15

| | | | | | | | | | Cigarettes/day <= 20: 1 (10.0/3.0)

| | | | | | | | | | Cigarettes/day > 20

| | | | | | | | | | | Cigarettes/day <= 35: 0 (81.0/1.0)

| | | | | | | | | | | Cigarettes/day > 35

| | | | | | | | | | | | Years\_smoked <= 22: 0 (5.0)

| | | | | | | | | | | | Years\_smoked > 22

| | | | | | | | | | | | | Years\_smoked <= 27: 1 (4.0)

| | | | | | | | | | | | | Years\_smoked > 27: 0 (6.0/1.0)

| | | | | | | | Age > 44

| | | | | | | | | Cigarettes/day <= 30

| | | | | | | | | | Cigarettes/day <= 5

| | | | | | | | | | | Years\_smoked <= 20: 1 (4.0/1.0)

| | | | | | | | | | | Years\_smoked > 20: 0 (44.0/1.0)

| | | | | | | | | | Cigarettes/day > 5

| | | | | | | | | | | Years\_smoked <= 24

| | | | | | | | | | | | Years\_smoked <= 20: 1 (7.0/1.0)

| | | | | | | | | | | | Years\_smoked > 20

| | | | | | | | | | | | | Years\_smoked <= 21: 0 (22.0/1.0)

| | | | | | | | | | | | | Years\_smoked > 21

| | | | | | | | | | | | | | Cigarettes/day <= 20

| | | | | | | | | | | | | | | Cigarettes/day <= 15: 0 (25.0/5.0)

| | | | | | | | | | | | | | | Cigarettes/day > 15: 1 (7.0/1.0)

| | | | | | | | | | | | | | Cigarettes/day > 20: 0 (27.0/5.0)

| | | | | | | | | | | Years\_smoked > 24

| | | | | | | | | | | | Age <= 45: 2 (6.0/1.0)

| | | | | | | | | | | | Age > 45

| | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | Years\_smoked <= 25: 1 (5.0)

| | | | | | | | | | | | | | Years\_smoked > 25

| | | | | | | | | | | | | | | Years\_smoked <= 27

| | | | | | | | | | | | | | | | Age <= 48

| | | | | | | | | | | | | | | | | Age <= 46: 0 (2.0/1.0)

| | | | | | | | | | | | | | | | | Age > 46: 1 (4.0/1.0)

| | | | | | | | | | | | | | | | Age > 48: 0 (8.0/1.0)

| | | | | | | | | | | | | | | Years\_smoked > 27

| | | | | | | | | | | | | | | | Gender = 0

| | | | | | | | | | | | | | | | | Cigarettes/day <= 25: 0 (3.0/1.0)

| | | | | | | | | | | | | | | | | Cigarettes/day > 25: 2 (3.0/1.0)

| | | | | | | | | | | | | | | | Gender = 1: 1 (8.0/1.0)

| | | | | | | | | | | | | Education = 1: 1 (26.0/3.0)

| | | | | | | | | Cigarettes/day > 30

| | | | | | | | | | Cigarettes/day <= 35: 0 (69.0/2.0)

| | | | | | | | | | Cigarettes/day > 35

| | | | | | | | | | | Years\_smoked <= 24: 0 (11.0)

| | | | | | | | | | | Years\_smoked > 24: 1 (7.0/3.0)

| | | | | | | Age > 50

| | | | | | | | Years\_smoked <= 20

| | | | | | | | | Cigarettes/day <= 20

| | | | | | | | | | Gender = 0: 0 (38.0/6.0)

| | | | | | | | | | Gender = 1

| | | | | | | | | | | Age <= 61: 1 (10.0)

| | | | | | | | | | | Age > 61

| | | | | | | | | | | | Cigarettes/day <= 15

| | | | | | | | | | | | | Cigarettes/day <= 10

| | | | | | | | | | | | | | Age <= 65: 1 (2.0)

| | | | | | | | | | | | | | Age > 65: 0 (2.0)

| | | | | | | | | | | | | Cigarettes/day > 10: 0 (6.0)

| | | | | | | | | | | | Cigarettes/day > 15

| | | | | | | | | | | | | Age <= 70: 1 (3.0/1.0)

| | | | | | | | | | | | | Age > 70: 0 (3.0/1.0)

| | | | | | | | | Cigarettes/day > 20: 0 (78.0/7.0)

| | | | | | | | Years\_smoked > 20

| | | | | | | | | Years\_smoked <= 24: 0 (471.0/21.0)

| | | | | | | | | Years\_smoked > 24

| | | | | | | | | | Age <= 57

| | | | | | | | | | | Cigarettes/day <= 20

| | | | | | | | | | | | Cigarettes/day <= 15: 0 (92.0/9.0)

| | | | | | | | | | | | Cigarettes/day > 15

| | | | | | | | | | | | | Gender = 0

| | | | | | | | | | | | | | Age <= 56

| | | | | | | | | | | | | | | Education = 0: 0 (3.0/1.0)

