# **WEKA Results of Data Set G1-Data-Supervised - multiclass.csv**

This file contains 823,138 G1 instructions with multiclass classification, such that: -

* 185,560 Instructions with X,Y,Z parameters greater than threshold are labelled 'High'
* 123,728 Instructions with F,E,S parameters greater than threshold are labelled 'Medium'
* Rest of the instructions (513,850) have been labelled 'Low'

|  |  |
| --- | --- |
| ZeroR | === Run information ===  Scheme: weka.classifiers.rules.ZeroR  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  ZeroR predicts class value: Low  Time taken to build model: 0.06 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.35 seconds  === Summary ===  Correctly Classified Instances 174551 62.3693 %  Incorrectly Classified Instances 105316 37.6307 %  Kappa statistic 0  Mean absolute error 0.358  Root mean squared error 0.4232  Relative absolute error 100 %  Root relative squared error 100 %  Total Number of Instances 279867  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  1.000 1.000 0.624 1.000 0.768 0.000 0.500 0.624 Low  0.000 0.000 0.000 0.000 0.000 0.000 0.500 0.226 High  0.000 0.000 0.000 0.000 0.000 0.000 0.500 0.150 Medium  Weighted Avg. 0.624 0.624 0.389 0.624 0.479 0.000 0.500 0.463  === Confusion Matrix ===  a b c <-- classified as  174551 0 0 | a = Low  63357 0 0 | b = High  41959 0 0 | c = Medium |
| Naïve Bayes | === Run information ===  Scheme: weka.classifiers.bayes.NaiveBayes  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Naive Bayes Classifier  Class  Attribute Low High Medium  (0.62) (0.23) (0.15)  =============================================  G1  mean 1 1 1  std. dev. 0.0017 0.0017 0.0017  weight sum 513850 185560 123728  precision 0.01 0.01 0.01  Z  mean 0.0415 47.7196 0.0461  std. dev. 0.9104 66.9725 0.9475  weight sum 513850 185560 123728  precision 1.4472 1.4472 1.4472  F  mean 695.7429 694.1924 7426.1174  std. dev. 1534.1751 1538.837 7157.3368  weight sum 513850 185560 123728  precision 1.9065 1.9065 1.9065  E  mean 23.839 25.3612 813.3917  std. dev. 81.7663 90.4025 782.188  weight sum 513850 185560 123728  precision 0.0158 0.0158 0.0158  X  mean 96.3318 186.6329 96.3384  std. dev. 33.2417 135.3781 33.2182  weight sum 513850 185560 123728  precision 0.0088 0.0088 0.0088  Y  mean 95.7298 190.5919 95.7313  std. dev. 28.0155 135.939 28.0076  weight sum 513850 185560 123728  precision 0.0097 0.0097 0.0097  Time taken to build model: 0.79 seconds  === Evaluation on test split ===  Time taken to test model on training split: 1.74 seconds  === Summary ===  Correctly Classified Instances 275248 98.3496 %  Incorrectly Classified Instances 4619 1.6504 %  Kappa statistic 0.9697  Mean absolute error 0.0222  Root mean squared error 0.1142  Relative absolute error 6.2088 %  Root relative squared error 26.9795 %  Total Number of Instances 279867  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.976 0.001 0.999 0.976 0.988 0.968 0.994 0.997 Low  0.993 0.003 0.990 0.993 0.992 0.989 0.998 0.992 High  0.998 0.016 0.916 0.998 0.955 0.948 0.996 0.977 Medium  Weighted Avg. 0.983 0.004 0.985 0.983 0.984 0.970 0.995 0.993  === Confusion Matrix ===  a b c <-- classified as  170438 542 3571 | a = Low  144 62945 268 | b = High  0 94 41865 | c = Medium |
| Logistic Regression | === Run information ===  Scheme: weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Logistic Regression with ridge parameter of 1.0E-8  Coefficients...  Class  Variable Low High  ===============================  Z 4.3701 5.4708  F -0.0252 -0.024  E -0.1339 -0.1353  X 0.0948 0.1675  Y 0.0898 0.1699  Intercept 116.9703 91.4987  Odds Ratios...  Class  Variable Low High  ===============================  Z 79.0544 237.6446  F 0.9751 0.9763  E 0.8747 0.8735  X 1.0994 1.1823  Y 1.0939 1.1852  Time taken to build model: 53.27 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.64 seconds  === Summary ===  Correctly Classified Instances 277743 99.2411 %  Incorrectly Classified Instances 2124 0.7589 %  Kappa statistic 0.9858  Mean absolute error 0.0053  Root mean squared error 0.0683  Relative absolute error 1.4891 %  Root relative squared error 16.1374 %  Total Number of Instances 279867  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.996 0.013 0.992 0.996 0.994 0.985 0.999 0.999 Low  0.979 0.003 0.991 0.979 0.985 0.981 0.998 0.997 High  0.996 0.001 0.995 0.996 0.995 0.995 1.000 0.998 Medium  Weighted Avg. 0.992 0.009 0.992 0.992 0.992 0.985 0.999 0.998  === Confusion Matrix ===  a b c <-- classified as  173908 566 77 | a = Low  1178 62053 126 | b = High  177 0 41782 | c = Medium |
| MLP | === Run information ===  Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Sigmoid Node 0  Inputs Weights  Threshold -41.37746787531669  Node 3 15.968093353710374  Node 4 13.719523499395214  Node 5 20.313152614742933  Node 6 -15.7365430161216  Sigmoid Node 1  Inputs Weights  Threshold 24.01899888331354  Node 3 1.300612518327966  Node 4 -14.551778823304229  Node 5 -19.506146789209183  Node 6 16.261354920370096  Sigmoid Node 2  Inputs Weights  Threshold -14.996010099489586  Node 3 -17.82006324684913  Node 4 12.24188980156112  Node 5 10.70387016716158  Node 6 -13.076051394886326  Sigmoid Node 3  Inputs Weights  Threshold -53.99843471495498  Attrib G1 -0.015916173868014492  Attrib Z 56.461795306154265  Attrib F -75.28847495549718  Attrib E -92.05252224306088  Attrib X -0.13013780501797706  Attrib Y -1.6928962229324946  Sigmoid Node 4  Inputs Weights  Threshold -44.50977865954518  Attrib G1 0.016320232639976534  Attrib Z 24.93342887955914  Attrib F 0.0207484958089401  Attrib E -0.4861250830574653  Attrib X -124.11269426755591  Attrib Y -0.012030461849520733  Sigmoid Node 5  Inputs Weights  Threshold -58.85908675392389  Attrib G1 -0.047391629774234646  Attrib Z -67.62527520010809  Attrib F -0.6113556939094151  Attrib E 5.505213642189427  Attrib X -13.212210935306421  Attrib Y 9.27057376904345  Sigmoid Node 6  Inputs Weights  Threshold 103.25408089835771  Attrib G1 -0.029431187533203187  Attrib Z 42.77386022137825  Attrib F 0.5018643984960456  Attrib E 0.022302159128000154  Attrib X -2.370582581935577  Attrib Y 114.4191780277348  Class Low  Input  Node 0  Class High  Input  Node 1  Class Medium  Input  Node 2  Time taken to build model: 865.1 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.82 seconds  === Summary ===  Correctly Classified Instances 279792 99.9732 %  Incorrectly Classified Instances 75 0.0268 %  Kappa statistic 0.9995  Mean absolute error 0.0007  Root mean squared error 0.0126  Relative absolute error 0.1972 %  Root relative squared error 2.9761 %  Total Number of Instances 279867  === 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 Low  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 High  1.000 0.000 0.999 1.000 0.999 0.999 1.000 1.000 Medium  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000  === Confusion Matrix ===  a b c <-- classified as  174487 17 47 | a = Low  0 63350 7 | b = High  0 4 41955 | c = Medium |
| AdaBoostM1 | === Run information ===  Scheme: weka.classifiers.meta.AdaBoostM1 -P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  AdaBoostM1: Base classifiers and their weights:  Decision Stump  Classifications  E <= 1000.78522 : Low  E > 1000.78522 : Medium  E is missing : Low  Class distributions  E <= 1000.78522  Low High Medium  0.6780651818774429 0.2446418533209821 0.07729296480157505  E > 1000.78522  Low High Medium  0.0 0.0025413349663196572 0.9974586650336803  E is missing  Low High Medium  0.6242574149170613 0.22543000080180967 0.15031258428112904  Weight: 0.86  Decision Stump  Classifications  F <= 6900.5 : Low  F > 6900.5 : Medium  F is missing : Low  Class distributions  F <= 6900.5  Low High Medium  0.5058172530306365 0.43312242935696904 0.06106031761239459  F > 6900.5  Low High Medium  1.3503484697484826E-16 6.007287017759574E-4 0.9993992712982239  F is missing  Low High Medium  0.44373613999502565 0.38003719268730607 0.17622666731766837  Weight: 0.27  Decision Stump  Classifications  F <= 6900.5 : High  F > 6900.5 : Medium  F is missing : High  Class distributions  F <= 6900.5  Low High Medium  0.4393268304601013 0.49139741675648496 0.0692757527834138  F > 6900.5  Low High Medium  1.6080703372883978E-17 7.845607833434522E-4 0.9992154392166566  F is missing  Low High Medium  0.39171851900259397 0.43823141815866423 0.1700500628387418  Weight: 0.19  Decision Stump  Classifications  F <= 6900.5 : Low  F > 6900.5 : Medium  F is missing : Low  Class distributions  F <= 6900.5  Low High Medium  0.4793550393238587 0.44505731408231386 0.07558764659382744  F > 6900.5  Low High Medium  4.4799421102272655E-16 9.4502509966945E-4 0.9990549749003301  F is missing  Low High Medium  0.43181503106401087 0.4010125035462335 0.1671724653897556  Weight: 0.12  Decision Stump  Classifications  F <= 6900.5 : High  F > 6900.5 : Medium  F is missing : High  Class distributions  F <= 6900.5  Low High Medium  0.44858995696465337 0.4713558974865103 0.08005414554883639  F > 6900.5  Low High Medium  8.959360289682118E-17 0.0010693750275671651 0.9989306249724327  F is missing  Low High Medium  0.40668495096538565 0.4274241080225543 0.1658909410120601  Weight: 0.08  Decision Stump  Classifications  F <= 6900.5 : Low  F > 6900.5 : Medium  F is missing : Low  Class distributions  F <= 6900.5  Low High Medium  0.46600509176163624 0.4508329050255522 0.08316200321281156  F > 6900.5  Low High Medium  2.7378556647094443E-16 0.0011613536843342318 0.9988386463156654  F is missing  Low High Medium  0.42419509803700617 0.41048835713949894 0.16531654482349495  Weight: 0.06  Decision Stump  Classifications  F <= 6900.5 : High  F > 6900.5 : Medium  F is missing : High  Class distributions  F <= 6900.5  Low High Medium  0.45228376555509586 0.462417332627905 0.08529890181699905  F > 6900.5  Low High Medium  3.1839683698968197E-16 0.0012272527351914842 0.9987727472648082  F is missing  Low High Medium  0.4127929818163871 0.4221489014028135 0.16505811678079954  Weight: 0.04  Decision Stump  Classifications  F <= 6900.5 : Low  F > 6900.5 : Medium  F is missing : Low  Class distributions  F <= 6900.5  Low High Medium  0.4600103105007572 0.45323359233618754 0.08675609716305534  F > 6900.5  Low High Medium  -6.233553403014792E-16 0.0012734517642136135 0.998726548235787  F is missing  Low High Medium  0.420572533210263 0.414485975247662 0.164941491542075  Weight: 0.02  Decision Stump  Classifications  F <= 6900.5 : High  F > 6900.5 : Medium  F is missing : High  Class distributions  F <= 6900.5  Low High Medium  0.45386030379037745 0.4583955235517955 0.08774417265782701  F > 6900.5  Low High Medium  -1.3475309283726975E-16 0.0013053659036594463 0.9986946340963407  F is missing  Low High Medium  0.4154248099655636 0.4196865071761702 0.16488868285826627  Weight: 0.02  Decision Stump  Classifications  F <= 6900.5 : Low  F > 6900.5 : Medium  F is missing : Low  Class distributions  F <= 6900.5  Low High Medium  0.4573124661009792 0.4542759595685834 0.08841157433043745  F > 6900.5  Low High Medium  1.7286797517788696E-16 0.0013271934411143045 0.9986728065588856  F is missing  Low High Medium  0.418902666693741 0.4162326685184955 0.16486466478776354  Weight: 0.01  Number of performed Iterations: 10  Time taken to build model: 15.47 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.43 seconds  === Summary ===  Correctly Classified Instances 196578 70.2398 %  Incorrectly Classified Instances 83289 29.7602 %  Kappa statistic 0.2806  Mean absolute error 0.3335  Root mean squared error 0.3845  Relative absolute error 93.1491 %  Root relative squared error 90.8497 %  Total Number of Instances 279867  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  1.000 0.790 0.677 1.000 0.808 0.377 0.700 0.734 Low  0.000 0.000 0.000 0.000 0.000 0.000 0.500 0.244 High  0.525 0.000 0.997 0.525 0.688 0.695 1.000 0.998 Medium  Weighted Avg. 0.702 0.493 0.572 0.702 0.607 0.339 0.699 0.663  === Confusion Matrix ===  a b c <-- classified as  174551 0 0 | a = Low  63290 0 67 | b = High  19932 0 22027 | c = Medium |
