

Machine Learning for Shot Classification on the DIII-D Tokamak

M. Wolff, J. Abbate, R. Conlin, E. Kolemen



Introduction

- Plasma zoology
 - Unclear distinctions
 - ITER mode optimization
- Categorization
 - Human fallibility
 - Machine learning applications



Background: Plasma Modes

- Quiescent high confinement mode (QH mode)
 - Wide pedestal
 - Super-H
 - ITER baseline scenario
 - Unspecified
- Resonant magnetic perturbation (RMP mode)
 - ITER steady state
 - ITER baseline scenario
 - Unspecified
- Negative triangularity mode (Neg-D)
 - Inner wall limited
 - Diverted
- Low confinement mode (L mode)
 - Improved confinement
 - Enhanced D-Alpha
 - High power
 - Standard
 - Inner wall limited
- Edge localized mode (ELM)

Background: Machine Learning Models

- Tree classification
 - Decision Tree
 - Random Forest
 - Extreme Trees
 - Extra Trees
- Other classifiers
 - Gradient Boosting
 - Bagging
 - Gaussian Naive Bayes
 - Quadratic Discriminant Analysis
- Stacking ensemble



Methodology

- Data
 - Hand-classified database
 - Single mode shots
 - Signals: q_0 , q_{95} , β_N , toroid B field strength, upper & lower triangularity, I30 coil current, topology
 - 190 data points
- Model selection
 - Tested 41 models on 80-20 train-test split
 - Eliminated low performance and failed convergence models
 - Tested different ensemble configurations
 - Selected highest performance ensemble
- Class imbalance problem
 - GAN
 - ADASYN
 - SMOTE



Results

True Positive Distribution in Testing Dataset

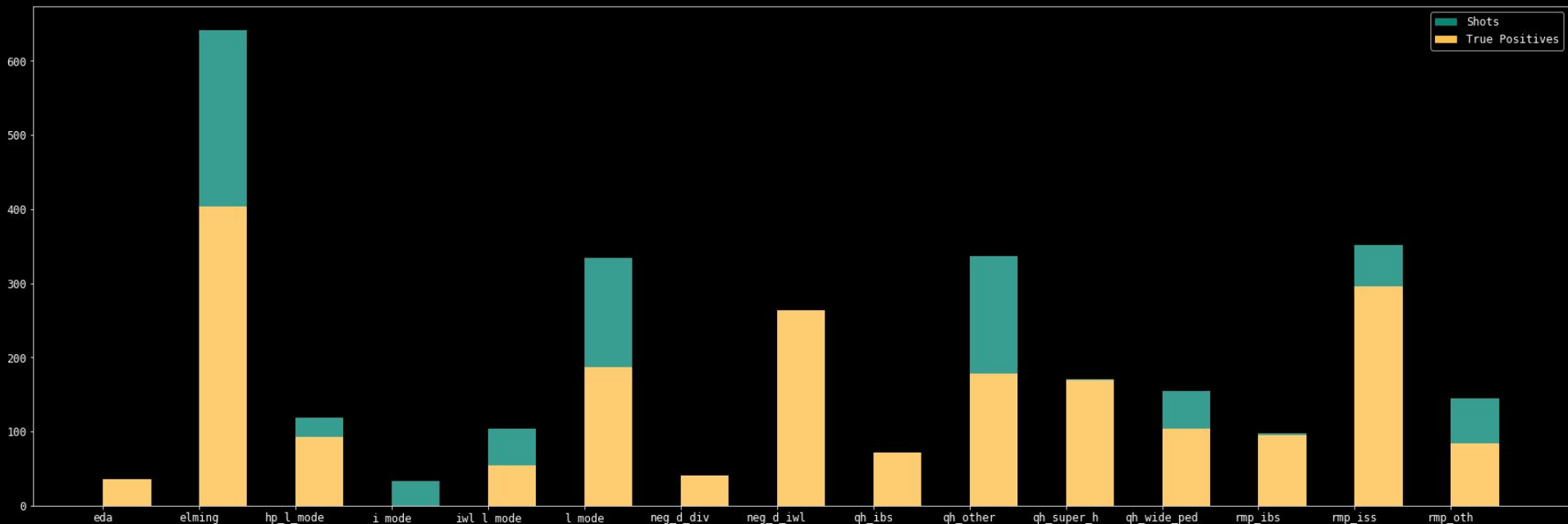


Figure 1.