| | | | | | | | | | | | | | | Education = 1

| | | | | | | | | | | | | | | | Years\_smoked <= 27: 2 (6.0/3.0)

| | | | | | | | | | | | | | | | Years\_smoked > 27: 1 (3.0)

| | | | | | | | | | | | | | Age > 56: 0 (3.0)

| | | | | | | | | | | | | Gender = 1

| | | | | | | | | | | | | | Age <= 55: 1 (8.0/1.0)

| | | | | | | | | | | | | | Age > 55: 2 (3.0/1.0)

| | | | | | | | | | | Cigarettes/day > 20: 0 (127.0/11.0)

| | | | | | | | | | Age > 57: 0 (450.0/33.0)

| | | | | Cigarettes/day > 40: 0 (673.0/8.0)

| | | Years\_smoked > 29

| | | | Cigarettes/day <= 40

| | | | | Age <= 69

| | | | | | Age <= 45

| | | | | | | Age <= 43: 0 (111.0)

| | | | | | | Age > 43

| | | | | | | | Years\_smoked <= 30

| | | | | | | | | Age <= 44

| | | | | | | | | | Gender = 0: 0 (2.0)

| | | | | | | | | | Gender = 1: 1 (5.0/1.0)

| | | | | | | | | Age > 44: 1 (3.0)

| | | | | | | | Years\_smoked > 30: 0 (88.0/4.0)

| | | | | | Age > 45

| | | | | | | Cigarettes/day <= 5

| | | | | | | | Years\_smoked <= 30

| | | | | | | | | Age <= 63

| | | | | | | | | | Education = 0

| | | | | | | | | | | Age <= 58: 1 (2.0)

| | | | | | | | | | | Age > 58: 0 (2.0)

| | | | | | | | | | Education = 1: 1 (4.0)

| | | | | | | | | Age > 63: 0 (7.0)

| | | | | | | | Years\_smoked > 30: 0 (394.0/52.0)

| | | | | | | Cigarettes/day > 5

| | | | | | | | Cigarettes/day <= 20

| | | | | | | | | Cigarettes/day <= 15

| | | | | | | | | | Years\_smoked <= 30

| | | | | | | | | | | Gender = 0

| | | | | | | | | | | | Age <= 60: 1 (9.0)

| | | | | | | | | | | | Age > 60: 0 (14.0/5.0)

| | | | | | | | | | | Gender = 1: 1 (14.0/3.0)

| | | | | | | | | | Years\_smoked > 30

| | | | | | | | | | | Age <= 48

| | | | | | | | | | | | Years\_smoked <= 33

| | | | | | | | | | | | | Years\_smoked <= 31: 1 (3.0/1.0)

| | | | | | | | | | | | | Years\_smoked > 31: 2 (3.0/1.0)

| | | | | | | | | | | | Years\_smoked > 33: 0 (38.0/1.0)

| | | | | | | | | | | Age > 48

| | | | | | | | | | | | Years\_smoked <= 34

| | | | | | | | | | | | | Age <= 55

| | | | | | | | | | | | | | Gender = 0

| | | | | | | | | | | | | | | Years\_smoked <= 31: 0 (4.0)

| | | | | | | | | | | | | | | Years\_smoked > 31

| | | | | | | | | | | | | | | | Age <= 54

| | | | | | | | | | | | | | | | | Age <= 50: 1 (2.0)

| | | | | | | | | | | | | | | | | Age > 50: 2 (4.0/1.0)

| | | | | | | | | | | | | | | | Age > 54: 0 (3.0/1.0)

| | | | | | | | | | | | | | Gender = 1: 1 (11.0/1.0)

| | | | | | | | | | | | | Age > 55: 0 (119.0/13.0)

| | | | | | | | | | | | Years\_smoked > 34

| | | | | | | | | | | | | Age <= 53

| | | | | | | | | | | | | | Years\_smoked <= 37

| | | | | | | | | | | | | | | Age <= 50

| | | | | | | | | | | | | | | | Gender = 0: 1 (6.0/2.0)

| | | | | | | | | | | | | | | | Gender = 1: 0 (3.0)

| | | | | | | | | | | | | | | Age > 50: 1 (6.0)

| | | | | | | | | | | | | | Years\_smoked > 37: 0 (56.0/2.0)

| | | | | | | | | | | | | Age > 53

| | | | | | | | | | | | | | Years\_smoked <= 50

| | | | | | | | | | | | | | | Years\_smoked <= 35

| | | | | | | | | | | | | | | | Age <= 60

| | | | | | | | | | | | | | | | | Age <= 58: 1 (10.0/1.0)

| | | | | | | | | | | | | | | | | Age > 58: 2 (4.0/1.0)

| | | | | | | | | | | | | | | | Age > 60

| | | | | | | | | | | | | | | | | Gender = 0: 0 (10.0/1.0)

| | | | | | | | | | | | | | | | | Gender = 1: 1 (6.0/1.0)

| | | | | | | | | | | | | | | Years\_smoked > 35

| | | | | | | | | | | | | | | | Age <= 58

| | | | | | | | | | | | | | | | | Years\_smoked <= 40

| | | | | | | | | | | | | | | | | | Years\_smoked <= 38: 1 (15.0/1.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 38