| LogitBoost | === Run information ===  Scheme: weka.classifiers.meta.LogitBoost -P 100 -L -1.7976931348623157E308 -H 1.0 -Z 3.0 -O 1 -E 1 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump -batch-size  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  LogitBoost: Base classifiers and their weights:  Iteration 1  Class 1 (Drop=Low)  Decision Stump  Classifications  E <= 1000.78522 : 1.5512933184566817  E > 1000.78522 : -1.5000000000304805  E is missing : 1.3091583671322615  Class 2 (Drop=High)  Decision Stump  Classifications  Y <= 176.6975 : -0.7848567523498508  Y > 176.6975 : 2.9965464313794836  Y is missing : -0.4855649963896224  Class 3 (Drop=Medium)  Decision Stump  Classifications  E <= 1000.78522 : -1.1521816583942617  E > 1000.78522 : 2.988563992638124  E is missing : -0.8235933707381978  Iteration 2  Class 1 (Drop=Low)  Decision Stump  Classifications  Z <= 68.5 : 0.25679847695206637  Z > 68.5 : -2.9991680987504745  Z is missing : -0.061146238505125065  Class 2 (Drop=High)  Decision Stump  Classifications  Z <= 68.5 : -0.23276602727959095  Z > 68.5 : 2.9994748609885615  Z is missing : 0.23499557340367574  Class 3 (Drop=Medium)  Decision Stump  Classifications  F <= 6900.5 : -0.9444086644905488  F > 6900.5 : 2.9651499011901263  F is missing : -0.21260417858033187  Iteration 3  Class 1 (Drop=Low)  Decision Stump  Classifications  X <= 180.99450000000002 : 0.3165467769275217  X > 180.99450000000002 : -2.9070886253274626  X is missing : -0.07801575878346942  Class 2 (Drop=High)  Decision Stump  Classifications  X <= 180.99450000000002 : -0.510670487509703  X > 180.99450000000002 : 2.902573988595933  X is missing : 0.05875381420028744  Class 3 (Drop=Medium)  Decision Stump  Classifications  F <= 6900.5 : -0.5829862507804531  F > 6900.5 : 2.1421023386996785  F is missing : 0.0673000889348816  Iteration 4  Class 1 (Drop=Low)  Decision Stump  Classifications  Y <= 180.99450000000002 : 0.263921210850963  Y > 180.99450000000002 : -1.882716023833082  Y is missing : -0.07478349580964433  Class 2 (Drop=High)  Decision Stump  Classifications  Y <= 180.99450000000002 : -0.47177843584068635  Y > 180.99450000000002 : 1.9524898288412966  Y is missing : 0.02834917342096341  Class 3 (Drop=Medium)  Decision Stump  Classifications  E <= 1000.78522 : -0.3035222178284942  E > 1000.78522 : 1.3299196129369233  E is missing : 0.11175205122859326  Iteration 5  Class 1 (Drop=Low)  Decision Stump  Classifications  F <= 6900.5 : 0.12623513789320726  F > 6900.5 : -1.4209525695375076  F is missing : -0.12831362508208752  Class 2 (Drop=High)  Decision Stump  Classifications  Z <= 68.5 : -0.2922109518909327  Z > 68.5 : 1.3078542920981886  Z is missing : 0.04920221298612628  Class 3 (Drop=Medium)  Decision Stump  Classifications  F <= 6900.5 : -0.451053194524476  F > 6900.5 : 1.45125581769797  F is missing : 0.184948950031954  Iteration 6  Class 1 (Drop=Low)  Decision Stump  Classifications  X <= 180.99450000000002 : 0.13380007848897071  X > 180.99450000000002 : -1.28147323573817  X is missing : -0.10274238529093249  Class 2 (Drop=High)  Decision Stump  Classifications  X <= 180.5785 : -0.28930631766610454  X > 180.5785 : 1.3452065979078132  X is missing : 0.10745636930804867  Class 3 (Drop=Medium)  Decision Stump  Classifications  E <= 1000.78522 : -0.548877061993477  E > 1000.78522 : 1.1789806022232867  E is missing : 0.04786029637655276  Iteration 7  Class 1 (Drop=Low)  Decision Stump  Classifications  Z <= 68.5 : 0.05520215377106049  Z > 68.5 : -1.1464448975879369  Z is missing : -0.13103613265783545  Class 2 (Drop=High)  Decision Stump  Classifications  Z <= 68.5 : -0.2301214972265533  Z > 68.5 : 1.1670353039464596  Z is missing : 0.08613942672254032  Class 3 (Drop=Medium)  Decision Stump  Classifications  F <= 6900.5 : -0.22144489590202282  F > 6900.5 : 1.102516981195682  F is missing : 0.1675931129899884  Iteration 8  Class 1 (Drop=Low)  Decision Stump  Classifications  Y <= 180.99450000000002 : 0.1630229800315196  Y > 180.99450000000002 : -1.1499123932358157  Y is missing : -0.09278372378361334  Class 2 (Drop=High)  Decision Stump  Classifications  Y <= 180.99450000000002 : -0.3887467087920442  Y > 180.99450000000002 : 1.159665074079725  Y is missing : 0.036165036200572875  Class 3 (Drop=Medium)  Decision Stump  Classifications  E <= 1001.5753050000001 : -0.3757678126338612  E > 1001.5753050000001 : 1.0816982330220342  E is missing : 0.16393148869635346  Iteration 9  Class 1 (Drop=Low)  Decision Stump  Classifications  X <= 180.99450000000002 : 0.04265092931160904  X > 180.99450000000002 : -1.0883185748800808  X is missing : -0.1317064242839745  Class 2 (Drop=High)  Decision Stump  Classifications  X <= 174.9105 : -0.2950008377784096  X > 174.9105 : 1.1101461411727278  X is missing : 0.05857349859870486  Class 3 (Drop=Medium)  Decision Stump  Classifications  F <= 6900.5 : -0.2091831560958078  F > 6900.5 : 1.0657993085553512  F is missing : 0.21143080423545774  Iteration 10  Class 1 (Drop=Low)  Decision Stump  Classifications  Z <= 68.5 : 0.06418743627337051  Z > 68.5 : -1.057183114822879  Z is missing : -0.10053430861602103  Class 2 (Drop=High)  Decision Stump  Classifications  Z <= 68.5 : -0.3047328477487498  Z > 68.5 : 1.0703878635017072  Z is missing : 0.02197673050497714  Class 3 (Drop=Medium)  Decision Stump  Classifications  E <= 1000.78522 : -0.29440482210429175  E > 1000.78522 : 1.023546927260955  E is missing : 0.19712671719869357  Number of performed iterations: 10  Time taken to build model: 60.09 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.91 seconds  === Summary ===  Correctly Classified Instances 279792 99.9732 %  Incorrectly Classified Instances 75 0.0268 %  Kappa statistic 0.9995  Mean absolute error 0.015  Root mean squared error 0.0222  Relative absolute error 4.1949 %  Root relative squared error 5.2565 %  Total Number of Instances 279867  === 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 Low  0.999 0.000 1.000 0.999 0.999 0.999 1.000 1.000 High  1.000 0.000 0.998 1.000 0.999 0.999 1.000 1.000 Medium  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000  === Confusion Matrix ===  a b c <-- classified as  174551 0 0 | a = Low  0 63282 75 | b = High  0 0 41959 | c = Medium |
| meta.MultiClassClassifier | === Run information ===  Scheme: weka.classifiers.meta.MultiClassClassifier -M 0 -R 2.0 -S 1 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M -1 -num-decimal-places 4  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  MultiClassClassifier  Classifier 1, using indicator values: Strings: 1  Invert: false  Cols: 1  Logistic Regression with ridge parameter of 1.0E-8  Coefficients...  Class  Variable neg\_Low  ====================  Z 0.2892  F 0.0014  E 0.012  X 0.0561  Y 0.0583  Intercept -20.6353  Odds Ratios...  Class  Variable neg\_Low  ====================  Z 1.3353  F 1.0014  E 1.0121  X 1.0577  Y 1.0601  Classifier 2, using indicator values: Strings: 2  Invert: false  Cols: 2  Logistic Regression with ridge parameter of 1.0E-8  Coefficients...  Class  Variable neg\_High  =====================  Z -0.7362  F -0  E 0.0409  X -0.2066  Y -0.2149  Intercept 58.3369  Odds Ratios...  Class  Variable neg\_High  =====================  Z 0.4789  F 1  E 1.0417  X 0.8133  Y 0.8066  Classifier 3, using indicator values: Strings: 3  Invert: false  Cols: 3  Logistic Regression with ridge parameter of 1.0E-8  Coefficients...  Class  Variable neg\_Medium  =======================  Z 0.3947  F -0.0027  E -0.0232  X 0.0183  Y 0.0245  Intercept 17.4379  Odds Ratios...  Class  Variable neg\_Medium  =======================  Z 1.4839  F 0.9973  E 0.9771  X 1.0184  Y 1.0248  Time taken to build model: 85.56 seconds  === Evaluation on test split ===  Time taken to test model on training split: 1.38 seconds  === Summary ===  Correctly Classified Instances 277657 99.2103 %  Incorrectly Classified Instances 2210 0.7897 %  Kappa statistic 0.9853  Mean absolute error 0.0101  Root mean squared error 0.0711  Relative absolute error 2.8334 %  Root relative squared error 16.789 %  Total Number of Instances 279867  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.995 0.012 0.993 0.995 0.994 0.983 0.997 0.998 Low  0.979 0.003 0.989 0.979 0.984 0.980 0.996 0.994 High  1.000 0.001 0.994 1.000 0.997 0.996 1.000 0.999 Medium  Weighted Avg. 0.992 0.008 0.992 0.992 0.992 0.985 0.998 0.997  === Confusion Matrix ===  a b c <-- classified as  173646 686 219 | a = Low  1264 62056 37 | b = High  0 4 41955 | c = Medium |
| OneR | === Run information ===  Scheme: weka.classifiers.rules.OneR -B 6  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  E:  < 59.40024 -> Low  < 59.454535 -> High  < 81.99895000000001 -> Low  < 82.02966 -> High  < 114.909625 -> Low  < 114.97967 -> High  < 122.575425 -> Low  < 122.64578 -> High  < 149.88584500000002 -> Low  < 150.078885 -> High  < 160.09302 -> Low  < 160.22501499999998 -> High  < 197.91489 -> Low  < 198.09047 -> High  < 198.55177 -> Low  < 198.68809 -> High  < 204.43301 -> Low  < 204.64257 -> High  < 205.428335 -> Low  < 205.571875 -> High  < 208.668315 -> Low  < 208.91586 -> High  < 215.82909 -> Low  < 216.052515 -> High  < 238.247915 -> Low  < 238.427605 -> High  < 249.633505 -> Low  < 249.87808 -> High  < 282.826515 -> Low  < 283.01716999999996 -> High  < 293.474485 -> Low  < 293.690405 -> High  < 293.82626 -> Low  < 294.002615 -> High  < 337.22308 -> Low  < 337.44555 -> High  < 338.00096499999995 -> Low  < 338.109465 -> High  < 348.258185 -> Low  < 348.43805499999996 -> High  < 370.603285 -> Low  < 370.77139999999997 -> High  < 381.00753 -> Low  < 381.18634499999996 -> High  < 435.241535 -> Low  < 435.441915 -> High  < 1000.78522 -> Low  >= 1000.78522 -> Medium  (579055/823138 instances correct)  Time taken to build model: 1.01 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.33 seconds  === Summary ===  Correctly Classified Instances 193991 69.3154 %  Incorrectly Classified Instances 85876 30.6846 %  Kappa statistic 0.2682  Mean absolute error 0.2046  Root mean squared error 0.4523  Relative absolute error 57.1422 %  Root relative squared error 106.8742 %  Total Number of Instances 279867  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.982 0.783 0.675 0.982 0.800 0.331 0.599 0.674 Low  0.010 0.015 0.157 0.010 0.018 -0.020 0.497 0.226 High  0.525 0.000 0.995 0.525 0.687 0.694 0.762 0.594 Medium  Weighted Avg. 0.693 0.492 0.606 0.693 0.606 0.306 0.601 0.561  === Confusion Matrix ===  a b c <-- classified as  171348 3171 32 | a = Low  62674 613 70 | b = High  19814 115 22030 | c = Medium |
| Decision Table | === Run information ===  Scheme: weka.classifiers.rules.DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N 5"  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Decision Table:  Number of training instances: 823138  Number of Rules : 62  Non matches covered by Majority class.  Best first.  Start set: no attributes  Search direction: forward  Stale search after 5 node expansions  Total number of subsets evaluated: 22  Merit of best subset found: 99.998  Evaluation (for feature selection): CV (leave one out)  Feature set: 2,3,4,5,6,7  Time taken to build model: 10.53 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.45 seconds  === Summary ===  Correctly Classified Instances 279861 99.9979 %  Incorrectly Classified Instances 6 0.0021 %  Kappa statistic 1  Mean absolute error 0.0001  Root mean squared error 0.004  Relative absolute error 0.0262 %  Root relative squared error 0.9547 %  Total Number of Instances 279867  === 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 Low  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 High  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 Medium  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000  === Confusion Matrix ===  a b c <-- classified as  174551 0 0 | a = Low  5 63352 0 | b = High  1 0 41958 | c = Medium |
