Module-5

Antara Paul, Labani Roy

Activity 1: Decision tree classifier and Adaboost Algorithm

A decision tree creates an upside down tree to make predictions of the input data or to classify the input data points.

- It starts with all training example and calculates Gini impurity for the whole training dataset.

$$G(S) = 1 - \sum_i (p_i)^2 \tag{1}$$

  where p_i is the probability. A feature/attribute with minimum gini index means a feature/attribute with maximum information. So, it looks for the feature with the minimum Gini index for the whole training dataset and assigns it to the root node.

- To get the child nodes, it will now calculate the Gini index for other features. It will continue until it meets the stopping criteria (maximum depth, leaf nodes, etc).

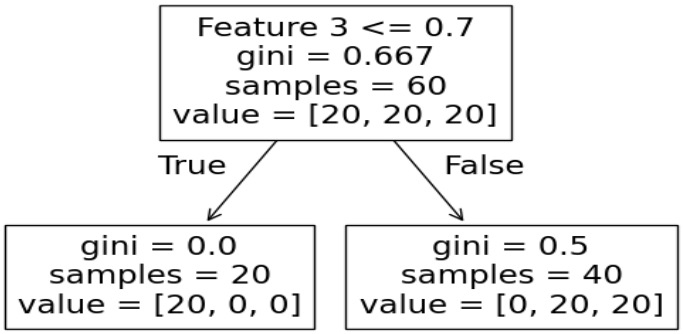


Figure 1: Decision tree (number of maximum leaf nodes = 2)

Adaboost is a machine learning algorithm that takes decision trees as estimators for classification.

- It takes the original training data and assigns the same weights to all the training samples. If N is the total number of training data then it will assign $1/N$ weights to each training sample. Weights will add up to 1.

- Then, this weighted training data is fed into a stump or a decision tree of depth 1 (we call it weak learner).

- Then, we will preserve the correctly classified points and evaluate the misclassified data points from the previous model's classification. The algorithm will try to give more weight to the wrongly classified data points than those that are correctly classified.
- Now, these new weighted misclassified data points will be fed again into a stump. This model will again give us classes that are classified correctly and incorrectly. We will again assign more weights to the misclassified points and feed it to another decision tree/stump.
- the above step continues. At the end of iteration, it will come up with a weighted sum. Naturally, some of the algorithms will be better than the other in classifying the data points. So, with the weighted some, we will give more weight to those better algorithms than the others.

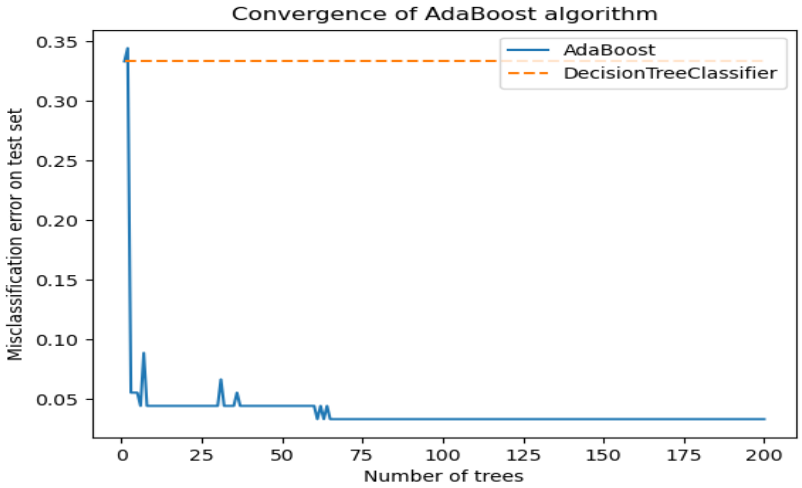


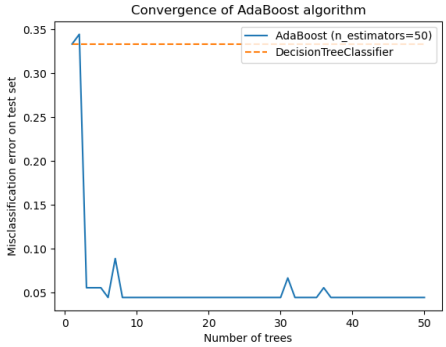
Figure 2

Fig(2) shows how Adaboost reduces the number of misclassifications of the Decision Tree classifier. As Adaboost is a machine learning algorithm, it tries to update weights on misclassified data points with single decision trees at each iteration, It gives us better convergence results than the Decision tree, which just gives us a point.