Results

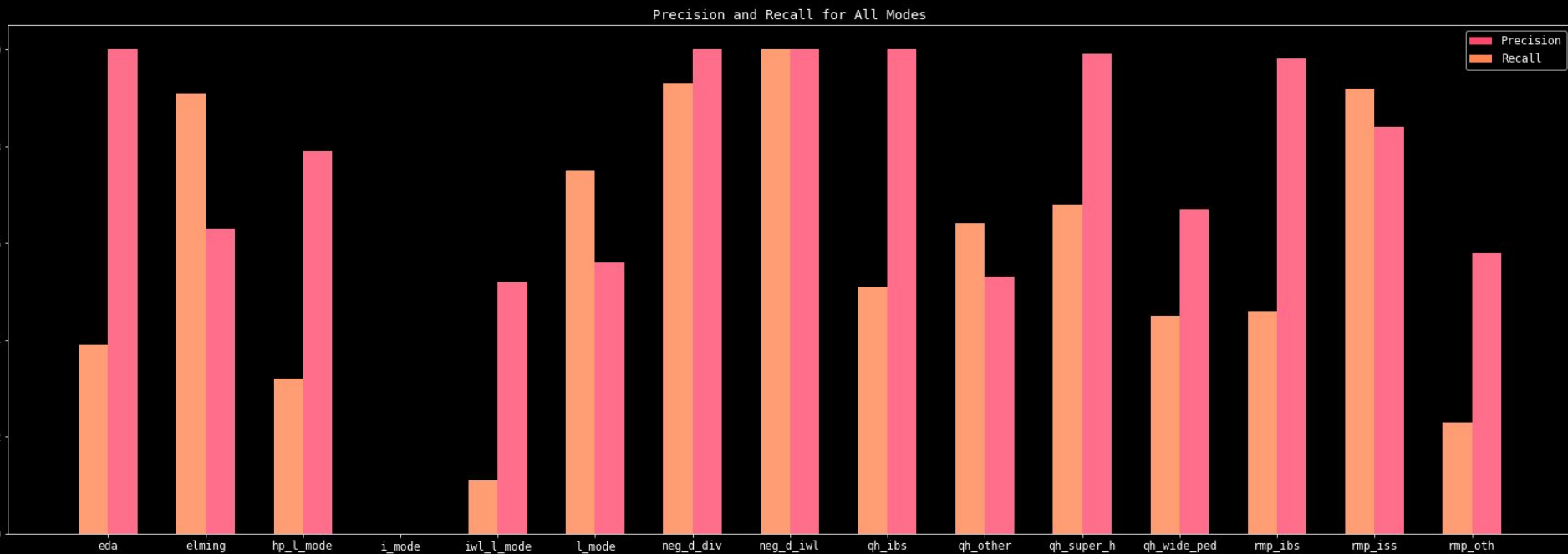


Figure 2.