| J48 Tree | === Run information ===  Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  J48 pruned tree  ------------------  E <= 1000.57044  | F <= 4800  | | Y <= 180.989  | | | Z <= 46  | | | | X <= 180.989: Low (513850.0)  | | | | X > 180.989: High (55646.0)  | | | Z > 46: High (64763.0)  | | Y > 180.989: High (64951.0)  | F > 4800  | | X <= 174.821  | | | Y <= 175.822: Medium (58540.0/2.0)  | | | Y > 175.822  | | | | Y <= 181.429: Medium (18.0)  | | | | Y > 181.429: High (7.0)  | | X > 174.821  | | | X <= 181.448  | | | | Y <= 181.059: Medium (18.0)  | | | | Y > 181.059: High (6.0)  | | | X > 181.448: High (19.0)  E > 1000.57044  | X <= 180.314  | | Y <= 181.059  | | | Z <= 46: Medium (65154.0)  | | | Z > 46: High (44.0)  | | Y > 181.059: High (56.0)  | X > 180.314: High (66.0)  Number of Leaves : 14  Size of the tree : 27  Time taken to build model: 7.76 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.33 seconds  === Summary ===  Correctly Classified Instances 279860 99.9975 %  Incorrectly Classified Instances 7 0.0025 %  Kappa statistic 1  Mean absolute error 0  Root mean squared error 0.0041  Relative absolute error 0.005 %  Root relative squared error 0.9649 %  Total Number of Instances 279867  === 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 Low  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 High  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 Medium  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000  === Confusion Matrix ===  a b c <-- classified as  174550 1 0 | a = Low  0 63356 1 | b = High  0 5 41954 | c = Medium |
|  |  |
| RandomForest | === 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: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === 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: 236.87 seconds  === Evaluation on test split ===  Time taken to test model on training split: 3.9 seconds  === Summary ===  Correctly Classified Instances 279862 99.9982 %  Incorrectly Classified Instances 5 0.0018 %  Kappa statistic 1  Mean absolute error 0  Root mean squared error 0.0032  Relative absolute error 0.0074 %  Root relative squared error 0.7449 %  Total Number of Instances 279867  === 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 Low  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 High  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 Medium  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000  === Confusion Matrix ===  a b c <-- classified as  174551 0 0 | a = Low  0 63356 1 | b = High  0 4 41955 | c = Medium |
| randomTree | === Run information ===  Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  RandomTree  ==========  F < 6900.5  | E < 1000.79  | | Y < 180.99  | | | X < 180.99  | | | | Z < 68.5 : Low (513850/0)  | | | | Z >= 68.5 : High (61661/0)  | | | X >= 180.99 : High (58748/0)  | | Y >= 180.99 : High (64951/0)  | E >= 1000.79  | | Z < 68.1  | | | X < 180.38  | | | | Y < 181.11 : Medium (62030/0)  | | | | Y >= 181.11 : High (55/0)  | | | X >= 180.38 : High (63/0)  | | Z >= 68.1 : High (46/0)  F >= 6900.5  | Y < 177.13  | | Z < 78  | | | X < 181.45 : Medium (61694/0)  | | | X >= 181.45 : High (12/0)  | | Z >= 78 : High (3/0)  | Y >= 177.13  | | X < 18.4 : Medium (2/0)  | | X >= 18.4  | | | Y < 181.08  | | | | X < 180.42 : Medium (2/0)  | | | | X >= 180.42 : High (5/0)  | | | Y >= 181.08 : High (16/0)  Size of the tree : 29  Time taken to build model: 2.29 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.22 seconds  === Summary ===  Correctly Classified Instances 279860 99.9975 %  Incorrectly Classified Instances 7 0.0025 %  Kappa statistic 1  Mean absolute error 0  Root mean squared error 0.0041  Relative absolute error 0.0047 %  Root relative squared error 0.9649 %  Total Number of Instances 279867  === 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 Low  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 High  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 Medium  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000  === Confusion Matrix ===  a b c <-- classified as  174547 0 4 | a = Low  0 63357 0 | b = High  3 0 41956 | c = Medium |
| RepTree | === Run information ===  Scheme: weka.classifiers.trees.REPTree -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  REPTree  ============  E < 1000.79  | F < 6900.5  | | Z < 68.5  | | | Y < 180.99  | | | | X < 180.84 : Low (342567/0) [171281/0]  | | | | X >= 180.84 : High (36968/0) [18680/2]  | | | Y >= 180.99 : High (43299/0) [21636/0]  | | Z >= 68.5 : High (43311/0) [21468/0]  | F >= 6900.5  | | X < 174.9  | | | Z < 76.8  | | | | Y < 182.75 : Medium (38979/0) [19578/1]  | | | | Y >= 182.75 : High (2/0) [4/0]  | | | Z >= 76.8 : High (2/0) [0/0]  | | X >= 174.9  | | | X < 181.45  | | | | Y < 181.26 : Medium (12/0) [7/1]  | | | | Y >= 181.26 : High (4/0) [1/0]  | | | X >= 181.45 : High (9/0) [10/0]  E >= 1000.79  | X < 179.81  | | Y < 180.09  | | | Z < 70.1 : Medium (43494/0) [21658/0]  | | | Z >= 70.1 : High (29/0) [15/0]  | | Y >= 180.09 : High (38/0) [19/1]  | X >= 179.81 : High (44/0) [23/1]  Size of the tree : 27  Time taken to build model: 5 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.21 seconds  === Summary ===  Correctly Classified Instances 279860 99.9975 %  Incorrectly Classified Instances 7 0.0025 %  Kappa statistic 1  Mean absolute error 0  Root mean squared error 0.0041  Relative absolute error 0.006 %  Root relative squared error 0.9704 %  Total Number of Instances 279867  === 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 Low  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 High  1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 Medium  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000  === Confusion Matrix ===  a b c <-- classified as  174550 1 0 | a = Low  0 63356 1 | b = High  0 5 41954 | c = Medium |
|  |  |
| SimpleKMeans | === Run information ===  Scheme: weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Ignored:  Drop  Test mode: Classes to clusters evaluation on training data  === Clustering model (full training set) ===  kMeans  ======  Number of iterations: 11  Within cluster sum of squared errors: 30245.041410039325  Initial starting points (random):  Cluster 0: 1,0,0,74.13867,63.122,340.85  Cluster 1: 1,0,10529,53.67962,144.396,173.298  Cluster 2: 1,0,3600,6.38519,124.302,390.782  Missing values globally replaced with mean/mode  Final cluster centroids:  Cluster#  Attribute Full Data 0 1 2  (823138.0) (642526.0) (61734.0) (118878.0)  =======================================================  G1 1 1 1 1  Z 10.7896 13.1812 0.0559 3.4373  F 1707.0224 717.7451 14189.0673 571.9917  E 142.862 168.7855 100.5891 24.6999  X 116.6894 96.8017 96.3935 234.7207  Y 117.1147 95.7345 95.7446 243.771  Time taken to build model (full training data) : 4.09 seconds  === Model and evaluation on training set ===  Clustered Instances  0 642526 ( 78%)  1 61734 ( 7%)  2 118878 ( 14%)  Class attribute: Drop  Classes to Clusters:  0 1 2 <-- assigned to cluster  512958 0 892 | Low  67538 36 117986 | High  62030 61698 0 | Medium  Cluster 0 <-- Low  Cluster 1 <-- Medium  Cluster 2 <-- High  Incorrectly clustered instances : 130496.0 15.8535 % |
| MakeDensityBasedClusterer | === Run information ===  Scheme: weka.clusterers.MakeDensityBasedClusterer -M 1.0E-6 -W weka.clusterers.SimpleKMeans -- -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Ignored:  Drop  Test mode: Classes to clusters evaluation on training data  === Clustering model (full training set) ===  MakeDensityBasedClusterer:  Wrapped clusterer:  kMeans  ======  Number of iterations: 11  Within cluster sum of squared errors: 30245.041410039325  Initial starting points (random):  Cluster 0: 1,0,0,74.13867,63.122,340.85  Cluster 1: 1,0,10529,53.67962,144.396,173.298  Cluster 2: 1,0,3600,6.38519,124.302,390.782  Missing values globally replaced with mean/mode  Final cluster centroids:  Cluster#  Attribute Full Data 0 1 2  (823138.0) (642526.0) (61734.0) (118878.0)  =======================================================  G1 1 1 1 1  Z 10.7896 13.1812 0.0559 3.4373  F 1707.0224 717.7451 14189.0673 571.9917  E 142.862 168.7855 100.5891 24.6999  X 116.6894 96.8017 96.3935 234.7207  Y 117.1147 95.7345 95.7446 243.771  Fitted estimators (with ML estimates of variance):  Cluster: 0 Prior probability: 0.7806  Attribute: G1  Normal Distribution. Mean = 1 StdDev = 0  Attribute: Z  Normal Distribution. Mean = 13.1812 StdDev = 41.0852  Attribute: F  Normal Distribution. Mean = 717.7451 StdDev = 1549.6194  Attribute: E  Normal Distribution. Mean = 168.7855 StdDev = 459.3622  Attribute: X  Normal Distribution. Mean = 96.8017 StdDev = 35.8826  Attribute: Y  Normal Distribution. Mean = 95.7345 StdDev = 30.0233  Cluster: 1 Prior probability: 0.075  Attribute: G1  Normal Distribution. Mean = 1 StdDev = 0  Attribute: Z  Normal Distribution. Mean = 0.0559 StdDev = 4.1648  Attribute: F  Normal Distribution. Mean = 14189.0673 StdDev = 3018.9034  Attribute: E  Normal Distribution. Mean = 100.5891 StdDev = 347.8173  Attribute: X  Normal Distribution. Mean = 96.3935 StdDev = 33.4108  Attribute: Y  Normal Distribution. Mean = 95.7446 StdDev = 28.1301  Cluster: 2 Prior probability: 0.1444  Attribute: G1  Normal Distribution. Mean = 1 StdDev = 0  Attribute: Z  Normal Distribution. Mean = 3.4373 StdDev = 21.8053  Attribute: F  Normal Distribution. Mean = 571.9917 StdDev = 1432.5605  Attribute: E  Normal Distribution. Mean = 24.6999 StdDev = 85.53  Attribute: X  Normal Distribution. Mean = 234.7207 StdDev = 143.4192  Attribute: Y  Normal Distribution. Mean = 243.771 StdDev = 141.0728  Time taken to build model (full training data) : 8.2 seconds  === Model and evaluation on training set ===  Clustered Instances  0 606622 ( 74%)  1 61729 ( 7%)  2 154787 ( 19%)  Log likelihood: -18.67943  Class attribute: Drop  Classes to Clusters:  0 1 2 <-- assigned to cluster  481491 0 32359 | Low  63101 31 122428 | High  62030 61698 0 | Medium  Cluster 0 <-- Low  Cluster 1 <-- Medium  Cluster 2 <-- High  Incorrectly clustered instances : 157521.0 19.1366 % |
| FarthestFirst | === Run information ===  Scheme: weka.clusterers.FarthestFirst -N 3 -S 1  Relation: G1-Data-Supervised - multiclass  Instances: 823138  Attributes: 7  G1  Z  F  E  X  Y  Ignored:  Drop  Test mode: Classes to clusters evaluation on training data  === Clustering model (full training set) ===  FarthestFirst  ==============  Cluster centroids:  Cluster 0  1.0 93.0 0.0 4.93384 134.226 89.687  Cluster 1  1.0 0.0 17554.0 3054.46939 91.8 108.2  Cluster 2  1.0 1000.0 16876.0 2.72811 130.922 109.355  Time taken to build model (full training data) : 3.95 seconds  === Model and evaluation on training set ===  Clustered Instances  0 819642 (100%)  1 3495 ( 0%)  2 1 ( 0%)  Class attribute: Drop  Classes to Clusters:  0 1 2 <-- assigned to cluster  513850 0 0 | Low  185557 2 1 | High  120235 3493 0 | Medium  Cluster 0 <-- Low  Cluster 1 <-- Medium  Cluster 2 <-- High  Incorrectly clustered instances : 305794.0 37.1498 % |
|  |  |
|  |  |

* **All-Instr-****SemiSupervised.csv**. This file contains all instructions but only G1 instructions have been labelled (single class i.e. Drop column -> Yes or No) while other instructions have not been labelled.