| | | | | | | | | | | | | | | | | | | Cigarettes/day <= 10: 1 (4.0/1.0)

| | | | | | | | | | | | | | | | | | | Cigarettes/day > 10

| | | | | | | | | | | | | | | | | | | | Gender = 0: 2 (3.0)

| | | | | | | | | | | | | | | | | | | | Gender = 1: 1 (4.0/1.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 40

| | | | | | | | | | | | | | | | | | Years\_smoked <= 44

| | | | | | | | | | | | | | | | | | | Age <= 56: 0 (20.0/1.0)

| | | | | | | | | | | | | | | | | | | Age > 56

| | | | | | | | | | | | | | | | | | | | Gender = 0: 1 (5.0)

| | | | | | | | | | | | | | | | | | | | Gender = 1: 0 (2.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 44: 0 (36.0/1.0)

| | | | | | | | | | | | | | | | Age > 58

| | | | | | | | | | | | | | | | | Years\_smoked <= 39: 0 (96.0/16.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 39

| | | | | | | | | | | | | | | | | | Years\_smoked <= 40: 1 (25.0/3.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 40

| | | | | | | | | | | | | | | | | | | Age <= 60

| | | | | | | | | | | | | | | | | | | | Years\_smoked <= 45: 1 (10.0)

| | | | | | | | | | | | | | | | | | | | Years\_smoked > 45: 0 (26.0/2.0)

| | | | | | | | | | | | | | | | | | | Age > 60

| | | | | | | | | | | | | | | | | | | | Years\_smoked <= 42

| | | | | | | | | | | | | | | | | | | | | Age <= 62: 2 (6.0/3.0)

| | | | | | | | | | | | | | | | | | | | | Age > 62: 0 (32.0/8.0)

| | | | | | | | | | | | | | | | | | | | Years\_smoked > 42

| | | | | | | | | | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | | | | | | | | | Age <= 65: 1 (40.0/11.0)

| | | | | | | | | | | | | | | | | | | | | | Age > 65

| | | | | | | | | | | | | | | | | | | | | | | Years\_smoked <= 45: 0 (9.0/2.0)

| | | | | | | | | | | | | | | | | | | | | | | Years\_smoked > 45: 1 (20.0/9.0)

| | | | | | | | | | | | | | | | | | | | | Education = 1

| | | | | | | | | | | | | | | | | | | | | | Age <= 64

| | | | | | | | | | | | | | | | | | | | | | | Years\_smoked <= 47: 1 (12.0)

| | | | | | | | | | | | | | | | | | | | | | | Years\_smoked > 47

| | | | | | | | | | | | | | | | | | | | | | | | Age <= 62: 0 (7.0/1.0)

| | | | | | | | | | | | | | | | | | | | | | | | Age > 62: 1 (5.0/2.0)

| | | | | | | | | | | | | | | | | | | | | | Age > 64

| | | | | | | | | | | | | | | | | | | | | | | Years\_smoked <= 44: 0 (6.0)

| | | | | | | | | | | | | | | | | | | | | | | Years\_smoked > 44: 1 (25.0/9.0)

| | | | | | | | | | | | | | Years\_smoked > 50

| | | | | | | | | | | | | | | Age <= 66: 0 (74.0/10.0)

| | | | | | | | | | | | | | | Age > 66

| | | | | | | | | | | | | | | | Years\_smoked <= 55

| | | | | | | | | | | | | | | | | Age <= 68

| | | | | | | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | | | | | | Years\_smoked <= 54: 1 (9.0/2.0)

| | | | | | | | | | | | | | | | | | | Years\_smoked > 54: 0 (2.0)

| | | | | | | | | | | | | | | | | | Education = 1: 0 (16.0/5.0)

| | | | | | | | | | | | | | | | | Age > 68: 1 (9.0/1.0)

| | | | | | | | | | | | | | | | Years\_smoked > 55: 0 (33.0/4.0)

| | | | | | | | | Cigarettes/day > 15

| | | | | | | | | | Years\_smoked <= 51

| | | | | | | | | | | Age <= 68

| | | | | | | | | | | | Years\_smoked <= 30: 1 (42.0/8.0)

| | | | | | | | | | | | Years\_smoked > 30

| | | | | | | | | | | | | Age <= 48

| | | | | | | | | | | | | | Years\_smoked <= 34: 1 (8.0)

| | | | | | | | | | | | | | Years\_smoked > 34: 0 (10.0/2.0)

| | | | | | | | | | | | | Age > 48

| | | | | | | | | | | | | | Years\_smoked <= 33

| | | | | | | | | | | | | | | Age <= 58: 1 (35.0/7.0)

| | | | | | | | | | | | | | | Age > 58

| | | | | | | | | | | | | | | | Gender = 0: 0 (16.0)

| | | | | | | | | | | | | | | | Gender = 1

| | | | | | | | | | | | | | | | | Age <= 62: 1 (3.0)

