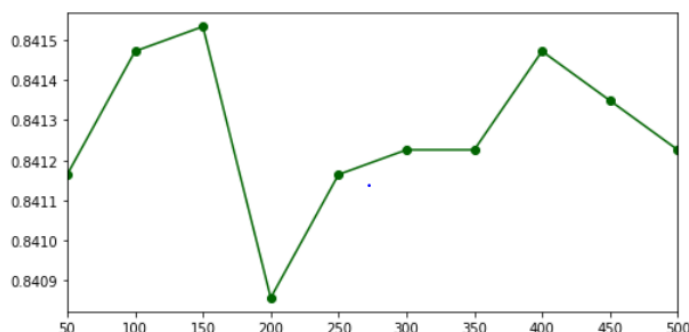


Elisabeth Webb
CA04 Ensemble Models
Machine Learning

Answers to CA04 Questions

Random Forest



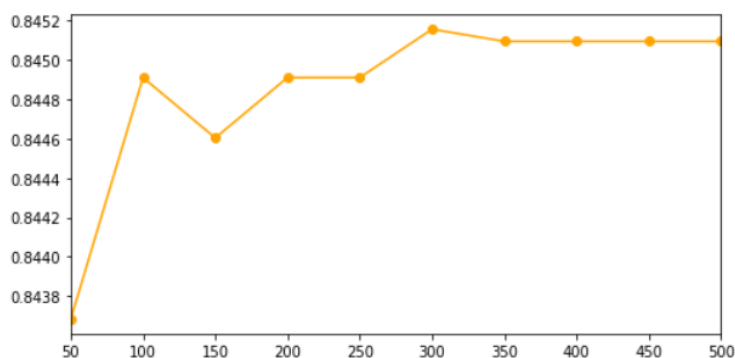
Write your observations about the Classifier's behavior with respect to the number of estimators.

The performance of this model with different number of estimators varies by going up and down. The trend is that it increases then decreases and increases then decreases. The random forest model has the highest accuracy using 150 estimators. It also has a high accuracy with both 100 and 400 estimators. The model has the lowest accuracy using 200 estimators.

Is there an optimal value of the estimator within the given range?

Yes, the model performs the best using 150 estimators. It also performs well with the 100 to 150 range as well as with 400 estimators.

AdaBoost



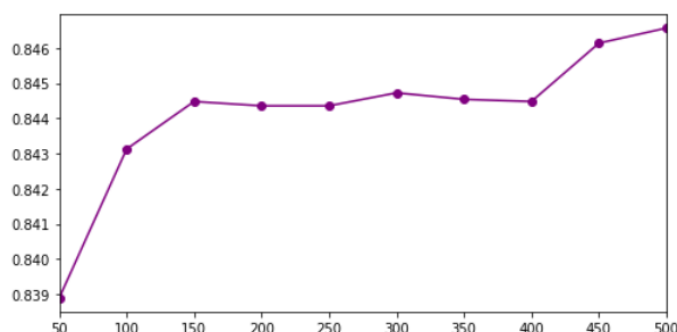
Write your observations about the Classifier's behavior with respect to the number of estimators.

The AdaBoost model tends to increase the more estimators used. It increases with 100 estimators, shows a small decrease with 150 estimators, and then continues to increase as the estimators used gets larger.

Is there an optimal value of the estimator within the given range?

The AdaBoost model reaches its optimal value within the 300 to 500 range of estimators used. The model has the highest accuracy using 300 estimators.

Gradient Boost



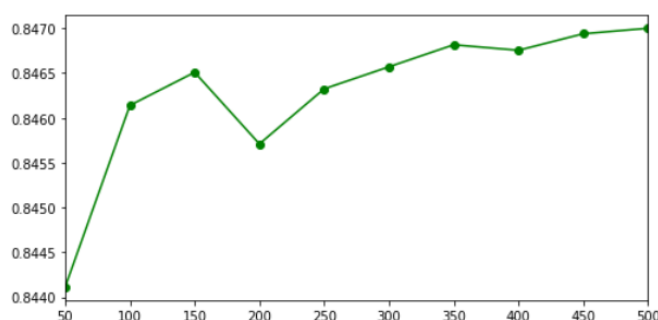
Write your observations about the Classifier's behavior with respect to the number of estimators.

The gradient boost improves the greater the estimators used. The model performs the worst using the least number of estimators, 50, and the best using the greatest number of estimators, 500.

Is there an optimal value of the estimator within the given range?

The optimal value for the gradient boost model is between 450 and 500 estimators. Specifically, the model performs the best using 500 estimators.

XBG



Write your observations about the Classifier's behavior with respect to the number of estimators.

This classifier tends to increase the more estimators used. It shows a decrease in accuracy at 200 estimators, and then overall continues to improve as the number of estimators increases.

Is there an optimal value of the estimator within the given range?

The XGB model reaches its optimal performance in the range of 350 to 500 estimators. This model achieves the highest accuracy using 500 estimators.