Results

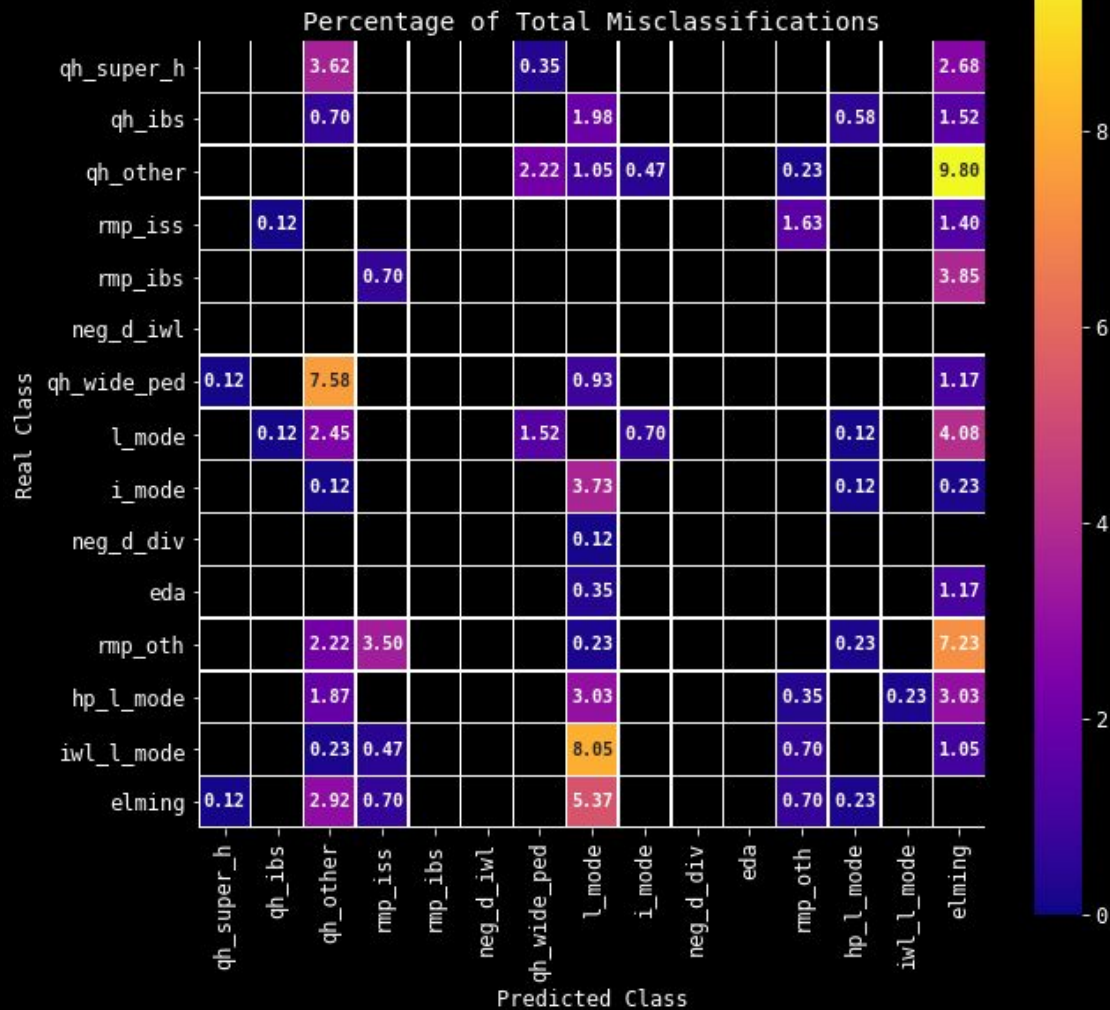


Figure 3.

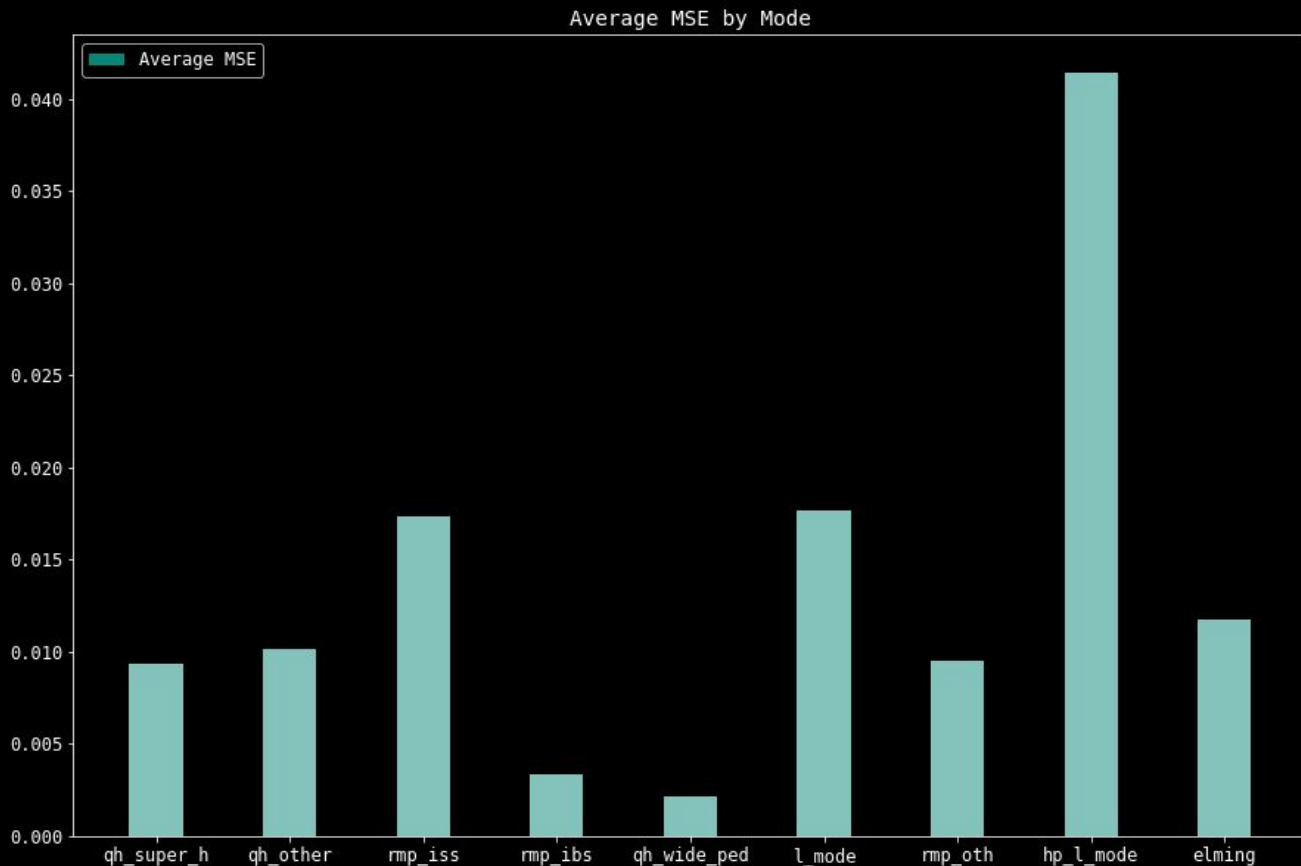


Applications

- Profile prediction validation

| | |
|------------------|-----|
| QH Super H | 3 |
| QH (other) | 26 |
| RMP ISS | 18 |
| RMP IBS | 2 |
| QH Wide Pedestal | 3 |
| L Mode | 62 |
| RMP (other) | 3 |
| HP L mode | 8 |
| ELMing | 212 |

Figure 4.



Discussion & Conclusions

- Implications
 - High recognition of some modes
 - Overidentification of ELMing modes
 - Confusion in L modes in QH modes
- Future research
 - Expand database
 - Include more features
 - Refine hyperparameters





Questions?