| | | | | | | | | | | | | | | | | Age > 62: 0 (10.0/2.0)

| | | | | | | | | | | | | | Years\_smoked > 33

| | | | | | | | | | | | | | | Age <= 54

| | | | | | | | | | | | | | | | Years\_smoked <= 40: 1 (50.0/19.0)

| | | | | | | | | | | | | | | | Years\_smoked > 40: 0 (25.0/3.0)

| | | | | | | | | | | | | | | Age > 54

| | | | | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | | | | Years\_smoked <= 36

| | | | | | | | | | | | | | | | | | Age <= 62

| | | | | | | | | | | | | | | | | | | Age <= 55: 2 (3.0/1.0)

| | | | | | | | | | | | | | | | | | | Age > 55: 1 (7.0)

| | | | | | | | | | | | | | | | | | Age > 62: 0 (13.0/4.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 36: 1 (144.0/53.0)

| | | | | | | | | | | | | | | | Education = 1

| | | | | | | | | | | | | | | | | Years\_smoked <= 46

| | | | | | | | | | | | | | | | | | Age <= 64

| | | | | | | | | | | | | | | | | | | Age <= 55

| | | | | | | | | | | | | | | | | | | | Years\_smoked <= 42

| | | | | | | | | | | | | | | | | | | | | Years\_smoked <= 34: 2 (3.0/1.0)

| | | | | | | | | | | | | | | | | | | | | Years\_smoked > 34: 1 (4.0)

| | | | | | | | | | | | | | | | | | | | Years\_smoked > 42: 0 (2.0)

| | | | | | | | | | | | | | | | | | | Age > 55

| | | | | | | | | | | | | | | | | | | | Gender = 0

| | | | | | | | | | | | | | | | | | | | | Years\_smoked <= 44: 1 (41.0/11.0)

| | | | | | | | | | | | | | | | | | | | | Years\_smoked > 44: 2 (6.0/1.0)

| | | | | | | | | | | | | | | | | | | | Gender = 1: 1 (32.0/7.0)

| | | | | | | | | | | | | | | | | | Age > 64

| | | | | | | | | | | | | | | | | | | Years\_smoked <= 41

| | | | | | | | | | | | | | | | | | | | Age <= 66: 1 (8.0/3.0)

| | | | | | | | | | | | | | | | | | | | Age > 66: 0 (8.0/2.0)

| | | | | | | | | | | | | | | | | | | Years\_smoked > 41: 1 (12.0/3.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 46

| | | | | | | | | | | | | | | | | | Age <= 63

| | | | | | | | | | | | | | | | | | | Age <= 60: 0 (10.0)

| | | | | | | | | | | | | | | | | | | Age > 60

| | | | | | | | | | | | | | | | | | | | Years\_smoked <= 48: 1 (4.0/1.0)

| | | | | | | | | | | | | | | | | | | | Years\_smoked > 48: 0 (4.0)

| | | | | | | | | | | | | | | | | | Age > 63: 1 (24.0/6.0)

| | | | | | | | | | | Age > 68

| | | | | | | | | | | | Years\_smoked <= 47: 0 (24.0/6.0)

| | | | | | | | | | | | Years\_smoked > 47: 1 (7.0/2.0)

| | | | | | | | | | Years\_smoked > 51

| | | | | | | | | | | Gender = 0

| | | | | | | | | | | | Education = 0: 2 (28.0/15.0)

| | | | | | | | | | | | Education = 1

| | | | | | | | | | | | | Years\_smoked <= 56

| | | | | | | | | | | | | | Age <= 66

| | | | | | | | | | | | | | | Years\_smoked <= 52: 2 (3.0/1.0)

| | | | | | | | | | | | | | | Years\_smoked > 52: 0 (9.0/1.0)

| | | | | | | | | | | | | | Age > 66: 1 (8.0/2.0)

| | | | | | | | | | | | | Years\_smoked > 56: 0 (9.0/1.0)

| | | | | | | | | | | Gender = 1

| | | | | | | | | | | | Age <= 66: 0 (18.0/2.0)

| | | | | | | | | | | | Age > 66

| | | | | | | | | | | | | Years\_smoked <= 54: 1 (6.0/3.0)

| | | | | | | | | | | | | Years\_smoked > 54: 0 (11.0/4.0)

| | | | | | | | Cigarettes/day > 20

| | | | | | | | | Cigarettes/day <= 35

| | | | | | | | | | Cigarettes/day <= 30

| | | | | | | | | | | Cigarettes/day <= 25

| | | | | | | | | | | | Age <= 63

| | | | | | | | | | | | | Gender = 0

| | | | | | | | | | | | | | Years\_smoked <= 38

| | | | | | | | | | | | | | | Age <= 56

| | | | | | | | | | | | | | | | Age <= 49

| | | | | | | | | | | | | | | | | Years\_smoked <= 33: 1 (3.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 33: 0 (11.0)

| | | | | | | | | | | | | | | | Age > 49

| | | | | | | | | | | | | | | | | Years\_smoked <= 33

| | | | | | | | | | | | | | | | | | Education = 0: 0 (6.0)