  + **Total instructions** – 844,802
  + **No** - 556,326
  + **Yes** - 265,204
  + **Blank** - 23,272

|  |  |
| --- | --- |
| ZeroR | === Run information ===  Scheme: weka.classifiers.rules.ZeroR  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  ZeroR predicts class value: No  Time taken to build model: 0.06 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.3 seconds  === Summary ===  Correctly Classified Instances 189249 67.7549 %  Incorrectly Classified Instances 90065 32.2451 %  Kappa statistic 0  Mean absolute error 0.4371  Root mean squared error 0.4674  Relative absolute error 100 %  Root relative squared error 100 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  1.000 1.000 0.678 1.000 0.808 0.000 0.500 0.659 No  0.000 0.000 0.000 0.000 0.000 0.000 0.500 0.314 Yes  Weighted Avg. 0.678 0.678 0.459 0.678 0.547 0.000 0.500 0.548  === Confusion Matrix ===  a b <-- classified as  189249 0 | a = No  90065 0 | b = Yes |
| NaiveBayes | === Run information ===  Scheme: weka.classifiers.bayes.NaiveBayes  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Naive Bayes Classifier  Class  Attribute No Yes  (0.68) (0.32)  ===================================  M107  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  M190  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  S  mean 42.7641 4.6242  std. dev. 134.9318 46.3373  weight sum 556326 265204  precision 1.9934 1.9934  M104  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  G28  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  F  mean 698.5244 3383.4927  std. dev. 1536.189 5725.6144  weight sum 556326 265204  precision 1.9535 1.9535  G1  mean 1 1  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  Z  mean 0.0418 27.3111  std. dev. 0.9119 55.8979  weight sum 556326 265204  precision 0.2997 0.2997  M109  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  G21  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  G90  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  M82  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  E  mean 23.8091 340.6006  std. dev. 81.7941 630.9198  weight sum 556326 265204  precision 0.0166 0.0166  G92  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  X  mean 96.2842 150.5614  std. dev. 33.2509 115.7767  weight sum 556326 265204  precision 0.0062 0.0062  Y  mean 95.6789 150.012  std. dev. 28.027 114.5772  weight sum 556326 265204  precision 0.0073 0.0073  M140  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  M106  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  M84  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  M220  mean 0 0  std. dev. 0.0017 0.0017  weight sum 556326 265204  precision 0.01 0.01  Time taken to build model: 1.91 seconds  === Evaluation on test split ===  Time taken to test model on training split: 3.92 seconds  === Summary ===  Correctly Classified Instances 267848 95.8949 %  Incorrectly Classified Instances 11466 4.1051 %  Kappa statistic 0.9086  Mean absolute error 0.0293  Root mean squared error 0.1459  Relative absolute error 6.7079 %  Root relative squared error 31.206 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.944 0.010 0.995 0.944 0.969 0.911 0.975 0.973 No  0.990 0.056 0.894 0.990 0.940 0.911 0.985 0.958 Yes  Weighted Avg. 0.959 0.025 0.962 0.959 0.959 0.911 0.979 0.968  === Confusion Matrix ===  a b <-- classified as  178660 10589 | a = No  877 89188 | b = Yes |
| Logistic Regression | === Run information ===  Scheme: weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Logistic Regression with ridge parameter of 1.0E-8  Coefficients...  Class  Variable No  ====================  S 0.0052  F -0.0014  Z -0.2903  E -0.0118  X -0.0566  Y -0.0579  Intercept 20.7326  Odds Ratios...  Class  Variable No  ====================  S 1.0052  F 0.9986  Z 0.7481  E 0.9882  X 0.945  Y 0.9437  Time taken to build model: 29.81 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.64 seconds  === Summary ===  Correctly Classified Instances 275816 98.7476 %  Incorrectly Classified Instances 3498 1.2524 %  Kappa statistic 0.9713  Mean absolute error 0.0227  Root mean squared error 0.1039  Relative absolute error 5.1912 %  Root relative squared error 22.2345 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.993 0.024 0.989 0.993 0.991 0.971 0.942 0.901 No  0.976 0.007 0.985 0.976 0.980 0.971 0.997 0.989 Yes  Weighted Avg. 0.987 0.019 0.987 0.987 0.987 0.971 0.960 0.929  === Confusion Matrix ===  a b <-- classified as  187935 1314 | a = No  2184 87881 | b = Yes |
| MLP | === Run information ===  Scheme: weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Sigmoid Node 0  Inputs Weights  Threshold -19.85067942687471  Node 2 -15.100123111921517  Node 3 12.837401929060047  Node 4 -13.26395013147198  Node 5 13.277333179612805  Node 6 13.419566633502306  Node 7 1.8567388118409593  Node 8 -24.208085634982915  Node 9 -15.097802511277548  Node 10 -14.867305492560098  Node 11 -3.162005486744766  Node 12 3.236967889549179  Sigmoid Node 1  Inputs Weights  Threshold 19.85084434546159  Node 2 15.099661990672745  Node 3 -12.838310127553527  Node 4 13.263913364848273  Node 5 -13.276471711385316  Node 6 -13.41937952041084  Node 7 -1.882195809857957  Node 8 24.209109167832427  Node 9 15.098164125222532  Node 10 14.867015494478617  Node 11 3.1569932533527765  Node 12 -3.2149923329548775  Sigmoid Node 2  Inputs Weights  Threshold 11.954121015405205  Attrib M107 -0.039910987219127225  Attrib M190 -0.020160478178097685  Attrib S -3.077812019531936  Attrib M104 -0.0037132854805144166  Attrib G28 0.014325201159616288  Attrib F 25.955556729601486  Attrib G1 -0.015055062115807726  Attrib Z 1.909809612359326  Attrib M109 0.030911548727283036  Attrib G21 0.014573937396722705  Attrib G90 0.024662013475019853  Attrib M82 0.0023120360213984767  Attrib E 1.070279212320967  Attrib G92 0.022704960228617843  Attrib X 4.00213699413126  Attrib Y -8.384018429943898  Attrib M140 -0.04585767849763711  Attrib M106 0.003692678377666389  Attrib M84 -0.027872098686640014  Attrib M220 0.021748740534021357  Sigmoid Node 3  Inputs Weights  Threshold -16.791880987768735  Attrib M107 -0.009442679614649077  Attrib M190 0.026885018519084244  Attrib S -0.886004826051193  Attrib M104 -0.015035571872998657  Attrib G28 -0.0014175235494634925  Attrib F 1.7262048476256882  Attrib G1 -0.004389767040305803  Attrib Z 7.517132000811289  Attrib M109 0.019982471248294234  Attrib G21 0.01737812792148455  Attrib G90 -0.007229001284765654  Attrib M82 -0.047348817205845155  Attrib E -0.07369162407482105  Attrib G92 0.0033270772957177744  Attrib X -1.9891917524288047  Attrib Y -72.96596976993837  Attrib M140 0.021678240795154097  Attrib M106 -0.022390614666537714  Attrib M84 0.042379650575486524  Attrib M220 -0.0497882043113446  Sigmoid Node 4  Inputs Weights  Threshold 21.751152188155075  Attrib M107 -0.03557245852798849  Attrib M190 0.04852197966270158  Attrib S -21.369935440869835  Attrib M104 -0.03159883430201357  Attrib G28 -0.0322039573309635  Attrib F 0.12299517400844101  Attrib G1 0.03562783466085594  Attrib Z 12.369482411897808  Attrib M109 -0.012863447782479442  Attrib G21 0.027494884617053628  Attrib G90 0.022193115053612197  Attrib M82 0.020339204203689815  Attrib E 8.123382812152725  Attrib G92 0.01871297699690977  Attrib X 69.0090203554719  Attrib Y -0.15652040883751422  Attrib M140 0.03338155270190325  Attrib M106 -0.016707170176783658  Attrib M84 0.0013122411162361375  Attrib M220 -0.038829410498645106  Sigmoid Node 5  Inputs Weights  Threshold -10.140573227432428  Attrib M107 -0.04180823067211196  Attrib M190 -0.005006934440935119  Attrib S 9.885575274135501  Attrib M104 -0.018807563855809364  Attrib G28 -0.02124705973845925  Attrib F 9.025540122460319  Attrib G1 -0.012417738572855316  Attrib Z -3.285819901540864  Attrib M109 0.031961924658614  Attrib G21 -0.029939810341620188  Attrib G90 -0.0032269939774173667  Attrib M82 0.022613370318051954  Attrib E -0.00956499204032945  Attrib G92 0.02293604834819782  Attrib X -2.6800106609843226  Attrib Y -73.5352822756414  Attrib M140 -0.04794260230003679  Attrib M106 0.0040725484796537564  Attrib M84 -0.029945025704165885  Attrib M220 0.008569642488035836  Sigmoid Node 6  Inputs Weights  Threshold -23.654316896235585  Attrib M107 0.04771042138089912  Attrib M190 -0.04283614553476739  Attrib S 23.88280464731806  Attrib M104 -0.03749737555903  Attrib G28 0.023352217379551052  Attrib F 6.868462951711105  Attrib G1 0.006854067545700213  Attrib Z -41.335915173410825  Attrib M109 0.01615842501369727  Attrib G21 0.0030991581839103916  Attrib G90 0.03146044407584346  Attrib M82 -0.011269577070153622  Attrib E 7.6511025530169965  Attrib G92 -0.04979605762976115  Attrib X -60.851947677096405  Attrib Y 1.3750490409162088  Attrib M140 0.009753092018673096  Attrib M106 4.787814416399652E-4  Attrib M84 -0.03451931861365262  Attrib M220 0.017755284246809805  Sigmoid Node 7  Inputs Weights  Threshold -1.8148118594200182  Attrib M107 0.0394572001514477  Attrib M190 -0.013892221416409133  Attrib S 4.2780784050415095  Attrib M104 -0.015271302382555418  Attrib G28 -0.010365501854749201  Attrib F -0.6928947760223487  Attrib G1 0.046834657129544016  Attrib Z -0.020354954274624117  Attrib M109 2.244465141079538E-4  Attrib G21 -0.012144475169706924  Attrib G90 -0.014309043225704643  Attrib M82 -0.032095773945528906  Attrib E 0.6003402905339472  Attrib G92 -0.04534842245065504  Attrib X 0.020178205009175765  Attrib Y -1.6337982145852723  Attrib M140 -0.043557626681318545  Attrib M106 -0.03547286830012976  Attrib M84 -0.027150439155535767  Attrib M220 -0.04674970468384827  Sigmoid Node 8  Inputs Weights  Threshold 24.111761982412535  Attrib M107 -0.039428241899761546  Attrib M190 0.018249665560533868  Attrib S -0.9552376798101678  Attrib M104 0.037323630243030884  Attrib G28 -0.013400496658466668  Attrib F 2.7221366580159905  Attrib G1 0.008240167639980048  Attrib Z -16.657956832370793  Attrib M109 -0.0025201975633610263  Attrib G21 0.04388720218214308  Attrib G90 0.02404671774212136  Attrib M82 0.017673292510568178  Attrib E 105.35850229734869  Attrib G92 0.024147459905823107  Attrib X 5.572178388031088  Attrib Y -4.338779854586645  Attrib M140 0.023483237962975537  Attrib M106 -0.013922846676635285  Attrib M84 -0.027033410920233536  Attrib M220 -0.012852513888725117  Sigmoid Node 9  Inputs Weights  Threshold 7.389713431064906  Attrib M107 -0.0291608001528503  Attrib M190 -0.0220060597952924  Attrib S 10.82475968805993  Attrib M104 -0.042861225480774846  Attrib G28 -0.022603754221252693  Attrib F 16.13364709295929  Attrib G1 0.03916343547835731  Attrib Z -6.942450894902248  Attrib M109 0.006082771237505634  Attrib G21 0.022653963469668248  Attrib G90 0.04452673406383158  Attrib M82 -0.013152271904209534  Attrib E -6.64193568145738  Attrib G92 -0.04183720393240781  Attrib X -5.297941325360963  Attrib Y 8.774299947365718  Attrib M140 0.014273985026890992  Attrib M106 -0.03380711537244019  Attrib M84 0.04191917106474102  Attrib M220 -0.030938750922222338  Sigmoid Node 10  Inputs Weights  Threshold 1.565054233646416  Attrib M107 0.018048419034228907  Attrib M190 -0.04703631140734496  Attrib S -7.320355046949645  Attrib M104 0.02097350505787335  Attrib G28 0.010587588727425344  Attrib F -14.55752499894117  Attrib G1 0.03159438695063274  Attrib Z 27.37048439682671  Attrib M109 0.007012270594579505  Attrib G21 0.0166997856032573  Attrib G90 -0.03334513871978924  Attrib M82 -0.03006079925313525  Attrib E 19.544154536555826  Attrib G92 -0.009015487010529157  