Scott Tasker

Aim: To evaluate the performance of the proposed intrusion classifiers.

**Background / Intro**

802.11 wireless networks, commonly known as Wi-Fi, have become the default for wireless local area networks. Securing these networks is critical to protect the networks users from malicious acts. Binary classification of incoming data packets (deciding if each data packet is ‘safe’ or ‘malicious’) is a challenging task, as each data packet received by the router has multiple parameters (157 fields), and a malicious signature may consist of only small differences across multiple fields. Machine learning algorithms are an ideal candidate for this application. This is because these algorithms can determine the probability that an unlabeled observation fits into a negative or positive set, based on labelled observations they have previously seen (training data). Labels are assigned based on these probabilities using a threshold for classification (default = 0.5).

To assess the performance of our intrusion detection model, we have utilised the reduced CLS portion of the Aegean Wifi Intrusion Dataset [1]. As this dataset is labelled, we can assess the models predicted classifications against the actual classifications for each observation in the test data. The lower the number of incorrectly classified observations, the higher the classifiers accuracy. For this application, the sensitivity of the classifier (e.g. the ability to correctly identify positive observations) will be given more weight when ranking algorithms than the overall accuracy, as the network must be protected, even at the expense that some ‘safe’ traffic will be labelled false positive and blocked. Another key metric for intrusion detection is the time required to make predictions. As to work in a real network, the algorithm would need to make predictions as traffic is received. A delay would not be a workable solution.

**Methods**

The performance of each model was assessed using the equations in figure 1, and receiver operator characteristic curves.



**Figure 1**. Calculation of binary classifier performance from confusion matrix. [2]

To enable high-throughput testing, a function (model\_evaluator.py) was built using the Python programming language, and the open source libraries: Numpy, Pandas, SciKit Learn and MatPlotLib. The function outputs a report for each binary classification model, and the results were compiled in an excel workbook.

Computation time was assessed using a short Python script to record the time required to train the model, and the time required to make predictions when run on an Apple MacBook Pro, 2.7GHz i5, 8GB DDR3 RAM. Results were recorded in an excel workbook.

**Results**

We evaluated 96 permutations of the intrusion detection classifier using the model\_evaulator function. Figure 2 is an example of the evaluation report returned by this function. The 10 best performing models are shown in Table 1. The full database is included in the appendix.

The 5 models with the highest sensitivity scores were taken forward for further analysis. Table 2 shows the computation time required to train each of these models, and the time required to predict unlabeled data. Table 3 compares these models, with the published benchmarks.



**Figure 2**. Output from model\_evaluator function.



Table 1. Top performing algorithms (10/96). Algorithms ranked according by number of false negatives observed.



Table 2. Analysis of time required to build model, and time required to test model,



Table 3. Comparison of the top 5 performing algorithms from our project to published models [1][2][3]

**Conclusion**

* We identified 2 models, Robust Scaler-RFElog10-Decision Tree, and Robust Scaler FS-RFElog10-Random Forest giving the highest sensitivity (only 3 false negatives of 20079 positive observations).
* Tuning of the Robust Scaler-RFElog10 Decision Tree and Robust Scaler-RFElog10 Random Forest algorithms did not increase the sensitivity or the specificity of the algorithm.
* The Robust Scaler-RFElog10-Decision Tree required the lowest computation to use, as demonstrated in table 2.
* The top performing models had comparable performance to state-of-the-art models [2],[3].
* Future analysis would identify the 3 false-negative observations in the database, to see if the pre-processing, or feature selection could be adjusted to allow correct classification of these samples.

**References**

[1] C. Kolias, G. Kambourakis, A. Stavrou and S. Gritzalis, "Intrusion Detection in 802.11 Networks: Empirical Evaluation of Threats and a Public Dataset", *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 184-208, 2016. Available: 10.1109/comst.2015.2402161 [Accessed 13 January 2020].

[2] M. Aminanto, R. Choi, H. Tanuwidjaja, P. Yoo and K. Kim, "Deep Abstraction and Weighted Feature Selection for Wi-Fi Impersonation Detection", *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 3, pp. 621-636, 2018. Available: 10.1109/tifs.2017.2762828 [Accessed 13 January 2020].