| | | | | | | | | | | | | | | | | | Education = 1: 1 (2.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 33

| | | | | | | | | | | | | | | | | | Age <= 51: 0 (2.0/1.0)

| | | | | | | | | | | | | | | | | | Age > 51

| | | | | | | | | | | | | | | | | | | Age <= 53: 1 (3.0)

| | | | | | | | | | | | | | | | | | | Age > 53: 2 (3.0/1.0)

| | | | | | | | | | | | | | | Age > 56: 0 (39.0/4.0)

| | | | | | | | | | | | | | Years\_smoked > 38

| | | | | | | | | | | | | | | Age <= 53: 0 (14.0/1.0)

| | | | | | | | | | | | | | | Age > 53

| | | | | | | | | | | | | | | | Years\_smoked <= 45

| | | | | | | | | | | | | | | | | Age <= 56

| | | | | | | | | | | | | | | | | | Years\_smoked <= 40: 1 (2.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 40: 0 (10.0/2.0)

| | | | | | | | | | | | | | | | | Age > 56: 1 (32.0/11.0)

| | | | | | | | | | | | | | | | Years\_smoked > 45

| | | | | | | | | | | | | | | | | Age <= 61: 0 (19.0/3.0)

| | | | | | | | | | | | | | | | | Age > 61

| | | | | | | | | | | | | | | | | | Years\_smoked <= 51

| | | | | | | | | | | | | | | | | | | Age <= 62: 2 (5.0/2.0)

| | | | | | | | | | | | | | | | | | | Age > 62: 1 (3.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 51: 0 (2.0)

| | | | | | | | | | | | | Gender = 1: 0 (114.0/24.0)

| | | | | | | | | | | | Age > 63: 0 (154.0/18.0)

| | | | | | | | | | | Cigarettes/day > 25

| | | | | | | | | | | | Age <= 63

| | | | | | | | | | | | | Years\_smoked <= 50

| | | | | | | | | | | | | | Years\_smoked <= 30

| | | | | | | | | | | | | | | Gender = 0: 1 (13.0/3.0)

| | | | | | | | | | | | | | | Gender = 1

| | | | | | | | | | | | | | | | Education = 0: 1 (3.0/1.0)

| | | | | | | | | | | | | | | | Education = 1: 2 (4.0/1.0)

| | | | | | | | | | | | | | Years\_smoked > 30

| | | | | | | | | | | | | | | Gender = 0

| | | | | | | | | | | | | | | | Years\_smoked <= 32

| | | | | | | | | | | | | | | | | Age <= 53: 1 (6.0)

| | | | | | | | | | | | | | | | | Age > 53: 0 (12.0/3.0)

| | | | | | | | | | | | | | | | Years\_smoked > 32: 1 (123.0/62.0)

| | | | | | | | | | | | | | | Gender = 1

| | | | | | | | | | | | | | | | Years\_smoked <= 44

| | | | | | | | | | | | | | | | | Age <= 56

| | | | | | | | | | | | | | | | | | Years\_smoked <= 36: 1 (22.0/10.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 36: 0 (24.0/5.0)

| | | | | | | | | | | | | | | | | Age > 56

| | | | | | | | | | | | | | | | | | Years\_smoked <= 35: 0 (15.0/2.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 35: 1 (34.0/12.0)

| | | | | | | | | | | | | | | | Years\_smoked > 44

| | | | | | | | | | | | | | | | | Age <= 60: 0 (13.0)

| | | | | | | | | | | | | | | | | Age > 60

| | | | | | | | | | | | | | | | | | Years\_smoked <= 47: 2 (5.0/1.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 47: 0 (4.0)

| | | | | | | | | | | | | Years\_smoked > 50: 0 (22.0/2.0)

| | | | | | | | | | | | Age > 63

| | | | | | | | | | | | | Years\_smoked <= 43: 0 (98.0/11.0)

| | | | | | | | | | | | | Years\_smoked > 43

| | | | | | | | | | | | | | Years\_smoked <= 52

| | | | | | | | | | | | | | | Years\_smoked <= 48

| | | | | | | | | | | | | | | | Age <= 65

| | | | | | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | | | | | Age <= 64: 2 (4.0/1.0)

| | | | | | | | | | | | | | | | | | Age > 64: 1 (3.0)

| | | | | | | | | | | | | | | | | Education = 1: 1 (8.0/1.0)

| | | | | | | | | | | | | | | | Age > 65: 0 (29.0/8.0)

| | | | | | | | | | | | | | | Years\_smoked > 48

| | | | | | | | | | | | | | | | Age <= 66

| | | | | | | | | | | | | | | | | Years\_smoked <= 51: 1 (15.0/4.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 51: 0 (3.0)

| | | | | | | | | | | | | | | | Age > 66

| | | | | | | | | | | | | | | | | Gender = 0: 1 (10.0/4.0)

| | | | | | | | | | | | | | | | | Gender = 1

| | | | | | | | | | | | | | | | | | Education = 0: 1 (2.0)