Attrib X -35.867317465672265  Attrib Y 23.55722733059263  Attrib M140 -0.007254995346911118  Attrib M106 0.022935706981833698  Attrib M84 0.011291560084488339  Attrib M220 -0.0064398772841314095  Sigmoid Node 11  Inputs Weights  Threshold -2.709913159505111  Attrib M107 0.04963336376696724  Attrib M190 0.024512707499340425  Attrib S -1.8965441957315774  Attrib M104 -0.03757543838862365  Attrib G28 0.010334401750429968  Attrib F -5.016930404375514  Attrib G1 0.022790700545982395  Attrib Z 1.5607380094770622  Attrib M109 0.03895403495196258  Attrib G21 -0.020898848301481143  Attrib G90 -0.03199663010160857  Attrib M82 -0.04107667585880648  Attrib E 11.712904537560538  Attrib G92 0.033677135096275165  Attrib X 1.038608317765133  Attrib Y -1.475564549592372  Attrib M140 0.012049949262758833  Attrib M106 -0.017059911359165993  Attrib M84 -0.042298628049646014  Attrib M220 0.002633562776410206  Sigmoid Node 12  Inputs Weights  Threshold -1.5434928832909651  Attrib M107 0.022784750421348005  Attrib M190 -0.04590973919276016  Attrib S 6.055790867616816  Attrib M104 0.026526934908149974  Attrib G28 -0.006637094819710122  Attrib F 0.015168461540243285  Attrib G1 0.03567414034339822  Attrib Z 0.008338591019505482  Attrib M109 -0.02010921131013803  Attrib G21 -0.016851602625313912  Attrib G90 -0.0122519169024334  Attrib M82 0.043046892944361564  Attrib E 0.592037659649291  Attrib G92 -0.0479232634921042  Attrib X 0.21479617269077872  Attrib Y -2.6277819424190625  Attrib M140 -0.014660775843414797  Attrib M106 -0.0061249640584449955  Attrib M84 -0.021550006826998314  Attrib M220 -0.01030565170645864  Class No  Input  Node 0  Class Yes  Input  Node 1  Time taken to build model: 1525.5 seconds  === Evaluation on test split ===  Time taken to test model on training split: 1.8 seconds  === Summary ===  Correctly Classified Instances 279270 99.9842 %  Incorrectly Classified Instances 44 0.0158 %  Kappa statistic 0.9996  Mean absolute error 0.0002  Root mean squared error 0.0115  Relative absolute error 0.0483 %  Root relative squared error 2.4604 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === 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 0.956 0.930 No  1.000 0.000 1.000 1.000 1.000 1.000 0.997 0.989 Yes  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 0.969 0.949  === Confusion Matrix ===  a b <-- classified as  189242 7 | a = No  37 90028 | b = Yes |
| AdaBoostM1 | === Run information ===  Scheme: weka.classifiers.meta.AdaBoostM1 -P 100 -S 1 -I 10 -W weka.classifiers.trees.DecisionStump  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  AdaBoostM1: Base classifiers and their weights:  Decision Stump  Classifications  E <= 1000.78522 : No  E > 1000.78522 : Yes  E is missing : No  Class distributions  E <= 1000.78522  No Yes  0.7268007498905865 0.2731992501094135  E > 1000.78522  No Yes  0.0 1.0  E is missing  No Yes  0.6771828174260222 0.32281718257397785  Weight: 1.07  Decision Stump  Classifications  X <= 180.99450000000002 : No  X > 180.99450000000002 : Yes  X is missing : Yes  Class distributions  X <= 180.99450000000002  No Yes  0.5202444908381655 0.47975550916183457  X > 180.99450000000002  No Yes  5.682864910075588E-17 1.0  X is missing  No Yes  0.45420967291632436 0.5457903270836757  Weight: 0.33  Decision Stump  Classifications  Y <= 180.99450000000002 : Yes  Y > 180.99450000000002 : Yes  Y is missing : Yes  Class distributions  Y <= 180.99450000000002  No Yes  0.4604788213178792 0.5395211786821207  Y > 180.99450000000002  No Yes  2.3787045354993445E-16 0.9999999999999998  Y is missing  No Yes  0.3907919357460916 0.6092080642539084  Weight: 0.44  Decision Stump  Classifications  Y <= 180.99450000000002 : No  Y > 180.99450000000002 : Yes  Y is missing : No  Class distributions  Y <= 180.99450000000002  No Yes  0.5709111037852413 0.4290888962147587  Y > 180.99450000000002  No Yes  -7.76309720609642E-17 1.0000000000000002  Y is missing  No Yes  0.5000000000068455 0.4999999999931546  Weight: 0.51  Decision Stump  Classifications  Z <= 68.5 : Yes  Z > 68.5 : Yes  Z is missing : Yes  Class distributions  Z <= 68.5  No Yes  0.47957153775925243 0.5204284622407477  Z > 68.5  No Yes  2.1222369853768694E-16 0.9999999999999998  Z is missing  No Yes  0.40050822284772547 0.5994917771522745  Weight: 0.4  Decision Stump  Classifications  Z <= 68.5 : No  Z > 68.5 : Yes  Z is missing : No  Class distributions  Z <= 68.5  No Yes  0.5797113567720782 0.4202886432279218  Z > 68.5  No Yes  -1.2211269763796421E-16 1.0000000000000002  Z is missing  No Yes  0.5000000000041388 0.4999999999958612  Weight: 0.56  Decision Stump  Classifications  F <= 6900.5 : Yes  F > 6900.5 : Yes  F is missing : Yes  Class distributions  F <= 6900.5  No Yes  0.48022125519074216 0.5197787448092578  F > 6900.5  No Yes  -1.8079237420475905E-16 1.0000000000000002  F is missing  No Yes  0.392155753650455 0.6078442463495449  Weight: 0.44  Decision Stump  Classifications  F <= 6900.5 : No  F > 6900.5 : Yes  F is missing : Yes  Class distributions  F <= 6900.5  No Yes  0.588823350833019 0.41117664916698105  F > 6900.5  No Yes  5.278641330646758E-17 1.0  F is missing  No Yes  0.4999999999967782 0.5000000000032219  Weight: 0.62  Decision Stump  Classifications  X <= 180.99450000000002 : Yes  X > 180.99450000000002 : Yes  X is missing : Yes  Class distributions  X <= 180.99450000000002  No Yes  0.45809889891831523 0.5419011010816849  X > 180.99450000000002  No Yes  9.095257068302642E-17 1.0  X is missing  No Yes  0.384113736020774 0.6158862639792261  Weight: 0.47  Decision Stump  Classifications  X <= 180.99450000000002 : No  X > 180.99450000000002 : Yes  X is missing : No  Class distributions  X <= 180.99450000000002  No Yes  0.5754506444065212 0.4245493555934788  X > 180.99450000000002  No Yes  -2.3997975114624073E-16 1.0000000000000002  X is missing  No Yes  0.5000000000022553 0.4999999999977447  Weight: 0.54  Number of performed Iterations: 10  Time taken to build model: 29.98 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.41 seconds  === Summary ===  Correctly Classified Instances 208425 74.6203 %  Incorrectly Classified Instances 70889 25.3797 %  Kappa statistic 0.2682  Mean absolute error 0.2819  Root mean squared error 0.3606  Relative absolute error 64.4786 %  Root relative squared error 77.158 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  1.000 0.787 0.727 1.000 0.842 0.394 0.972 0.972 No  0.213 0.000 1.000 0.213 0.351 0.394 0.994 0.974 Yes  Weighted Avg. 0.746 0.533 0.815 0.746 0.684 0.394 0.979 0.973  === Confusion Matrix ===  a b <-- classified as  189249 0 | a = No  70889 19176 | b = Yes |
| Bagging | === Run information ===  Scheme: weka.classifiers.meta.Bagging -P 100 -S 1 -num-slots 1 -I 10 -W weka.classifiers.trees.REPTree -- -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Bagging with 10 iterations and base learner  weka.classifiers.trees.REPTree -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0  Time taken to build model: 175.51 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.65 seconds  === Summary ===  Correctly Classified Instances 279310 99.9986 %  Incorrectly Classified Instances 4 0.0014 %  Kappa statistic 1  Mean absolute error 0  Root mean squared error 0.0037  Relative absolute error 0.0046 %  Root relative squared error 0.787 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === 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 0.972 0.972 No  1.000 0.000 1.000 1.000 1.000 1.000 0.994 0.973 Yes  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 0.979 0.972  === Confusion Matrix ===  a b <-- classified as  189245 4 | a = No  0 90065 | b = Yes |
| meta.MultiClassClassifier | === Run information ===  Scheme: weka.classifiers.meta.MultiClassClassifier -M 0 -R 2.0 -S 1 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M -1 -num-decimal-places 4  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  MultiClassClassifier  Classifier 1  Logistic Regression with ridge parameter of 1.0E-8  Coefficients...  Class  Variable No  ====================  S 0.0052  F -0.0014  Z -0.2903  E -0.0118  X -0.0566  Y -0.0579  Intercept 20.7326  Odds Ratios...  Class  Variable No  ====================  S 1.0052  F 0.9986  Z 0.7481  E 0.9882  X 0.945  Y 0.9437  Time taken to build model: 60.19 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.67 seconds  === Summary ===  Correctly Classified Instances 275816 98.7476 %  Incorrectly Classified Instances 3498 1.2524 %  Kappa statistic 0.9713  Mean absolute error 0.0227  Root mean squared error 0.1039  Relative absolute error 5.1912 %  Root relative squared error 22.2345 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.993 0.024 0.989 0.993 0.991 0.971 0.942 0.901 No  0.976 0.007 0.985 0.976 0.980 0.971 0.997 0.989 Yes  Weighted Avg. 0.987 0.019 0.987 0.987 0.987 0.971 0.960 0.929  === Confusion Matrix ===  a b <-- classified as  187935 1314 | a = No  2184 87881 | b = Yes |
| OneR | === Run information ===  Scheme: weka.classifiers.rules.OneR -B 6  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  E:  < 2.949635 -> No  < 2.950245 -> Yes  < 7.359265000000001 -> No  < 7.359484999999999 -> Yes  < 8.528765 -> No  < 8.52936 -> Yes  < 11.890754999999999 -> No  < 11.89176 -> Yes  < 13.84176 -> No  < 13.842369999999999 -> Yes  < 14.042265 -> No  < 14.04458 -> Yes  < 14.24344 -> No  < 14.244795 -> Yes  < 15.20128 -> No  < 15.203600000000002 -> Yes  < 18.012864999999998 -> No  < 18.01726 -> Yes  < 18.2687 -> No  < 18.27108 -> Yes  < 24.84082 -> No  < 24.85111 -> Yes  < 30.24996 -> No  < 30.268250000000002 -> Yes  < 36.548395 -> No  < 36.55815 -> Yes  < 59.920515 -> No  < 59.93567 -> Yes  < 67.7576 -> No  < 67.77168 -> Yes  < 70.01095000000001 -> No  < 70.03642500000001 -> Yes  < 70.149895 -> No  < 70.170745 -> Yes  < 70.311635 -> No  < 70.327045 -> Yes  < 71.10797 -> No  < 71.14335 -> Yes  < 89.599785 -> No  < 89.63648 -> Yes  < 101.00479999999999 -> No  < 101.032335 -> Yes  < 101.17459 -> No  < 101.20841 -> Yes  < 108.18589 -> No  < 108.20197 -> Yes  < 111.53202999999999 -> No  < 111.557265 -> Yes  < 114.909625 -> No  < 114.97967 -> Yes  < 115.10499 -> No  < 115.20228 -> Yes  < 115.23377500000001 -> No  < 115.266605 -> Yes  < 115.776555 -> No  < 115.847575 -> Yes  < 122.575425 -> No  < 122.628015 -> Yes  < 149.258775 -> No  < 149.345 -> Yes  < 155.39821 -> No  < 155.604985 -> Yes  < 171.24968 -> No  < 171.451785 -> Yes  < 187.762355 -> No  < 187.91027 -> Yes  < 193.74296 -> No  < 193.98691000000002 -> Yes  < 194.53761500000002 -> No  < 194.86158999999998 -> Yes  < 197.91489 -> No  < 198.09047 -> Yes  < 198.55177 -> No  < 198.68809 -> Yes  < 204.43301 -> No  < 204.64257 -> Yes  < 205.428335 -> No  < 205.571875 -> Yes  < 238.59273 -> No  < 238.796565 -> Yes  < 238.97042 -> No  < 239.177305 -> Yes  < 249.03796499999999 -> No  < 249.344715 -> Yes  < 260.39327000000003 -> No  < 260.584965 -> Yes  < 282.49318 -> No  < 282.683495 -> Yes  < 282.826515 -> No  < 283.04268 -> Yes  < 293.072455 -> No  < 293.23415 -> Yes  < 293.6305 -> No  < 294.09527 -> Yes  < 370.603285 -> No  < 370.81049499999995 -> Yes  < 380.76853 -> No  < 380.98308 -> Yes  < 381.32375 -> No  < 381.50106 -> Yes  < 381.69055000000003 -> No  < 381.91092000000003 -> Yes  < 414.19912999999997 -> No  < 414.436375 -> Yes  < 424.50912 -> No  < 424.77096 -> Yes  < 435.52711999999997 -> No  < 435.78555 -> Yes  < 1000.78522 -> No  >= 1000.78522 -> Yes  (612514/821530 instances correct)  Time taken to build model: 1.88 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.45 seconds  === Summary ===  Correctly Classified Instances 205143 73.4453 %  Incorrectly Classified Instances 74171 26.5547 %  Kappa statistic 0.2468  Mean absolute error 0.2655  Root mean squared error 0.5153  Relative absolute error 60.7451 %  Root relative squared error 110.2474 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.979 0.780 0.725 0.979 0.833 0.334 0.593 0.704 No  0.220 0.021 0.835 0.220 0.348 0.334 0.599 0.424 Yes  Weighted Avg. 0.734 0.535 0.761 0.734 0.677 0.334 0.595 0.614  === Confusion Matrix ===  a b <-- classified as  185352 3897 | a = No  70274 19791 | b = Yes |