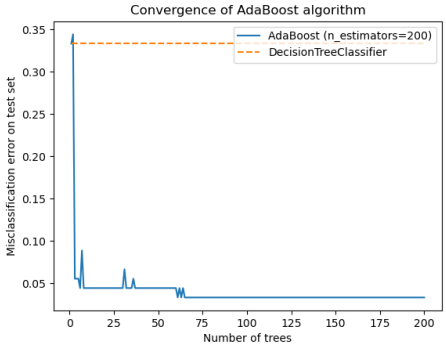
As we are increasing the number of estimators (Fig(3a), Fig(3b)), misclassified error on test dataset with Adaboost algorithm converges to more lower value.

If the learning rate is very low, Adaboost will not be able to classify datapoints properly. As learning rate is related to the weight update at each step, it will not be able to put more weights to the misclassified datapoints from each tree and the result would be similar to what one decision tree will give us (Fig(4a)). As the learning rate increases, Adaboost algorithm works well and provide lower value of misclassified error (Fig(4b)).

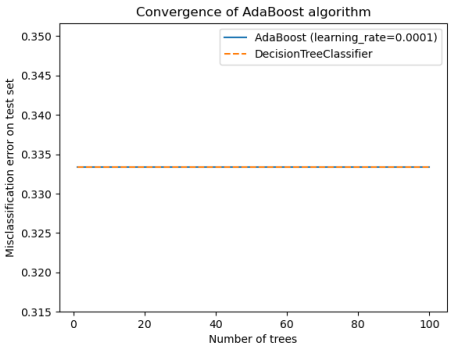
In Fig. 5, shown are the classified data of the three flowers trained on a small sized weak learner of 10 estimators and on an optimal weak learner of 200 estimators.


Convergence of AdaBoost algorithm
AdaBoost (n_estimators=50)
DecisionTreeClassifier
Misclassification error on test set
Number of trees

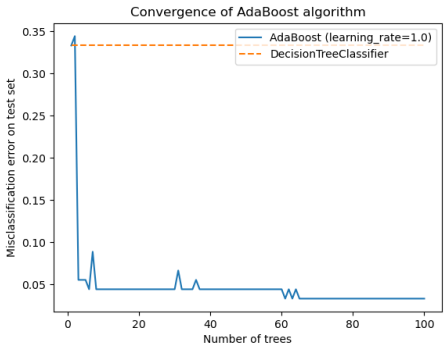
(a)


Convergence of AdaBoost algorithm
AdaBoost (n_estimators=200)
DecisionTreeClassifier
Misclassification error on test set
Number of trees

(b)


Convergence of AdaBoost algorithm
AdaBoost (learning_rate=0.0001)
DecisionTreeClassifier
Misclassification error on test set
Number of trees

(a)


Convergence of AdaBoost algorithm
AdaBoost (learning_rate=1.0)
DecisionTreeClassifier
Misclassification error on test set
Number of trees

(b)