| | | | | | | | | | | | | | | | | | Education = 1: 2 (2.0)

| | | | | | | | | | | | | | Years\_smoked > 52: 0 (48.0/12.0)

| | | | | | | | | | Cigarettes/day > 30: 0 (383.0/29.0)

| | | | | | | | | Cigarettes/day > 35

| | | | | | | | | | Age <= 61

| | | | | | | | | | | Gender = 0

| | | | | | | | | | | | Years\_smoked <= 48

| | | | | | | | | | | | | Age <= 47

| | | | | | | | | | | | | | Years\_smoked <= 35: 2 (8.0/3.0)

| | | | | | | | | | | | | | Years\_smoked > 35: 0 (5.0/1.0)

| | | | | | | | | | | | | Age > 47

| | | | | | | | | | | | | | Age <= 58

| | | | | | | | | | | | | | | Years\_smoked <= 40: 1 (73.0/24.0)

| | | | | | | | | | | | | | | Years\_smoked > 40

| | | | | | | | | | | | | | | | Age <= 55: 0 (18.0/4.0)

| | | | | | | | | | | | | | | | Age > 55: 1 (12.0/6.0)

| | | | | | | | | | | | | | Age > 58

| | | | | | | | | | | | | | | Years\_smoked <= 33: 0 (7.0/1.0)

| | | | | | | | | | | | | | | Years\_smoked > 33: 1 (39.0/19.0)

| | | | | | | | | | | | Years\_smoked > 48

| | | | | | | | | | | | | Years\_smoked <= 50

| | | | | | | | | | | | | | Age <= 60: 0 (7.0/2.0)

| | | | | | | | | | | | | | Age > 60: 1 (3.0)

| | | | | | | | | | | | | Years\_smoked > 50: 0 (4.0)

| | | | | | | | | | | Gender = 1

| | | | | | | | | | | | Years\_smoked <= 41

| | | | | | | | | | | | | Age <= 60

| | | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | | Age <= 58

| | | | | | | | | | | | | | | | Age <= 52

| | | | | | | | | | | | | | | | | Years\_smoked <= 36: 1 (8.0/4.0)

| | | | | | | | | | | | | | | | | Years\_smoked > 36: 0 (7.0/1.0)

| | | | | | | | | | | | | | | | Age > 52

| | | | | | | | | | | | | | | | | Age <= 54: 1 (5.0/1.0)

| | | | | | | | | | | | | | | | | Age > 54: 2 (9.0/4.0)

| | | | | | | | | | | | | | | Age > 58

| | | | | | | | | | | | | | | | Years\_smoked <= 34: 0 (4.0/1.0)

| | | | | | | | | | | | | | | | Years\_smoked > 34: 1 (4.0)

| | | | | | | | | | | | | | Education = 1

| | | | | | | | | | | | | | | Age <= 48: 0 (5.0)

| | | | | | | | | | | | | | | Age > 48

| | | | | | | | | | | | | | | | Years\_smoked <= 32

| | | | | | | | | | | | | | | | | Age <= 54: 2 (2.0)

| | | | | | | | | | | | | | | | | Age > 54: 0 (5.0)

| | | | | | | | | | | | | | | | Years\_smoked > 32

| | | | | | | | | | | | | | | | | Age <= 56

| | | | | | | | | | | | | | | | | | Years\_smoked <= 38: 1 (11.0/2.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 38: 0 (4.0/1.0)

| | | | | | | | | | | | | | | | | Age > 56

| | | | | | | | | | | | | | | | | | Years\_smoked <= 34: 0 (2.0/1.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 34

| | | | | | | | | | | | | | | | | | | Years\_smoked <= 36: 2 (2.0)

| | | | | | | | | | | | | | | | | | | Years\_smoked > 36: 1 (7.0/3.0)

| | | | | | | | | | | | | Age > 60: 0 (5.0)

| | | | | | | | | | | | Years\_smoked > 41: 0 (30.0/4.0)

| | | | | | | | | | Age > 61

| | | | | | | | | | | Years\_smoked <= 39: 0 (82.0/4.0)

| | | | | | | | | | | Years\_smoked > 39

| | | | | | | | | | | | Gender = 0

| | | | | | | | | | | | | Years\_smoked <= 52

| | | | | | | | | | | | | | Age <= 65: 1 (37.0/14.0)

| | | | | | | | | | | | | | Age > 65

| | | | | | | | | | | | | | | Years\_smoked <= 44: 0 (24.0/4.0)

| | | | | | | | | | | | | | | Years\_smoked > 44

| | | | | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | | | | Age <= 67: 2 (5.0/1.0)

| | | | | | | | | | | | | | | | | Age > 67

| | | | | | | | | | | | | | | | | | Years\_smoked <= 49: 0 (7.0/3.0)

| | | | | | | | | | | | | | | | | | Years\_smoked > 49: 1 (2.0/1.0)

| | | | | | | | | | | | | | | | Education = 1: 1 (6.0/1.0)