[3] M. Aminanto and K. Kim, "Detecting Impersonation Attack in WiFi Networks Using Deep Learning Approach", *Information Security Applications*, pp. 136-147, 2017. Available: 10.1007/978-3-319-56549-1\_12 [Accessed 13 January 2020].

**Appendix**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Preprocessing | FeatureSelection | Algorithms | Tuning | Accuracy | True Positive | True Negative | False Positive | False Negative | specificity = TN/(TN+FP) | recall (or sensitivity) = TP / (TP + FN) | precision = TP / (TP + FP) | F1score = (2\* True Positive)/((2\*True Positive)+False Positive+False Negative) | False Positives Rate = FP/(FP+TN) | MCC |
| RobustScaler | RFE, logistic classifier, 10 features | DT | default parameters | 0.97614423 | 20076 | 19124 | 955 | 3 | 0.999843154 | 0.99985059 | 0.954590842 | 0.976696667 | 0.04756213 | 0.976951366 |
| RobustScaler | RFE, logistic classifier, 10 features | DT | tuned parameters | 0.97614423 | 20076 | 19124 | 955 | 3 | 0.999843154 | 0.99985059 | 0.954590842 | 0.976696667 | 0.04756213 | 0.976951366 |
| RobustScaler | RFE, logistic classifier, 10 features | RF | default parameters | 0.97614423 | 20076 | 19124 | 955 | 3 | 0.999843154 | 0.99985059 | 0.954590842 | 0.976696667 | 0.04756213 | 0.976951366 |
| RobustScaler | RFE, logistic classifier, 10 features | RF | tuned parameters | 0.97614423 | 20076 | 19124 | 955 | 3 | 0.999843154 | 0.99985059 | 0.954590842 | 0.976696667 | 0.04756213 | 0.976951366 |
| RobustScaler | RFE, logistic classifier, 10 features | KNN | tuned parameters | 0.97552169 | 20076 | 19099 | 980 | 3 | 0.999842948 | 0.99985059 | 0.953457447 | 0.976103075 | 0.048807212 | 0.976371021 |
| RobustScaler | RFE, logistic classifier, 10 features | LR | tuned parameters | 0.97509836 | 20076 | 19082 | 997 | 3 | 0.999842808 | 0.99985059 | 0.952688274 | 0.975699844 | 0.049653867 | 0.975976977 |
| RobustScaler | RFE, logistic classifier, 10 features | LR | default parameters | 0.96269735 | 20076 | 18584 | 1495 | 3 | 0.999838597 | 0.99985059 | 0.930693987 | 0.964033613 | 0.074455899 | 0.964641156 |
| RobustScaler | RFE, logistic classifier, 10 features | SVC | default parameters | 0.96257284 | 20073 | 18582 | 1497 | 6 | 0.999677211 | 0.99970118 | 0.930598053 | 0.963912699 | 0.074555506 | 0.964507733 |
| StandardScaler | RFE, logistic classifier, 5 features | SVC | tuned parameters | 0.983564919 | 20067 | 19431 | 648 | 12 | 0.999382811 | 0.999402361 | 0.96871832 | 0.98382115 | 0.032272524 | 0.983921116 |
| StandardScaler | RFE, logistic classifier, 5 features | SVC | default parameters | 0.978509886 | 20067 | 19228 | 851 | 12 | 0.999376299 | 0.999402361 | 0.959317334 | 0.978949679 | 0.042382589 | 0.979128827 |
| StandardScaler | Decision tree classifier, 10 most important features | LR | default parameters | 0.927411724 | 20064 | 17179 | 2900 | 15 | 0.999127603 | 0.999252951 | 0.873715381 | 0.932277025 | 0.144429503 | 0.934260305 |
| RobustScaler | Decision tree classifier, 10 most important features | KNN | default parameters | 0.98819662 | 20053 | 19631 | 448 | 26 | 0.998677316 | 0.998705115 | 0.978147407 | 0.988319369 | 0.022311868 | 0.988343569 |
| StandardScaler | RFE, logistic classifier, 5 features | LR | tuned parameters | 0.965187509 | 20052 | 18708 | 1371 | 27 | 0.998558847 | 0.998655312 | 0.936003361 | 0.966314876 | 0.068280293 | 0.96672657 |