References

Picture credit:

<https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.pngguru.com%2Ffree-transparent-background-png-clipart-emdt-e&psig=AOvVaw3PqRBiJFXCOTDJTnXj9wqo&ust=1596661652244000&source=images&cd=vfe&ved=0CAMQjB1qFwoTKjlk7i6gusCFQAAAAAdAAAAABAE>

References

- [1] Edge stability and transport control with resonant magnetic perturbations in collisionless tokamak plasmas.
- [2] Quadratic discriminant analysis.
- [3] Decision tree classifier, 2019.
- [4] Decision trees, 2019.
- [5] Extra trees classifier, 2019.
- [6] sklearn.tree.extratreeclassifier, 2019.
- [7] S. Ide C. Kessel T.C. Luce J.A. Snipes J.K. Stober A.C.C. Sips, G. Giruzzi and the Integrated Operation Scenario Topical Group of the ITPA. Progress in preparing scenarios for iter operation.
- [8] Robert ARNOUX. How fritz wagner "discovered" the h-mode. *ITER NEWSLINE*, (86), June 2009.
- [9] L. Breiman. Bagging predictors. *Machine Learning*, 2(24):123-140, 1996.
- [10] K. H. Burrell, M. E. Austin, D. P. Brennan, J. C. DeBoo, E. J. Doyle, C. Fenzl, C. Fuchs, P. Ghil, C. M. Greenfield, R. J. Groebner, L. L. Lao, T. C. Luce, M. A. Makowski, G. R. McKee, R. A. Moyer, C. C. Petty, M. Porkolab, C. L. Rettig, T. L. Rhodes, J. C. Rost, B. W. Stallard, P. J. Strait, E. J. Synakowski, M. R. Wade, J. G. Watkins, and W. P. West. Quiescent double barrier high-confinement mode plasmas in the diti-d tokamak. *Physics of Plasmas*, 8(5):2153-2162, 2001.
- [11] J.W. Hughes B. Lipschultz J.E. Rice E.S. Marmar M. Greenwald I. Czigler A. Dominguez-F. Gollinopoulos N. Howard I. Lin R.M. McDermott M. Porkolab M.L. Reinke J. Terry N. Tsujii S. Wolfe S. Wakitch Y. Lin D.G. Whyte, A.E. Hubbard and the Alcator C-Mod Team.
- [12] T. E. Evans, K. H. Burrell, M. E. Fenstermacher, R. A. Moyer, T. H. Osborne, M. J. Schaffer, W. P. West, L. W. Yan, J. A. Boedo, E. J. Doyle, G. L. Jackson, I. Joseph, C. J. Lasnier, A. W. Leonard, T. L. Rhodes, P. R. Thomas, J. G. Watkins, and L. Zeng. The physics of edge resonant magnetic perturbations in hot tokamak plasmas. *Physics of Plasmas*, 13(5):056121, 2006.
- [13] Renuka Joshi. Accuracy, precision, recall f1 score: Interpretation of performance measures, September 2016.
- [14] R. Ohlen L. Breiman, J. Friedman and C. Stone. Classification and regression trees.
- [15] Hui Li, Chi Wang, Cui Tu, and Fei Xu. Machine learning approach for solar wind categorisation. *Earth and Space Science*, 7(5), 05 2020.
- [16] Behnaz; Fuqua Joshua D; de Carlo Alexandra R; Bartholomai James A; et al. Lynch, Chip M; Abdollahi. Prediction of lung cancer patient survival via supervised machine learning classification techniques. *International journal of medical informatics*, 108:1-8.
- [17] Earl S. Marmar, Réjean L. Boivin, Catherine L. Fiore, J. A. Goetz, Robert S. Granetz, M. J. Greenwald, Alan E. Hubbard, Jerry Hughes, I. H. Hutchinson, J. H. Irby, B. Bombard, Yi bing Lin, Bruce Lipschultz, Anton Mazurenko, Dmitri Moessesian, Thomas Sann Pedersen, Miklos Porkolab, John R. Rice, Govert Schilling, and Joseph A. Snipes. Enhanced d-alpha h-mode studies in the alcator c-mod tokamak. 2000.
- [18] F. Matos, V. Menkovski, F. Felici, A. Pau, and F. Jenko. Classification of tokamak plasma confinement states with convolutional recurrent neural networks, Nov 11 2019.
- [19] Milta K. Ghosh M.A. Hossain Nazia Hameed, Antesar M.Shahut. Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques.
- [20] Vyacheslav Oshchepkov, Yuri V. Khotyaintsev, Andrey Divin, Gian L. Delzanno, Soren Andersen, Pawel Herman, Steven W. D. Chien, Levon Avakov, and Stefano Markidis. Automated classification of plasma regions using 3d particle energy distribution, Aug 27 2019.
- [21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825-2830, 2011.
- [22] SciKit Learn. *Bagging Classifier*, 2019.
- [23] SciKit Learn. *Ensemble Methods: Bagging Meta-Estimator*, 2019.
- [24] SciKit Learn. *Gradient Boosting Classifier*, 2019.
- [25] V. Bobkov M. Cavedon M. Dunne R.M. McDermott H. Meyer R. Nazikian C. Paz-Soldan D.A. Ryan E. Viezzer M. Willensdorfer The ASDEX Upgrade W. Suttrop, A. Kirk and MST Teams. Experimental conditions to suppress edge localised modes by magnetic perturbations in the asdex upgrade tokamak. 58(9), 2018.
- [26] F. Wagner. A quarter-century of h-mode studies. *Plasma Physics and Controlled Fusion*, 49(12B).
- [27] T.H. Osborne W.M. Solomon K. Barada A.M. Garofalo R.J. Groebner N.C. Luhmann G.R. McKee C.M. Muscatello M. Ono C.C. Petty M. Porkolab T.L. Rhodes J.C. Rost P.B. Snyder G.M. Staebler B.J. Tobias Z. Yan Xi Chen, K.H. Burrell and the DIII-D Team. Stationary qh-mode plasmas with high and wide pedestal at low rotation on diti-d. *Nuclear Fusion*, 57(2), 2017.