| | | | | | | | | | | | | Years\_smoked > 52

| | | | | | | | | | | | | | Years\_smoked <= 56

| | | | | | | | | | | | | | | Education = 0

| | | | | | | | | | | | | | | | Age <= 64: 0 (4.0/1.0)

| | | | | | | | | | | | | | | | Age > 64: 1 (6.0)

| | | | | | | | | | | | | | | Education = 1

| | | | | | | | | | | | | | | | Years\_smoked <= 53: 1 (3.0/1.0)

| | | | | | | | | | | | | | | | Years\_smoked > 53: 0 (6.0/1.0)

| | | | | | | | | | | | | | Years\_smoked > 56: 0 (13.0/2.0)

| | | | | | | | | | | | Gender = 1: 0 (71.0/25.0)

| | | | | Age > 69

| | | | | | Years\_smoked <= 49: 0 (880.0/113.0)

| | | | | | Years\_smoked > 49

| | | | | | | Years\_smoked <= 50

| | | | | | | | Age <= 73

| | | | | | | | | Cigarettes/day <= 20: 1 (23.0/8.0)

| | | | | | | | | Cigarettes/day > 20

| | | | | | | | | | Cigarettes/day <= 25: 0 (5.0/2.0)

| | | | | | | | | | Cigarettes/day > 25

| | | | | | | | | | | Education = 0: 1 (7.0/2.0)

| | | | | | | | | | | Education = 1: 0 (4.0/1.0)

| | | | | | | | Age > 73: 0 (65.0/19.0)

| | | | | | | Years\_smoked > 50: 0 (1074.0/247.0)

| | | | Cigarettes/day > 40: 0 (1597.0/123.0)

| | Freq > 1

| | | Freq <= 3

| | | | Years\_smoked <= 29: 0 (3622.0/9.0)

| | | | Years\_smoked > 29

| | | | | Cigarettes/day <= 15: 0 (1544.0/47.0)

| | | | | Cigarettes/day > 15

| | | | | | Cigarettes/day <= 20

| | | | | | | Age <= 71

| | | | | | | | Freq <= 2

| | | | | | | | | Education = 0

| | | | | | | | | | Age <= 46: 0 (9.0)

| | | | | | | | | | Age > 46

| | | | | | | | | | | Gender = 0

| | | | | | | | | | | | Years\_smoked <= 49

| | | | | | | | | | | | | Age <= 62

| | | | | | | | | | | | | | Years\_smoked <= 42: 1 (19.0/3.0)

| | | | | | | | | | | | | | Years\_smoked > 42

| | | | | | | | | | | | | | | Age <= 58: 0 (5.0)

| | | | | | | | | | | | | | | Age > 58: 1 (5.0/1.0)

| | | | | | | | | | | | | Age > 62

| | | | | | | | | | | | | | Years\_smoked <= 44: 0 (14.0/4.0)

| | | | | | | | | | | | | | Years\_smoked > 44: 1 (7.0)

| | | | | | | | | | | | Years\_smoked > 49

| | | | | | | | | | | | | Years\_smoked <= 50: 2 (5.0/1.0)

| | | | | | | | | | | | | Years\_smoked > 50

| | | | | | | | | | | | | | Age <= 67: 0 (8.0/1.0)

| | | | | | | | | | | | | | Age > 67: 1 (5.0/1.0)

| | | | | | | | | | | Gender = 1

| | | | | | | | | | | | Age <= 66: 1 (33.0/14.0)

| | | | | | | | | | | | Age > 66: 0 (25.0/7.0)

| | | | | | | | | Education = 1: 0 (131.0/33.0)

| | | | | | | | Freq > 2: 0 (167.0/32.0)

| | | | | | | Age > 71: 0 (177.0/11.0)

| | | | | | Cigarettes/day > 20: 0 (2182.0/69.0)

| | | Freq > 3: 0 (11486.0/16.0)

Number of Leaves : 266

Size of the tree : 530

Time taken to build model: 11.04 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 178825 95.5726 %