| DecisionTable | === Run information ===  Scheme: weka.classifiers.rules.DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N 5"  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  Decision Table:  Number of training instances: 821530  Number of Rules : 44  Non matches covered by Majority class.  Best first.  Start set: no attributes  Search direction: forward  Stale search after 5 node expansions  Total number of subsets evaluated: 160  Merit of best subset found: 99.998  Evaluation (for feature selection): CV (leave one out)  Feature set: 6,8,13,15,16,21  Time taken to build model: 118.14 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.38 seconds  === Summary ===  Correctly Classified Instances 279299 99.9946 %  Incorrectly Classified Instances 15 0.0054 %  Kappa statistic 0.9999  Mean absolute error 0.0001  Root mean squared error 0.0056  Relative absolute error 0.019 %  Root relative squared error 1.2042 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === 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 0.972 0.972 No  1.000 0.000 1.000 1.000 1.000 1.000 0.994 0.974 Yes  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 0.979 0.972  === Confusion Matrix ===  a b <-- classified as  189246 3 | a = No  12 90053 | b = Yes |
| RandomForest | === 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: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === 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: 508.6 seconds  === Evaluation on test split ===  Time taken to test model on training split: 4.9 seconds  === Summary ===  Correctly Classified Instances 279311 99.9989 %  Incorrectly Classified Instances 3 0.0011 %  Kappa statistic 1  Mean absolute error 0.0001  Root mean squared error 0.0034  Relative absolute error 0.0262 %  Root relative squared error 0.7171 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === 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 0.976 0.976 No  1.000 0.000 1.000 1.000 1.000 1.000 0.998 0.992 Yes  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 0.983 0.981  === Confusion Matrix ===  a b <-- classified as  189246 3 | a = No  0 90065 | b = Yes |
| randomTree | === Run information ===  Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  RandomTree  ==========  Z < 68.5  | E < 1000.79  | | X < 180.99  | | | E < 45.13  | | | | E < 45.08  | | | | | F < 6900.5  | | | | | | S < 150.5  | | | | | | | F < 4500  | | | | | | | | X < 174.9  | | | | | | | | | Y < 180.92 : No (423470/0)  | | | | | | | | | Y >= 180.92 : Yes (44316/0)  | | | | | | | | X >= 174.9  | | | | | | | | | E < 18.95  | | | | | | | | | | E < 13.27 : No (3/0)  | | | | | | | | | | E >= 13.27  | | | | | | | | | | | Y < 153.34 : No (2/0)  | | | | | | | | | | | Y >= 153.34 : Yes (1/0)  | | | | | | | | | E >= 18.95 : Yes (7/0)  | | | | | | | F >= 4500  | | | | | | | | Y < 178.9 : No (39825/0)  | | | | | | | | Y >= 178.9 : Yes (4234/0)  | | | | | | S >= 150.5  | | | | | | | S < 448.5  | | | | | | | | X < 176.66  | | | | | | | | | F < 225  | | | | | | | | | | Y < 179.57 : No (18931/0)  | | | | | | | | | | Y >= 179.57 : Yes (1/0)  | | | | | | | | | F >= 225 : No (4785/0)  | | | | | | | | X >= 176.66 : Yes (1/0)  | | | | | | | S >= 448.5 : No (24788/0)  | | | | | F >= 6900.5 : Yes (46178/0)  | | | | E >= 45.08  | | | | | E < 45.13 : No (34/0)  | | | | | E >= 45.13  | | | | | | F < 8436 : No (6/0)  | | | | | | F >= 8436 : Yes (1/0)  | | | E >= 45.13  | | | | X < 174.81  | | | | | S < 150.5  | | | | | | E < 45.64  | | | | | | | E < 45.64  | | | | | | | | X < 166.81  | | | | | | | | | F < 6804.5  | | | | | | | | | | Y < 218.25 : No (204/0)  | | | | | | | | | | Y >= 218.25 : Yes (21/0)  | | | | | | | | | F >= 6804.5 : Yes (26/0)  | | | | | | | | X >= 166.81  | | | | | | | | | E < 45.22  | | | | | | | | | | Y < 228 : No (9/0)  | | | | | | | | | | Y >= 228 : Yes (1/0)  | | | | | | | | | E >= 45.22  | | | | | | | | | | E < 45.23  | | | | | | | | | | | F < 6098 : No (1/0)  | | | | | | | | | | | F >= 6098 : Yes (1/0)  | | | | | | | | | | E >= 45.23  | | | | | | | | | | | E < 45.26  | | | | | | | | | | | | E < 45.25  | | | | | | | | | | | | | Y < 246.07 : No (5/0)  | | | | | | | | | | | | | Y >= 246.07 : Yes (2/0)  | | | | | | | | | | | | E >= 45.25 : No (2/0)  | | | | | | | | | | | E >= 45.26  | | | | | | | | | | | | F < 5603  | | | | | | | | | | | | | X < 169.26 : No (2/0)  | | | | | | | | | | | | | X >= 169.26  | | | | | | | | | | | | | | X < 174.39  | | | | | | | | | | | | | | | E < 45.44  | | | | | | | | | | | | | | | | Y < 190.19 : No (1/0)  | | | | | | | | | | | | | | | | Y >= 190.19 : Yes (1/0)  | | | | | | | | | | | | | | | E >= 45.44  | | | | | | | | | | | | | | | | Y < 269.29 : No (4/0)  | | | | | | | | | | | | | | | | Y >= 269.29 : Yes (2/0)  | | | | | | | | | | | | | | X >= 174.39  | | | | | | | | | | | | | | | Y < 312.94 : No (7/0)  | | | | | | | | | | | | | | | Y >= 312.94 : Yes (1/0)  | | | | | | | | | | | | F >= 5603 : Yes (3/0)  | | | | | | | E >= 45.64  | | | | | | | | Y < 195.57  | | | | | | | | | F < 7786 : No (4/0)  | | | | | | | | | F >= 7786 : Yes (2/0)  | | | | | | | | Y >= 195.57 : Yes (2/0)  | | | | | | E >= 45.64  | | | | | | | Y < 180.99  | | | | | | | | Y < 180.8  | | | | | | | | | X < 16.57  | | | | | | | | | | E < 52.21  | | | | | | | | | | | E < 48.2  | | | | | | | | | | | | X < 6.14  | | | | | | | | | | | | | E < 46.58  | | | | | | | | | | | | | | F < 12140 : No (8/0)  | | | | | | | | | | | | | | F >= 12140 : Yes (2/0)  | | | | | | | | | | | | | E >= 46.58 : No (6/0)  | | | | | | | | | | | | X >= 6.14 : No (9/0)  | | | | | | | | | | | E >= 48.2  | | | | | | | | | | | | E < 49.76  | | | | | | | | | | | | | F < 12446 : No (2/0)  | | | | | | | | | | | | | F >= 12446 : Yes (1/0)  | | | | | | | | | | | | E >= 49.76  | | | | | | | | | | | | | F < 8884 : No (2/0)  | | | | | | | | | | | | | F >= 8884 : Yes (2/0)  | | | | | | | | | | E >= 52.21  | | | | | | | | | | | F < 7520 : No (169/0)  | | | | | | | | | | | F >= 7520 : Yes (8/0)  | | | | | | | | | X >= 16.57  | | | | | | | | | | Y < 179.47  | | | | | | | | | | | E < 45.89  | | | | | | | | | | | | X < 30.97  | | | | | | | | | | | | | F < 7241.5 : No (20/0)  | | | | | | | | | | | | | F >= 7241.5 : Yes (2/0)  | | | | | | | | | | | | X >= 30.97  | | | | | | | | | | | | | X < 75.98 : No (21/0)  | | | | | | | | | | | | | X >= 75.98  | | | | | | | | | | | | | | E < 45.78  | | | | | | | | | | | | | | | X < 76.43  | | | | | | | | | | | | | | | | F < 8547.5 : No (2/0)  | | | | | | | | | | | | | | | | F >= 8547.5 : Yes (1/0)  | | | | | | | | | | | | | | | X >= 76.43  | | | | | | | | | | | | | | | | F < 5491 : No (44/0)  | | | | | | | | | | | | | | | | F >= 5491 : Yes (2/0)  | | | | | | | | | | | | | | E >= 45.78 : No (33/0)  | | | | | | | | | | | E >= 45.89  | | | | | | | | | | | | X < 16.71  | | | | | | | | | | | | | F < 7294.5 : No (2/0)  | | | | | | | | | | | | | F >= 7294.5 : Yes (2/0)  | | | | | | | | | | | | X >= 16.71  | | | | | | | | | | | | | E < 45.9  | | | | | | | | | | | | | | E < 45.9  | | | | | | | | | | | | | | | F < 7206.5 : No (5/0)  | | | | | | | | | | | | | | | F >= 7206.5 : Yes (1/0)  | | | | | | | | | | | | | | E >= 45.9  | | | | | | | | | | | | | | | Y < 107.02  | | | | | | | | | | | | | | | | F < 8577 : No (2/0)  | | | | | | | | | | | | | | | | F >= 8577 : Yes (2/0)  | | | | | | | | | | | | | | | Y >= 107.02  | | | | | | | | | | | | | | | | X < 85.25 : No (1/0)  | | | | | | | | | | | | | | | | X >= 85.25  | | | | | | | | | | | | | | | | | F < 7406 : No (1/0)  | | | | | | | | | | | | | | | | | F >= 7406 : Yes (1/0)  | | | | | | | | | | | | | E >= 45.9  | | | | | | | | | | | | | | E < 46.09  | | | | | | | | | | | | | | | X < 117.35 : No (50/0)  | | | | | | | | | | | | | | | X >= 117.35  | | | | | | | | | | | | | | | | F < 5849 : No (36/0)  | | | | | | | | | | | | | | | | F >= 5849 : Yes (3/0)  | | | | | | | | | | | | | | E >= 46.09  | | | | | | | | | | | | | | | E < 46.51  | | | | | | | | | | | | | | | | E < 46.48  | | | | | | | | | | | | | | | | | E < 46.47  | | | | | | | | | | | | | | | | | | F < 6060 : No (192/0)  | | | | | | | | | | | | | | | | | | F >= 6060 : Yes (25/0)  | | | | | | | | | | | | | | | | | E >= 46.47 : No (8/0)  | | | | | | | | | | | | | | | | E >= 46.48  | | | | | | | | | | | | | | | | | F < 4963.5 : No (6/0)  | | | | | | | | | | | | | | | | | F >= 4963.5 : Yes (3/0)  | | | | | | | | | | | | | | | E >= 46.51  | | | | | | | | | | | | | | | | F < 6600.5 : No (39341/0)  | | | | | | | | | | | | | | | | F >= 6600.5 : Yes (3733/0)  | | | | | | | | | | Y >= 179.47 : No (8/0)  | | | | | | | | Y >= 180.8  | | | | | | | | | F < 5455.5 : No (1/0)  | | | | | | | | | F >= 5455.5 : Yes (1/0)  | | | | | | | Y >= 180.99 : Yes (4286/0)  | | | | | S >= 150.5  | | | | | | E < 45.9 : No (34/0)  | | | | | | E >= 45.9  | | | | | | | S < 320.5  | | | | | | | | X < 24.78  | | | | | | | | | S < 314.5 : Yes (2/0)  | | | | | | | | | S >= 314.5  | | | | | | | | | | E < 72.94 : No (1/0)  | | | | | | | | | | E >= 72.94 : Yes (1/0)  | | | | | | | | X >= 24.78  | | | | | | | | | X < 49.79 : No (32/0)  | | | | | | | | | X >= 49.79  | | | | | | | | | | F < 5575 : No (218/0)  | | | | | | | | | | F >= 5575 : Yes (20/0)  | | | | | | | S >= 320.5  | | | | | | | | S < 339.5  | | | | | | | | | E < 87.35 : No (80/0)  | | | | | | | | | E >= 87.35  | | | | | | | | | | E < 88.12 : Yes (1/0)  | | | | | | | | | | E >= 88.12  | | | | | | | | | | | S < 323.5  | | | | | | | | | | | | S < 322.5 : No (16/0)  | | | | | | | | | | | | S >= 322.5  | | | | | | | | | | | | | E < 302.08 : No (7/0)  | | | | | | | | | | | | | E >= 302.08  | | | | | | | | | | | | | | Y < 91.79  | | | | | | | | | | | | | | | F < 8367.5 : No (1/0)  | | | | | | | | | | | | | | | F >= 8367.5 : Yes (1/0)  | | | | | | | | | | | | | | Y >= 91.79 : No (2/0)  | | | | | | | | | | | S >= 323.5 : No (139/0)  | | | | | | | | S >= 339.5  | | | | | | | | | F < 6332.5  | | | | | | | | | | F < 450  | | | | | | | | | | | S < 346.5  | | | | | | | | | | | | E < 162.55 : No (52/0)  | | | | | | | | | | | | E >= 162.55  | | | | | | | | | | | | | S < 345.5  | | | | | | | | | | | | | | X < 21.63 : Yes (1/0)  | | | | | | | | | | | | | | X >= 21.63 : No (37/0)  | | | | | | | | | | | | | S >= 345.5  | | | | | | | | | | | | | | X < 27.65 : Yes (1/0)  | | | | | | | | | | | | | | X >= 27.65 : No (5/0)  | | | | | | | | | | | S >= 346.5  | | | | | | | | | | | | S < 531.5  | | | | | | | | | | | | | Y < 181.81 : No (2425/0)  | | | | | | | | | | | | | Y >= 181.81 : Yes (1/0)  | | | | | | | | | | | | S >= 531.5  | | | | | | | | | | | | | S < 532.5  | | | | | | | | | | | | | | E < 179.46 : No (9/0)  | | | | | | | | | | | | | | E >= 179.46  | | | | | | | | | | | | | | | E < 226.6 : Yes (1/0)  | | | | | | | | | | | | | | | E >= 226.6 : No (1/0)  | | | | | | | | | | | | | S >= 532.5  | | | | | | | | | | | | | | Y < 181.35 : No (935/0)  | | | | | | | | | | | | | | Y >= 181.35 : Yes (1/0)  | | | | | | | | | | F >= 450 : No (142/0)  | | | | | | | | | F >= 6332.5 : Yes (187/0)  | | | | X >= 174.81  | | | | | Y < 181.08  | | | | | | S < 154.5  | | | | | | | X < 179.04  | | | | | | | | E < 769.99  | | | | | | | | | F < 5206.5 : No (86/0)  | | | | | | | | | F >= 5206.5 : Yes (11/0)  | | | | | | | | E >= 769.99 : No (18/0)  | | | | | | | X >= 179.04 : No (18/0)  | | | | | | S >= 154.5 : No (16/0)  | | | | | Y >= 181.08 : Yes (69/0)  | | X >= 180.99 : Yes (53052/0)  | E >= 1000.79 : Yes (56052/0)  Z >= 68.5 : Yes (52927/0)  Size of the tree : 233  Time taken to build model: 5.28 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.33 seconds  === Summary ===  Correctly Classified Instances 279278 99.9871 %  Incorrectly Classified Instances 36 0.0129 %  Kappa statistic 0.9997  Mean absolute error 0.0001  Root mean squared error 0.0114  Relative absolute error 0.0295 %  Root relative squared error 2.4289 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === 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 0.970 0.969 No  1.000 0.000 1.000 1.000 1.000 1.000 0.995 0.978 Yes  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 0.978 0.972  === Confusion Matrix ===  a b <-- classified as  189240 9 | a = No  27 90038 | b = Yes |