| StandardScaler | Decision tree classifier, 10 most important features | LR | tuned parameters | 0.929354051 | 20047 | 17274 | 2805 | 32 | 0.99815093 | 0.998406295 | 0.877253632 | 0.933917216 | 0.139698192 | 0.935628964 |
| StandardScaler | RFE, logistic classifier, 10 features | LR | tuned parameters | 0.98055182 | 20032 | 19345 | 734 | 47 | 0.99757632 | 0.997659246 | 0.964653761 | 0.980878933 | 0.036555605 | 0.980930375 |
| StandardScaler | RFE, logistic classifier, 10 features | LR | default parameters | 0.980003984 | 20032 | 19323 | 756 | 47 | 0.997573567 | 0.997659246 | 0.963632865 | 0.980350894 | 0.037651277 | 0.980408459 |
| StandardScaler | RFE, logistic classifier, 5 features | LR | default parameters | 0.960779919 | 20024 | 18559 | 1520 | 55 | 0.997045235 | 0.99726082 | 0.929446714 | 0.962160344 | 0.075700981 | 0.962540291 |
| RobustScaler | Decision tree classifier, 10 most important features | KNN | tuned parameters | 0.98142338 | 19911 | 19501 | 578 | 168 | 0.991458641 | 0.991633049 | 0.971789741 | 0.981611122 | 0.028786294 | 0.981415759 |
| StandardScaler | RFE, logistic classifier, 5 features | NBC | tuned parameters | 0.940659395 | 19698 | 18077 | 2002 | 381 | 0.979358544 | 0.981024951 | 0.907741935 | 0.942961775 | 0.099706161 | 0.941650897 |
| RobustScaler | Decision tree classifier, 10 most important features | NB | default parameters | 0.9558992 | 18613 | 19774 | 305 | 1466 | 0.930979284 | 0.926988396 | 0.983877788 | 0.95458625 | 0.01519 | 0.953849383 |
| RobustScaler | Decision tree classifier, 10 most important features | NB | tuned parameters | 0.9558992 | 18613 | 19774 | 305 | 1466 | 0.930979284 | 0.926988396 | 0.983877788 | 0.95458625 | 0.01519 | 0.953849383 |
| StandardScaler | Decision tree classifier, 10 most important features | NBC | tuned parameters | 0.944842871 | 18613 | 19330 | 749 | 1466 | 0.929505674 | 0.926988396 | 0.96131598 | 0.943840166 | 0.037302655 | 0.9411152 |
| RobustScaler | RFE, logistic classifier, 10 features | NB | tuned parameters | 0.91715225 | 18613 | 18218 | 1861 | 1466 | 0.925523268 | 0.926988396 | 0.90910423 | 0.917959214 | 0.092683899 | 0.910616797 |
| RobustScaler | RFE, logistic classifier, 5 features | RF | default parameters | 0.93821903 | 18609 | 19068 | 1011 | 1470 | 0.928425358 | 0.926789183 | 0.948470948 | 0.937504723 | 0.050351113 | 0.933640556 |
| RobustScaler | RFE, logistic classifier, 5 features | KNN | tuned parameters | 0.93821903 | 18609 | 19068 | 1011 | 1470 | 0.928425358 | 0.926789183 | 0.948470948 | 0.937504723 | 0.050351113 | 0.933640556 |
| RobustScaler | RFE, logistic classifier, 5 features | DT | default parameters | 0.93821903 | 18609 | 19068 | 1011 | 1470 | 0.928425358 | 0.926789183 | 0.948470948 | 0.937504723 | 0.050351113 | 0.933640556 |
| RobustScaler | RFE, logistic classifier, 5 features | DT | tuned parameters | 0.93821903 | 18609 | 19068 | 1011 | 1470 | 0.928425358 | 0.926789183 | 0.948470948 | 0.937504723 | 0.050351113 | 0.933640556 |
| RobustScaler | Decision tree classifier, 10 most important features | SVM | tuned parameters | 0.96331989 | 18606 | 20079 | 0 | 1473 | 0.931653675 | 0.926639773 | 1 | 0.961923226 | 0 | 0.962621303 |