Incorrectly Classified Instances 8284 4.4274 %

Kappa statistic 0.8118

Mean absolute error 0.0514

Root mean squared error 0.1621

Relative absolute error 30.4104 %

Root relative squared error 55.7768 %

Total Number of Instances 187109

=== Detailed Accuracy By Class ===

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

0.998 0.166 0.973 0.998 0.985 0.892 0.937 0.981 0

0.825 0.022 0.838 0.825 0.832 0.809 0.922 0.775 1

0.072 0.001 0.706 0.072 0.130 0.220 0.893 0.229 2

Weighted Avg. 0.956 0.145 0.950 0.956 0.947 0.867 0.934 0.938

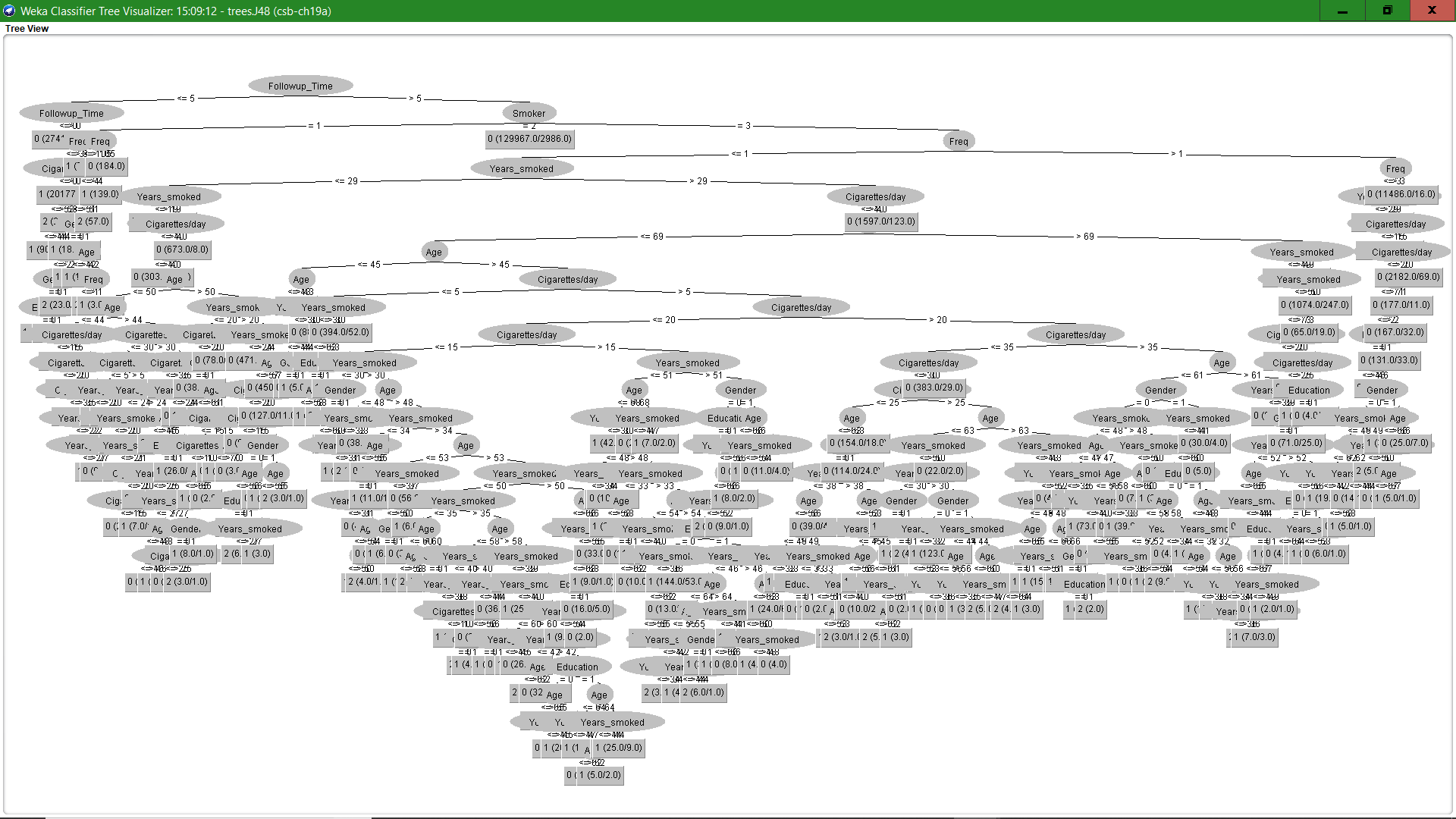
=== Confusion Matrix ===

a b c <-- classified as

159612 317 43 | a = 0

3921 18911 83 | b = 1

591 3329 302 | c = 2



OneR without follow up time:

=== Run information ===

Scheme: weka.classifiers.rules.OneR -B 6

Relation: csb-ch19a-weka.filters.unsupervised.attribute.Remove-R8

Instances: 187109

Attributes: 9

Age

Gender

Education

Smoker

Cigarettes/day

Years\_smoked

Years\_quit

Death\_codes

Freq

Test mode: 10-fold cross-validation

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

Years\_smoked:

< 2.5 -> 1

< 64.5 -> 0

< 65.5 -> 1

>= 65.5 -> 0

(160752/187109 instances correct)

Time taken to build model: 0.33 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances 160748 85.9114 %

Incorrectly Classified Instances 26361 14.0886 %

Kappa statistic 0.0678

Mean absolute error 0.0939

Root mean squared error 0.3065

Relative absolute error 55.5694 %

Root relative squared error 105.4239 %

Total Number of Instances 187109

=== Detailed Accuracy By Class ===

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

0.999 0.951 0.861 0.999 0.925 0.183 0.524 0.861 0

0.044 0.003 0.651 0.044 0.083 0.148 0.520 0.146 1

0.000 0.000 ? 0.000 ? ? 0.500 0.023 2

Weighted Avg. 0.859 0.814 ? 0.859 ? ? 0.523 0.754

=== Confusion Matrix ===

a b c <-- classified as

159733 239 0 | a = 0

21900 1015 0 | b = 1

3918 304 0 | c = 2