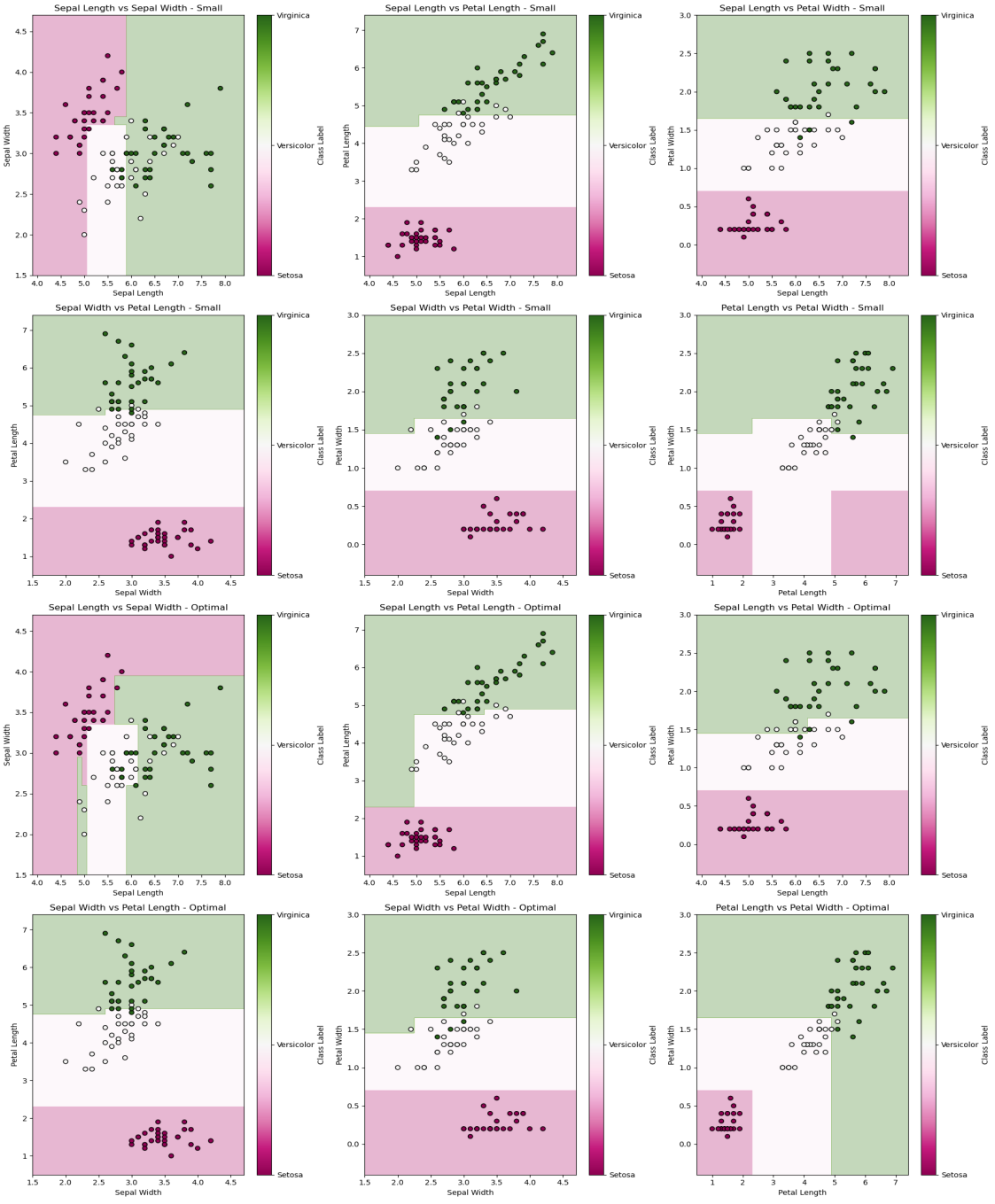

Sepal Length vs Sepal Width - Small
Sepal Length vs Petal Length - Small
Sepal Length vs Petal Width - Small
Sepal Width vs Petal Length - Small
Sepal Width vs Petal Width - Small
Petal Length vs Petal Width - Small
Sepal Length vs Sepal Width - Optimal
Sepal Length vs Petal Length - Optimal
Sepal Length vs Petal Width - Optimal
Sepal Width vs Petal Length - Optimal
Sepal Width vs Petal Width - Optimal
Petal Length vs Petal Width - Optimal
Sepal Length
Sepal Width
Petal Length
Petal Width
Class Label
Virginica
Versicolor
Setosa

Figure 5