| RobustScaler | Decision tree classifier, 10 most important features | RF | tuned parameters | 0.96331989 | 18606 | 20079 | 0 | 1473 | 0.931653675 | 0.926639773 | 1 | 0.961923226 | 0 | 0.962621303 |
| RobustScaler | Decision tree classifier, 10 most important features | LR | default parameters | 0.95634743 | 18606 | 19799 | 280 | 1473 | 0.930754043 | 0.926639773 | 0.985174203 | 0.955010907 | 0.013944918 | 0.954389107 |
| RobustScaler | Decision tree classifier, 10 most important features | SVM | default parameters | 0.95231336 | 18606 | 19637 | 442 | 1473 | 0.930222643 | 0.926639773 | 0.976795464 | 0.951056815 | 0.022013048 | 0.94969183 |
| RobustScaler | Decision tree classifier, 10 most important features | LR | tuned parameters | 0.95099358 | 18606 | 19584 | 495 | 1473 | 0.930047015 | 0.926639773 | 0.974085126 | 0.949770291 | 0.024652622 | 0.948165213 |
| StandardScaler | RFE, logistic classifier, 10 features | NBC | tuned parameters | 0.940759002 | 18606 | 19173 | 906 | 1473 | 0.928654461 | 0.926639773 | 0.953567036 | 0.939910586 | 0.045121769 | 0.936490423 |
| RobustScaler | RFE, logistic classifier, 5 features | SVC | default parameters | 0.9376961 | 18606 | 19050 | 1029 | 1473 | 0.928226867 | 0.926639773 | 0.947593583 | 0.936999547 | 0.051247572 | 0.933050951 |
| RobustScaler | RFE, logistic classifier, 5 features | LR | default parameters | 0.9376214 | 18606 | 19047 | 1032 | 1473 | 0.928216374 | 0.926639773 | 0.947448824 | 0.936928771 | 0.051396982 | 0.932967365 |
| RobustScaler | RFE, logistic classifier, 5 features | LR | tuned parameters | 0.9376214 | 18606 | 19047 | 1032 | 1473 | 0.928216374 | 0.926639773 | 0.947448824 | 0.936928771 | 0.051396982 | 0.932967365 |
| RobustScaler | RFE, logistic classifier, 10 features | SVC | tuned parameters | 0.85756263 | 15339 | 19099 | 980 | 4740 | 0.801166156 | 0.763932467 | 0.939947301 | 0.842848508 | 0.048807212 | 0.833945818 |
| RobustScaler | RFE, logistic classifier, 5 features | NB | default parameters | 0.8055182 | 15337 | 17011 | 3068 | 4742 | 0.782007079 | 0.76383286 | 0.833306167 | 0.797058518 | 0.152796454 | 0.753325122 |
| RobustScaler | RFE, logistic classifier, 5 features | NB | tuned parameters | 0.79710145 | 15337 | 16673 | 3406 | 4742 | 0.778566425 | 0.76383286 | 0.818278824 | 0.790119005 | 0.169629962 | 0.740652653 |
| RobustScaler | RFE, logistic classifier, 10 features | NB | default parameters | 0.81866627 | 15336 | 17540 | 2539 | 4743 | 0.787147153 | 0.763783057 | 0.857958042 | 0.808136165 | 0.12645052 | 0.77326187 |
| RobustScaler | RFE, logistic classifier, 5 features | SVC | tuned parameters | 0.82025997 | 13872 | 19068 | 1011 | 6207 | 0.754421365 | 0.690871059 | 0.932070147 | 0.79354728 | 0.050351113 | 0.783421334 |
| RobustScaler | RFE, logistic classifier, 5 features | RF | tuned parameters | 0.82025997 | 13872 | 19068 | 1011 | 6207 | 0.754421365 | 0.690871059 | 0.932070147 | 0.79354728 | 0.050351113 | 0.783421334 |
| StandardScaler | PCA, top 5 features | KNN | tuned parameters | 0.335997809 | 1473 | 12020 | 8059 | 18606 | 0.39247698 | 0.073360227 | 0.154532102 | 0.099490054 | 0.40136461 | -0.795236456 |
| StandardScaler | PCA, top 5 features | KNN | default parameters | 0.335524678 | 1472 | 12002 | 8077 | 18607 | 0.392106897 | 0.073310424 | 0.154152267 | 0.099365465 | 0.402261069 | -0.798015436 |
| StandardScaler | PCA, top 5 features | RF | default parameters | 0.262737188 | 1470 | 9081 | 10998 | 18609 | 0.327952329 | 0.073210817 | 0.117901829 | 0.090330906 | 0.547736441 | -1.331498602 |
| StandardScaler | PCA, top 5 features | LR | tuned parameters | 0.112580308 | 1469 | 3052 | 17027 | 18610 | 0.140891884 | 0.073161014 | 0.079422578 | 0.076163318 | 0.848000398 | -5.311306526 |
| StandardScaler | PCA, top 5 features | LR | default parameters | 0.112480701 | 1469 | 3048 | 17031 | 18610 | 0.140733216 | 0.073161014 | 0.079405405 | 0.076155421 | 0.848199612 | -5.319068958 |
| StandardScaler | RFE, logistic classifier, 5 features | DT | tuned parameters | 0.533019573 | 1468 | 19937 | 142 | 18611 | 0.517199336 | 0.073111211 | 0.911801242 | 0.135368159 | 0.007072065 | 0.234877717 |
| StandardScaler | RFE, logistic classifier, 5 features | DT | default parameters | 0.532745655 | 1468 | 19926 | 153 | 18611 | 0.517061525 | 0.073111211 | 0.905613819 | 0.135299539 | 0.007619901 | 0.232265742 |
| RobustScaler | RFE, logistic classifier, 10 features | KNN | default parameters | 0.5364062 | 1467 | 20074 | 5 | 18612 | 0.518895725 | 0.073061407 | 0.996603261 | 0.136142174 | 0.000249016 | 0.268986564 |
| StandardScaler | RFE, logistic classifier, 10 features | DT | tuned parameters | 0.533293491 | 1467 | 19949 | 130 | 18612 | 0.517336169 | 0.073061407 | 0.91859737 | 0.135357077 | 0.006474426 | 0.237645097 |
| StandardScaler | RFE, logistic classifier, 10 features | RF | default parameters | 0.533094278 | 1467 | 19941 | 138 | 18612 | 0.517236013 | 0.073061407 | 0.914018692 | 0.135307139 | 0.006872852 | 0.235728216 |
| StandardScaler | RFE, logistic classifier, 10 features | DT | default parameters | 0.533044474 | 1467 | 19939 | 140 | 18612 | 0.517210967 | 0.073061407 | 0.912881145 | 0.13529466 | 0.006972459 | 0.235250552 |
| StandardScaler | RFE, logistic classifier, 5 features | RF | tuned parameters | 0.532994671 | 1467 | 19937 | 142 | 18612 | 0.517185919 | 0.073061407 | 0.911746426 | 0.135282184 | 0.007072065 | 0.234773507 |
| StandardScaler | RFE, logistic classifier, 10 features | RF | tuned parameters | 0.531600179 | 1467 | 19881 | 198 | 18612 | 0.516483516 | 0.073061407 | 0.881081081 | 0.134933775 | 0.009861049 | 0.221659994 |
| StandardScaler | PCA, top 5 features | DT | tuned parameters | 0.510508491 | 1467 | 19034 | 1045 | 18612 | 0.505604845 | 0.073061407 | 0.583996815 | 0.129874729 | 0.052044425 | 0.06268211 |
| StandardScaler | RFE, logistic classifier, 5 features | RF | default parameters | 0.532895064 | 1464 | 19936 | 143 | 18615 | 0.517133148 | 0.072911998 | 0.911014312 | 0.135017984 | 0.007121869 | 0.234222113 |
| StandardScaler | RFE, logistic classifier, 5 features | KNN | default parameters | 0.52492654 | 1463 | 19617 | 462 | 18616 | 0.513090785 | 0.072862194 | 0.76 | 0.132975823 | 0.023009114 | 0.164799981 |
| StandardScaler | RFE, logistic classifier, 5 features | KNN | tuned parameters | 0.525673589 | 1461 | 19649 | 430 | 18618 | 0.513471137 | 0.072762588 | 0.772607086 | 0.132999545 | 0.021415409 | 0.170979282 |
| StandardScaler | RFE, logistic classifier, 10 features | KNN | default parameters | 0.535011704 | 1444 | 20041 | 38 | 18635 | 0.518176647 | 0.071915932 | 0.974358974 | 0.13394555 | 0.001892525 | 0.258233672 |