Thank you for listening!

This work was made possible by funding from the Department of Energy for the Summer Undergraduate Laboratory Internship (SULI) program. This work is supported by the US DOE Contract No. DE-AC02-09CH11466.

Selection Process: Model with Topology Features

- Model and accuracy score

```
ExtraTreeClassifier() 0.5517241379310345
DecisionTreeClassifier() 0.6275862068965516
RadiusNeighborsClassifier() nan
KNeighborsClassifier() 0.4206896551724138
ClassifierChain(base_estimator=RandomForestClassifier()) nan
OutputCodeClassifier(estimator=RandomForestClassifier()) 0.7241379310344828
OneVsOneClassifier(estimator=RandomForestClassifier()) 0.7206896551724138
OneVsRestClassifier(estimator=RandomForestClassifier()) 0.7344827586206897
SGDClassifier() 0.06206896551724138
RidgeClassifierCV(alphas=array([ 0.1, 1. , 10. ])) 0.46551724137931033
RidgeClassifier() 0.4913793103448276
PassiveAggressiveClassifier() 0.1172413793103448
AdaBoostClassifier() 0.18226600985221678
GradientBoostingClassifier() 0.603448275862069
BaggingClassifier() 0.6655172413793103
ExtraTreesClassifier() 0.7931034482758621
RandomForestClassifier() 0.7758620689655172
BernoulliNB() 0.23793103448275868
CalibratedClassifierCV() 0.45517241379310347
GaussianNB() 0.6482758620689655
LabelPropagation() 0.1586206896551724
LabelSpreading() 0.16896551724137931
LinearDiscriminantAnalysis() 0.4827586206896552
LinearSVC() 0.2
LogisticRegression() 0.3931034482758621
LogisticRegressionCV() 0.4525862068965517
MultinomialNB() nan
NearestCentroid() 0.3275862068965517
NuSVC() 0.3448275862068966
Perceptron() 0.07931034482758621
QuadraticDiscriminantAnalysis() 0.013793103448275862
GaussianMixture() 0.0
```