| RepTree | === Run information ===  Scheme: weka.classifiers.trees.REPTree -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  REPTree  ============  E < 1000.79  | Y < 180.99  | | X < 180.99  | | | Z < 68.5  | | | | F < 6900.5 : No (370884/0) [185442/0]  | | | | F >= 6900.5 : Yes (33384/0) [16833/0]  | | | Z >= 68.5 : Yes (35166/0) [17699/0]  | | X >= 180.99 : Yes (35256/0) [17760/0]  | Y >= 180.99 : Yes (35428/0) [17593/0]  E >= 1000.79 : Yes (37568/0) [18517/0]  Size of the tree : 11  Time taken to build model: 13.49 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.63 seconds  === Summary ===  Correctly Classified Instances 279304 99.9964 %  Incorrectly Classified Instances 10 0.0036 %  Kappa statistic 0.9999  Mean absolute error 0  Root mean squared error 0.006  Relative absolute error 0.0111 %  Root relative squared error 1.2801 %  Total Number of Instances 279314  Ignored Class Unknown Instances 7919  === 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 0.972 0.972 No  1.000 0.000 1.000 1.000 1.000 1.000 0.994 0.973 Yes  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 0.979 0.972  === Confusion Matrix ===  a b <-- classified as  189239 10 | a = No  0 90065 | b = Yes |
|  |  |
| SimpleKMeans | === Run information ===  Scheme: weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -M -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66% train, remainder test  === Clustering model (full training set) ===  kMeans  ======  Number of iterations: 6  Within cluster sum of squared errors: 124757.1676057535  Initial starting points (random):  Cluster 0: 0,0,0,0,0,0,1,0,0,0,0,0,1931.02917,0,103.26,89.349,0,0,0,0,Yes  Cluster 1: 0,0,0,0,0,0,1,0,0,0,0,0,3.57203,0,62.311,96.362,0,0,0,0,No  Cluster 2: 0,0,0,0,0,900,1,0,0,0,0,0,6.59796,0,141.826,89.687,0,0,0,0,No  Final cluster centroids:  Cluster#  Attribute Full Data 0 1 2  (844802.0) (265265.0) (21664.0) (557873.0)  =======================================================  M107 0.0006 0 0.0238 0  M190 0 0 0.0015 0  S 31.2872 4.6246 61.0089 42.8108  M104 0.0001 0 0.0044 0  G28 0.0001 0 0.0047 0  F 1545.0127 3383.5909 858.4084 697.4438  G1 0.9744 1 0 1  Z 8.8255 27.3108 8.6727 0.0417  M109 0 0 0.0016 0  G21 0 0 0.0016 0  G90 0 0 0.0016 0  M82 0 0 0.0016 0  E 125.2462 340.5792 100.1415 23.8317  G92 0.0209 0 0.8148 0  X 111.3278 150.5603 17.6113 96.3124  Y 110.7632 150.0129 18.0025 95.7024  M140 0.0001 0 0.0031 0  M106 0.0036 0 0.1397 0  M84 0 0 0.0016 0  M220 0 0 0.0001 0  Drop No Yes missing No  Time taken to build model (full training data) : 3.64 seconds  === Model and evaluation on test split ===  kMeans  ======  Number of iterations: 9  Within cluster sum of squared errors: 95790.84642128192  Initial starting points (random):  Cluster 0: 0,0,0,0,0,11746,1,0,0,0,0,0,1027.57295,0,87.524,100.793,0,0,0,0,Yes  Cluster 1: 0,0,0,0,0,0,1,0,0,0,0,0,8.61491,0,66.367,289.907,0,0,0,0,Yes  Cluster 2: 0,0,0,0,0,3600,1,0,0,0,0,0,2.5,0,0,0,0,0,0,0,No  Final cluster centroids:  Cluster#  Attribute Full Data 0 1 2  (557569.0) (72490.0) (104540.0) (380539.0)  =======================================================  M107 0.0006 0.0004 0.0001 0.0008  M190 0 0 0 0  S 31.3115 12.1934 0.1427 43.5159  M104 0.0001 0.0001 0 0.0001  G28 0.0001 0.0001 0 0.0001  F 1544.1463 7363.3432 663.5725 677.5377  G1 0.9744 0.9778 0.9976 0.9674  Z 8.8438 0.0458 46.1754 0.2642  M109 0 0 0 0  G21 0 0.0001 0 0  G90 0 0 0 0.0001  M82 0 0 0 0  E 125.0495 802.3234 24.8089 23.5713  G92 0.0209 0.0182 0.0018 0.0266  X 111.3633 94.3158 187.3129 93.7462  Y 110.7481 93.7754 186.4449 93.1862  M140 0.0001 0 0 0.0001  M106 0.0036 0.0032 0.0003 0.0045  M84 0 0.0001 0 0  M220 0 0 0 0  Drop No Yes Yes No  Time taken to build model (percentage split) : 4.78 seconds  Clustered Instances  0 37645 ( 13%)  1 53393 ( 19%)  2 196195 ( 68%) |
| FilteredClusterer | === Run information ===  Scheme: weka.clusterers.FilteredClusterer -F "weka.filters.AllFilter " -W weka.clusterers.SimpleKMeans -- -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -M -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66% train, remainder test  === Clustering model (full training set) ===  FilteredClusterer using weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -M -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10 on data filtered through weka.filters.AllFilter  Filtered Header  @relation 'Copy of All-Instr-SemiSupervised-1-weka.filters.AllFilter'  @attribute M107 numeric  @attribute M190 numeric  @attribute S numeric  @attribute M104 numeric  @attribute G28 numeric  @attribute F numeric  @attribute G1 numeric  @attribute Z numeric  @attribute M109 numeric  @attribute G21 numeric  @attribute G90 numeric  @attribute M82 numeric  @attribute E numeric  @attribute G92 numeric  @attribute X numeric  @attribute Y numeric  @attribute M140 numeric  @attribute M106 numeric  @attribute M84 numeric  @attribute M220 numeric  @attribute Drop {No,Yes}  @data  Clusterer Model  kMeans  ======  Number of iterations: 6  Within cluster sum of squared errors: 124757.1676057535  Initial starting points (random):  Cluster 0: 0,0,0,0,0,0,1,0,0,0,0,0,1931.02917,0,103.26,89.349,0,0,0,0,Yes  Cluster 1: 0,0,0,0,0,0,1,0,0,0,0,0,3.57203,0,62.311,96.362,0,0,0,0,No  Cluster 2: 0,0,0,0,0,900,1,0,0,0,0,0,6.59796,0,141.826,89.687,0,0,0,0,No  Final cluster centroids:  Cluster#  Attribute Full Data 0 1 2  (844802.0) (265265.0) (21664.0) (557873.0)  =======================================================  M107 0.0006 0 0.0238 0  M190 0 0 0.0015 0  S 31.2872 4.6246 61.0089 42.8108  M104 0.0001 0 0.0044 0  G28 0.0001 0 0.0047 0  F 1545.0127 3383.5909 858.4084 697.4438  G1 0.9744 1 0 1  Z 8.8255 27.3108 8.6727 0.0417  M109 0 0 0.0016 0  G21 0 0 0.0016 0  G90 0 0 0.0016 0  M82 0 0 0.0016 0  E 125.2462 340.5792 100.1415 23.8317  G92 0.0209 0 0.8148 0  X 111.3278 150.5603 17.6113 96.3124  Y 110.7632 150.0129 18.0025 95.7024  M140 0.0001 0 0.0031 0  M106 0.0036 0 0.1397 0  M84 0 0 0.0016 0  M220 0 0 0.0001 0  Drop No Yes missing No  Time taken to build model (full training data) : 3.65 seconds  === Model and evaluation on test split ===  FilteredClusterer using weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -M -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10 on data filtered through weka.filters.AllFilter  Filtered Header  @relation 'Copy of All-Instr-SemiSupervised-1-weka.filters.AllFilter'  @attribute M107 numeric  @attribute M190 numeric  @attribute S numeric  @attribute M104 numeric  @attribute G28 numeric  @attribute F numeric  @attribute G1 numeric  @attribute Z numeric  @attribute M109 numeric  @attribute G21 numeric  @attribute G90 numeric  @attribute M82 numeric  @attribute E numeric  @attribute G92 numeric  @attribute X numeric  @attribute Y numeric  @attribute M140 numeric  @attribute M106 numeric  @attribute M84 numeric  @attribute M220 numeric  @attribute Drop {No,Yes}  @data  Clusterer Model  kMeans  ======  Number of iterations: 9  Within cluster sum of squared errors: 95790.84642128192  Initial starting points (random):  Cluster 0: 0,0,0,0,0,11746,1,0,0,0,0,0,1027.57295,0,87.524,100.793,0,0,0,0,Yes  Cluster 1: 0,0,0,0,0,0,1,0,0,0,0,0,8.61491,0,66.367,289.907,0,0,0,0,Yes  Cluster 2: 0,0,0,0,0,3600,1,0,0,0,0,0,2.5,0,0,0,0,0,0,0,No  Final cluster centroids:  Cluster#  Attribute Full Data 0 1 2  (557569.0) (72490.0) (104540.0) (380539.0)  =======================================================  M107 0.0006 0.0004 0.0001 0.0008  M190 0 0 0 0  S 31.3115 12.1934 0.1427 43.5159  M104 0.0001 0.0001 0 0.0001  G28 0.0001 0.0001 0 0.0001  F 1544.1463 7363.3432 663.5725 677.5377  G1 0.9744 0.9778 0.9976 0.9674  Z 8.8438 0.0458 46.1754 0.2642  M109 0 0 0 0  G21 0 0.0001 0 0  G90 0 0 0 0.0001  M82 0 0 0 0  E 125.0495 802.3234 24.8089 23.5713  G92 0.0209 0.0182 0.0018 0.0266  X 111.3633 94.3158 187.3129 93.7462  Y 110.7481 93.7754 186.4449 93.1862  M140 0.0001 0 0 0.0001  M106 0.0036 0.0032 0.0003 0.0045  M84 0 0.0001 0 0  M220 0 0 0 0  Drop No Yes Yes No  Time taken to build model (percentage split) : 5.48 seconds  Clustered Instances  0 37645 ( 13%)  1 53393 ( 19%)  2 196195 ( 68%) |
| FarthestFirst | === Run information ===  Scheme: weka.clusterers.FarthestFirst -N 3 -S 1  Relation: Copy of All-Instr-SemiSupervised-1  Instances: 844802  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66% train, remainder test  === Clustering model (full training set) ===  FarthestFirst  ==============  Cluster centroids:  Cluster 0  0.0 0.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 0.0 0.0 0.0 9.28492 0.0 102.631 354.073 0.0 0.0 0.0 0.0 Yes  Cluster 1  0.0 0.0 687.3 0.0 0.0 16694.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 No  Cluster 2  0.0 0.0 0.0 0.0 0.0 0.0 0.0 182.0 0.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 0.0 0.0 0.0 0.0 No  Time taken to build model (full training data) : 1.9 seconds  === Model and evaluation on test split ===  FarthestFirst  ==============  Cluster centroids:  Cluster 0  0.0 0.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 0.0 0.0 0.0 4.88982 0.0 94.359 94.355 0.0 0.0 0.0 0.0 No  Cluster 1  0.0 0.0 687.3 0.0 0.0 16694.