| StandardScaler | RFE, logistic classifier, 10 features | KNN | tuned parameters | 0.534762687 | 1444 | 20031 | 48 | 18635 | 0.518052035 | 0.071915932 | 0.967828418 | 0.133883455 | 0.002390557 | 0.255663904 |
| StandardScaler | RFE, logistic classifier, 10 features | SVC | default parameters | 0.521788934 | 1444 | 19510 | 569 | 18635 | 0.511469393 | 0.071915932 | 0.717337308 | 0.130726055 | 0.028338065 | 0.141644572 |
| StandardScaler | PCA, top 5 features | DT | default parameters | 0.47903282 | 1438 | 17799 | 2280 | 18641 | 0.488446762 | 0.071617112 | 0.386767079 | 0.12085557 | 0.113551472 | -0.109934204 |
| StandardScaler | Decision tree classifier, 10 most important features | RF | tuned parameters | 0.49818218 | 1395 | 18611 | 1468 | 18684 | 0.499021317 | 0.069475571 | 0.487251135 | 0.121611019 | 0.073111211 | -0.010387559 |
| StandardScaler | Decision tree classifier, 10 most important features | DT | tuned parameters | 0.51200259 | 1394 | 19167 | 912 | 18685 | 0.506366903 | 0.069425768 | 0.604509974 | 0.124547688 | 0.045420589 | 0.07420519 |
| StandardScaler | Decision tree classifier, 10 most important features | KNN | default parameters | 0.511380049 | 1351 | 19185 | 894 | 18728 | 0.506026956 | 0.067284227 | 0.601781737 | 0.121035657 | 0.04452413 | 0.07123891 |
| StandardScaler | Decision tree classifier, 10 most important features | KNN | tuned parameters | 0.511529459 | 1346 | 19196 | 883 | 18733 | 0.506103509 | 0.067035211 | 0.603858232 | 0.120674198 | 0.043976294 | 0.07239128 |
| StandardScaler | Decision tree classifier, 10 most important features | SVC | default parameters | 0.503013098 | 1332 | 18868 | 1211 | 18747 | 0.501608401 | 0.066337965 | 0.523790798 | 0.117761471 | 0.060311769 | 0.018020107 |
| StandardScaler | Decision tree classifier, 10 most important features | SVC | tuned parameters | 0.504382688 | 1321 | 18934 | 1145 | 18758 | 0.502334713 | 0.065790129 | 0.53568532 | 0.117187847 | 0.057024752 | 0.026524353 |
| RobustScaler | PCA, top 5 features | LR | tuned parameters | 0.28783804 | 666 | 10893 | 9186 | 19413 | 0.359433775 | 0.033168983 | 0.067600487 | 0.044502355 | 0.457492903 | -1.116608497 |
| RobustScaler | PCA, top 5 features | LR | default parameters | 0.287887843 | 662 | 10899 | 9180 | 19417 | 0.359513128 | 0.032969769 | 0.067262751 | 0.044249858 | 0.457194083 | -1.116298501 |
| RobustScaler | Decision tree classifier, 10 most important features | DT | default parameters | 0.48585587 | 459 | 19052 | 1027 | 19620 | 0.492656185 | 0.022859704 | 0.308882907 | 0.042568978 | 0.051147966 | -0.109589614 |
| RobustScaler | Decision tree classifier, 10 most important features | DT | tuned parameters | 0.48585587 | 459 | 19052 | 1027 | 19620 | 0.492656185 | 0.022859704 | 0.308882907 | 0.042568978 | 0.051147966 | -0.109589614 |
| StandardScaler | Decision tree classifier, 10 most important features | DT | default parameters | 0.490412869 | 148 | 19546 | 533 | 19931 | 0.495123743 | 0.007370885 | 0.21732746 | 0.014258189 | 0.026545147 | -0.10695481 |
| RobustScaler | PCA, top 5 features | DT | default parameters | 0.35063997 | 140 | 13941 | 6138 | 19939 | 0.4114817 | 0.006972459 | 0.022300096 | 0.010623364 | 0.305692515 | -0.769436736 |
| RobustScaler | PCA, top 5 features | DT | tuned parameters | 0.37367399 | 127 | 14879 | 5200 | 19952 | 0.42717694 | 0.006325016 | 0.023840811 | 0.009997638 | 0.258977041 | -0.661943328 |
| StandardScaler | Decision tree classifier, 10 most important features | RF | default parameters | 0.495443 | 89 | 19807 | 272 | 19990 | 0.497700832 | 0.004432492 | 0.246537396 | 0.008708415 | 0.013546491 | -0.068904856 |
| StandardScaler | PCA, top 5 features | RF | tuned parameters | 0.456546641 | 78 | 18256 | 1823 | 20001 | 0.477193716 | 0.003884656 | 0.041031036 | 0.007097361 | 0.090791374 | -0.310648875 |
| RobustScaler | PCA, top 5 features | RF | default parameters | 0.49681259 | 12 | 19939 | 140 | 20067 | 0.49840024 | 0.000597639 | 0.078947368 | 0.001186298 | 0.006972459 | -0.073782921 |
| StandardScaler | Decision tree classifier, 10 most important features | NBC | default parameters | 0.499551771 | 7 | 20054 | 25 | 20072 | 0.499775707 | 0.000348623 | 0.21875 | 0.000696136 | 0.001245082 | -0.022483688 |
| RobustScaler | Decision tree classifier, 10 most important features | RF | default parameters | 0.5000249 | 1 | 20079 | 0 | 20078 | 0.500012451 | 4.98033E-05 | 1 | 9.96016E-05 | 0 | 0.007057144 |
| RobustScaler | PCA, top 5 features | RF | tuned parameters | 0.47920713 | 1 | 19243 | 836 | 20078 | 0.489382264 | 4.98033E-05 | 0.001194743 | 9.56206E-05 | 0.04163554 | -0.212530863 |
| RobustScaler | PCA, top 5 features | SVM | default parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| RobustScaler | PCA, top 5 features | SVM | tuned parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| RobustScaler | PCA, top 5 features | NB | default parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| RobustScaler | PCA, top 5 features | NB | tuned parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| RobustScaler | PCA, top 5 features | KNN | default parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| RobustScaler | PCA, top 5 features | KNN | tuned parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| StandardScaler | RFE, logistic classifier, 10 features | NBC | default parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| StandardScaler | RFE, logistic classifier, 5 features | NBC | default parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| RobustScaler | RFE, logistic classifier, 5 features | KNN | default parameters | 0.5 | 0 | 20079 | 0 | 20079 | 0.5 | 0 | #DIV/0! | 0 | 0 | 0 |
| StandardScaler | RFE, logistic classifier, 10 features | SVC | tuned parameters | 0.496563574 | 0 | 19941 | 138 | 20079 | 0.498275862 | 0 | 0 | 0 | 0.006872852 | -0.083476388 |
| StandardScaler | PCA, top 5 features | NBC | default parameters | 0.487673689 | 0 | 19584 | 495 | 20079 | 0.493759927 | 0 | 0 | 0 | 0.024652622 | -0.160980117 |
| StandardScaler | PCA, top 5 features | NBC | tuned parameters | 0.487673689 | 0 | 19584 | 495 | 20079 | 0.493759927 | 0 | 0 | 0 | 0.024652622 | -0.160980117 |
| StandardScaler | PCA, top 5 features | SVC | default parameters | 0.48070123 | 0 | 19304 | 775 | 20079 | 0.490160729 | 0 | 0 | 0 | 0.03859754 | -0.204349972 |
| StandardScaler | PCA, top 5 features | SVC | tuned parameters | 0.446137756 | 0 | 17916 | 2163 | 20079 | 0.471535728 | 0 | 0 | 0 | 0.107724488 | -0.3678394 |