Alena Sanchez

Professor Brahma

BSAN 6070

March 14, 2022

CA-04: Ensemble Models

Part 3:

1. Random Forest
   1. Write your observations about the Classifier’s behavior with respect to the   
      number of estimators.
      1. From the graphic below we can see how the accuracy of the various RandomForestClassifier Models differ when the number of estimators changes. As we can see there is no clear trend, but we can clearly see that with 250 estimators the model has the lowest accuracy. Between 150 and 250 the accuracy decreases and anything over 350 the accuracy seems to improve as the number of estimators increase.

Chart, line chart

Description automatically generated

* 1. Is there an optimal value of the estimator within the given range?
     1. With the given range, the optimal value of estimators is 450, here we can see that this model has the highest accuracy!

1. AdaBoost
   1. Write your observations about the Classifier’s behavior with respect to the   
      number of estimators.
      1. For the AdaBoost algorithm, the accuracy seems to stagnate once the number of estimators hits about 300. Between 50-150 the accuracy drastically increases as the number of estimators increases.

Chart, line chart

Description automatically generated

* 1. Is there an optimal value of the estimator within the given range?
     1. With the given range, the optimal value of estimators is 150, here we can see that this model has the highest accuracy!

1. Gradient Boost (Classifier)
   1. Write your observations about the Classifier’s behavior with respect to the   
      number of estimators.
      1. With the GradientBoostingClassifier, we can see that the pattern is somewhat like that of the AdaBoost algorithm, between 50 and 250 the accuracy seems to improve as the number of estimators increases. Between 250 and 500 estimators the accuracy seems to decrease gradually and slowly.

Chart, line chart

Description automatically generated

* 1. Is there an optimal value of the estimator within the given range?
     1. With the given range, the optimal value of estimators is 250, here we can see that this model has the highest accuracy as it continuously increasing from 0-250 and once it hits 250 the accuracy only decreases.

1. XGB Model
   1. Write your observations about the Classifier’s behavior with respect to the   
      number of estimators.
      1. With the XGB model, we can say that the accuracy generally increased as the number of estimators increased. Throughout the range between 50 and 500 there were only small dips in the accuracy such as at 250 estimators, but overall, the accuracy increases as number of estimators increases.

Chart, line chart

Description automatically generated

* 1. Is there an optimal value of the estimator within the given range?
     1. For the XGB model, the optimal number of estimators in this range is 400!