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0 0.0 0.0 No  Cluster 2  0.0 0.0 0.0 0.0 0.0 17196.0 0.0 0.0 0.0 0.0 0.0 0.0 2001.0 1.0 0.0 0.0 0.0 0.0 0.0 0.0 No  Time taken to build model (percentage split) : 1.85 seconds  Clustered Instances  0 280153 ( 98%)  1 1051 ( 0%)  2 6029 ( 2%) |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| RandomTRee with highly labelled semi-supervised | === Run information ===  Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1  Relation: highly labelled-semi supervised  Instances: 103052  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  RandomTree  ==========  Z < 58.7  | F < 6900.5  | | X < 142.06  | | | E < 1000.2  | | | | X < 82.7  | | | | | F < 1050  | | | | | | E < 10.13  | | | | | | | E < 6.8  | | | | | | | | S < 151  | | | | | | | | | E < 5.85  | | | | | | | | | | Z < 0.2  | | | | | | | | | | | Y < 194.66 : No (5743/0)  | | | | | | | | | | | Y >= 194.66 : Yes (710/0)  | | | | | | | | | | Z >= 0.2 : No (2/0)  | | | | | | | | | E >= 5.85  | | | | | | | | | | Y < 196.04 : No (1043/0)  | | | | | | | | | | Y >= 196.04 : Yes (95/0)  | | | | | | | | S >= 151 : No (430/0)  | | | | | | | E >= 6.8  | | | | | | | | E < 6.81 : Yes (3/0)  | | | | | | | | E >= 6.81  | | | | | | | | | F < 270  | | | | | | | | | | X < 74.52  | | | | | | | | | | | Y < 195.93 : No (1083/0)  | | | | | | | | | | | Y >= 195.93 : Yes (98/0)  | | | | | | | | | | X >= 74.52  | | | | | | | | | | | X < 81.31  | | | | | | | | | | | | X < 76.33  | | | | | | | | | | | | | E < 7.33  | | | | | | | | | | | | | | E < 6.91 : No (4/0)  | | | | | | | | | | | | | | E >= 6.91  | | | | | | | | | | | | | | | E < 6.93 : Yes (1/0)  | | | | | | | | | | | | | | | E >= 6.93  | | | | | | | | | | | | | | | | Y < 193.41 : No (14/0)  | | | | | | | | | | | | | | | | Y >= 193.41 : Yes (8/0)  | | | | | | | | | | | | | E >= 7.33  | | | | | | | | | | | | | | X < 75.72  | | | | | | | | | | | | | | | E < 9.7  | | | | | | | | | | | | | | | | S < 151.5  | | | | | | | | | | | | | | | | | E < 7.73 : No (5/0)  | | | | | | | | | | | | | | | | | E >= 7.73  | | | | | | | | | | | | | | | | | | Y < 200.29 : No (14/0)  | | | | | | | | | | | | | | | | | | Y >= 200.29 : Yes (8/0)  | | | | | | | | | | | | | | | | S >= 151.5 : No (9/0)  | | | | | | | | | | | | | | | E >= 9.7 : Yes (2/0)  | | | | | | | | | | | | | | X >= 75.72  | | | | | | | | | | | | | | | E < 7.47 : Yes (1/0)  | | | | | | | | | | | | | | | E >= 7.47  | | | | | | | | | | | | | | | | X < 76.31  | | | | | | | | | | | | | | | | | Y < 215.74 : No (83/0)  | | | | | | | | | | | | | | | | | Y >= 215.74 : Yes (10/0)  | | | | | | | | | | | | | | | | X >= 76.31 : No (10/0)  | | | | | | | | | | | | X >= 76.33  | | | | | | | | | | | | | E < 8.32  | | | | | | | | | | | | | | Y < 211.52 : No (55/0)  | | | | | | | | | | | | | | Y >= 211.52 : Yes (35/0)  | | | | | | | | | | | | | E >= 8.32  | | | | | | | | | | | | | | E < 9.42  | | | | | | | | | | | | | | | Y < 210.41 : No (40/0)  | | | | | | | | | | | | | | | Y >= 210.41 : Yes (6/0)  | | | | | | | | | | | | | | E >= 9.42  | | | | | | | | | | | | | | | S < 182  | | | | | | | | | | | | | | | | Y < 217.96 : No (10/0)  | | | | | | | | | | | | | | | | Y >= 217.96 : Yes (8/0)  | | | | | | | | | | | | | | | S >= 182 : No (2/0)  | | | | | | | | | | | X >= 81.31  | | | | | | | | | | | | Y < 239.78 : No (58/0)  | | | | | | | | | | | | Y >= 239.78 : Yes (4/0)  | | | | | | | | | F >= 270  | | | | | | | | | | F < 720  | | | | | | | | | | | X < 76.42 : No (2/0)  | | | | | | | | | | | X >= 76.42 : Yes (3/0)  | | | | | | | | | | F >= 720  | | | | | | | | | | | X < 78.63 : No (1/0)  | | | | | | | | | | | X >= 78.63 : Yes (1/0)  | | | | | | E >= 10.13  | | | | | | | Y < 196.12 : No (620/0)  | | | | | | | Y >= 196.12 : Yes (232/0)  | | | | | F >= 1050  | | | | | | Y < 152.85 : No (3300/0)  | | | | | | Y >= 152.85 : Yes (404/0)  | | | | X >= 82.7  | | | | | Y < 195.8 : No (26332/0)  | | | | | Y >= 195.8 : Yes (4552/0)  | | | E >= 1000.2 : Yes (6775/0)  | | X >= 142.06 : Yes (6298/0)  | F >= 6900.5 : Yes (6261/0)  Z >= 58.7 : Yes (6259/0)  Size of the tree : 89  Time taken to build model: 0.38 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.1 seconds  === Summary ===  Correctly Classified Instances 23913 99.9833 %  Incorrectly Classified Instances 4 0.0167 %  Kappa statistic 0.9997  Mean absolute error 0.0002  Root mean squared error 0.0129  Relative absolute error 0.0338 %  Root relative squared error 2.5977 %  Total Number of Instances 23917  Ignored Class Unknown Instances 11121  === 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 0.806 0.605 No  1.000 0.000 1.000 1.000 1.000 1.000 0.946 0.807 Yes  Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 0.869 0.696  === Confusion Matrix ===  a b <-- classified as  13074 4 | a = No  0 10839 | b = Yes |
| RandomTRee with balanced semi-supervised | === Run information ===  Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1  Relation: balanced-semi supervised  Instances: 64537  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  RandomTree  ==========  E < 1001.8  | S < 150.5  | | Z < 55.65  | | | E < 6.74  | | | | Y < 152.85  | | | | | E < 3.73  | | | | | | E < 2.51  | | | | | | | E < -0.5 : No (1/0)  | | | | | | | E >= -0.5  | | | | | | | | X < 164.24  | | | | | | | | | Z < 17  | | | | | | | | | | F < 8751.5 : No (1466/0)  | | | | | | | | | | F >= 8751.5 : Yes (314/0)  | | | | | | | | | Z >= 17  | | | | | | | | | | Z < 18.4  | | | | | | | | | | | Z < 18  | | | | | | | | | | | | Z < 17.3 : Yes (1/0)  | | | | | | | | | | | | Z >= 17.3 : No (2/0)  | | | | | | | | | | | Z >= 18  | | | | | | | | | | | | F < 10416.5 : No (1/0)  | | | | | | | | | | | | F >= 10416.5 : Yes (3/0)  | | | | | | | | | | Z >= 18.4 : No (4/0)  | | | | | | | | X >= 164.24 : Yes (353/0)  | | | | | | E >= 2.51  | | | | | | | E < 2.78  | | | | | | | | E < 2.52  | | | | | | | | | Y < 110.51  | | | | | | | | | | X < 110 : No (56/0)  | | | | | | | | | | X >= 110 : Yes (1/0)  | | | | | | | | | Y >= 110.51 : Yes (1/0)  | | | | | | | | E >= 2.52  | | | | | | | | | X < 169.76  | | | | | | | | | | F < 6326.5 : No (1536/0)  | | | | | | | | | | F >= 6326.5 : Yes (237/0)  | | | | | | | | | X >= 169.76 : Yes (260/0)  | | | | | | | E >= 2.78  | | | | | | | | E < 2.78  | | | | | | | | | X < 120.59  | | | | | | | | | | F < 4728.5 : No (8/0)  | | | | | | | | | | F >= 4728.5 : Yes (5/0)  | | | | | | | | | X >= 120.59 : Yes (8/0)  | | | | | | | | E >= 2.78  | | | | | | | | | E < 2.79 : No (16/0)  | | | | | | | | | E >= 2.79  | | | | | | | | | | F < 6326.5  | | | | | | | | | | | E < 2.79  | | | | | | | | | | | | X < 273.47 : No (21/0)  | | | | | | | | | | | | X >= 273.47 : Yes (1/0)  | | | | | | | | | | | E >= 2.79  | | | | | | | | | | | | Y < 111.9  | | | | | | | | | | | | | E < 2.8  | | | | | | | | | | | | | | X < 192.4 : No (40/0)  | | | | | | | | | | | | | | X >= 192.4 : Yes (17/0)  | | | | | | | | | | | | | E >= 2.8  | | | | | | | | | | | | | | X < 177.98 : No (2033/0)  | | | | | | | | | | | | | | X >= 177.98 : Yes (398/0)  | | | | | | | | | | | | Y >= 111.9  | | | | | | | | | | | | | X < 192.03 : No (1090/0)  | | | | | | | | | | | | | X >= 192.03 : Yes (297/0)  | | | | | | | | | | F >= 6326.5 : Yes (638/0)  | | | | | E >= 3.73  | | | | | | F < 6301.5  | | | | | | | X < 161.98 : No (4023/0)  | | | | | | | X >= 161.98 : Yes (1046/0)  | | | | | | F >= 6301.5 : Yes (932/0)  | | | | Y >= 152.85 : Yes (2368/0)  | | | E >= 6.74  | | | | Y < 133.34  | | | | | X < 141.56  | | | | | | E < 43.42  | | | | | | | X < 110.02  | | | | | | | | F < 6300.5 : No (893/0)  | | | | | | | | F >= 6300.5 : Yes (393/0)  | | | | | | | X >= 110.02  | | | | | | | | Y < 117.73  | | | | | | | | | E < 13.72  | | | | | | | | | | Y < 87.02  | | | | | | | | | | | F < 6324 : No (159/0)  | | | | | | | | | | | F >= 6324 : Yes (22/0)  | | | | | | | | | | Y >= 87.02  | | | | | | | | | | | F < 6390 : No (497/0)  | | | | | | | | | | | F >= 6390 : Yes (120/0)  | | | | | | | | | E >= 13.72  | | | | | | | | | | E < 31.25  | | | | | | | | | | | Y < 87.27  | | | | | | | | | | | | Y < 72.9 : No (1/0)  | | | | | | | | | | | | Y >= 72.9  | | | | | | | | | | | | | X < 121.43  | | | | | | | | | | | | | | F < 5388.5 : No (1/0)  | | | | | | | | | | | | | | F >= 5388.5 : Yes (2/0)  | | | | | | | | | | | | | X >= 121.43 : Yes (10/0)  | | | | | | | | | | | Y >= 87.27  | | | | | | | | | | | | Y < 113.91  | | | | | | | | | | | | | F < 6738.5 : No (28/0)  | | | | | | | | | | | | | F >= 6738.5 : Yes (24/0)  | | | | | | | | | | | | Y >= 113.91 : No (4/0)  | | | | | | | | | | E >= 31.25 : Yes (17/0)  | | | | | | | | Y >= 117.73  | | | | | | | | | F < 6547 : No (91/0)  | | | | | | | | | F >= 6547 : Yes (41/0)  | | | | | | E >= 43.42  | | | | | | | E < 905.73  | | | | | | | | X < 90.52  | | | | | | | | | X < 89.93  | | | | | | | | | | Y < 125.47  | | | | | | | | | | | F < 6974.5 : No (58/0)  | | | | | | | | | | | F >= 6974.5 : Yes (3/0)  | | | | | | | | | | Y >= 125.47  | | | | | | | | | | | E < 89.4  | | | | | | | | | | | | F < 7236.5 : No (3/0)  | | | | | | | | | | | | F >= 7236.5 : Yes (2/0)  | | | | | | | | | | | E >= 89.4 : Yes (4/0)  | | | | | | | | | X >= 89.93  | | | | | | | | | | F < 5713 : No (146/0)  | | | | | | | | | | F >= 5713 : Yes (66/0)  | | | | | | | | X >= 90.52  | | | | | | | | | X < 90.78  | | | | | | | | | | X < 90.66  | | | | | | | | | | | E < 51.97  | | | | | | | | | | | | E < 46.13 : No (1/0)  | | | | | | | | | | | | E >= 46.13 : Yes (2/0)  | | | | | | | | | | | E >= 51.97 : No (9/0)  | | | | | | | | | | X >= 90.66 : No (72/0)  | | | | | | | | | X >= 90.78  | | | | | | | | | | Y < 110.24  | | | | | | | | | | | X < 110.52  | | | | | | | | | | | | F < 5400.5 : No (557/0)  | | | | | | | | | | | | F >= 5400.5 : Yes (127/0)  | | | | | | | | | | | X >= 110.52 : No (28/0)  | | | | | | | | | | Y >= 110.24  | | | | | | | | | | | E < 93.15  | | | | | | | | | | | | F < 6611 : No (5/0)  | | | | | | | | | | | | F >= 6611 : Yes (2/0)  | | | | | | | | | | | E >= 93.15 : No (45/0)  | | | | | | | E >= 905.73 : No (54/0)  | | | | | X >= 141.56 : Yes (944/0)  | | | | Y >= 133.34 : Yes (921/0)  | | Z >= 55.65 : Yes (3300/0)  | S >= 150.5  | | F < 6969 : No (2459/0)  | | F >= 6969 : Yes (162/0)  E >= 1001.8 : Yes (3669/0)  Size of the tree : 143  Time taken to build model: 0.15 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0.02 seconds  === Summary ===  Correctly Classified Instances 10931 99.8812 %  Incorrectly Classified Instances 13 0.1188 %  Kappa statistic 0.9976  Mean absolute error 0.0012  Root mean squared error 0.0345  Relative absolute error 0.2379 %  Root relative squared error 6.8965 %  Total Number of Instances 10944  Ignored Class Unknown Instances 10999  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  0.999 0.002 0.998 0.999 0.999 0.998 0.728 0.369 No  0.998 0.001 0.999 0.998 0.999 0.998 0.939 0.740 Yes  Weighted Avg. 0.999 0.001 0.999 0.999 0.999 0.998 0.837 0.561  === Confusion Matrix ===  a b <-- classified as  5286 4 | a = No  9 5645 | b = Yes |
|  | === Run information ===  Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1  Relation: unbalanced-semi supervised-2  Instances: 42543  Attributes: 21  M107  M190  S  M104  G28  F  G1  Z  M109  G21  G90  M82  E  G92  X  Y  M140  M106  M84  M220  Drop  Test mode: split 66.0% train, remainder test  === Classifier model (full training set) ===  RandomTree  ==========  E < 641.19  | F < 6901.5  | | S < 151.5  | | | X < 161.98  | | | | Z < 48.75  | | | | | E < 43.74  | | | | | | Y < 157.07 : No (3517/0)  | | | | | | Y >= 157.07 : Yes (1252/0)  | | | | | E >= 43.74 : No (182/0)  | | | | Z >= 48.75 : Yes (1253/0)  | | | X >= 161.98 : Yes (1261/0)  | | S >= 151.5 : No (136/0)  | F >= 6901.5 : Yes (1196/0)  E >= 641.19 : Yes (1328/0)  Size of the tree : 15  Time taken to build model: 0.02 seconds  === Evaluation on test split ===  Time taken to test model on training split: 0 seconds  === Summary ===  Correctly Classified Instances 3410 99.7659 %  Incorrectly Classified Instances 8 0.2341 %  Kappa statistic 0.995  Mean absolute error 0.0023  Root mean squared error 0.0484  Relative absolute error 0.4979 %  Root relative squared error 9.995 %  Total Number of Instances 3418  Ignored Class Unknown Instances 11047  === Detailed Accuracy By Class ===  TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class  1.000 0.004 0.994 1.000 0.997 0.995 0.674 0.129 No  0.996 0.000 1.000 0.996 0.998 0.995 0.899 0.464 Yes  Weighted Avg. 0.998 0.001 0.998 0.998 0.998 0.995 0.815 0.339  === Confusion Matrix ===  a b <-- classified as  1280 0 | a = No  8 2130 | b = Yes |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |