BSAN 6070 Machine Learning

Rongxing (Vincent) Chen

CA04 Random Forest and other ensemble classifications

For each model:

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

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

Random Forest

The accuracy is increasing while the number of trees is increasing, but could not be improved while the accuracy has reach 250 and show some ups and downs.

The estimators 250 or 400 show greatest results which are at the middle of the range, thus these estimators are believed to be the best in this range of estimators.

AdaBoost

It is very clear that as n\_estimators is allowed to be larger, the accuracy is increasing significant However, the accuracy is not increasing for the model after the n\_estimator has reached to 200. Optimal estimator shall be over 200.

Gradient Boost

Like the AdaBoost model, the accuracy is increasing as the n\_estimators gets larger, but shows some vibration after the n\_estimator is greater than 100. Optimal estimator shall be 250 or 450.

XGB

Like the previous models, XGB Classifier is mainly showing an increasing trend within the range, but the accuracy is decreased when the estimator is greater than 150. Optimal n\_estimator shall be 100 to 200.

