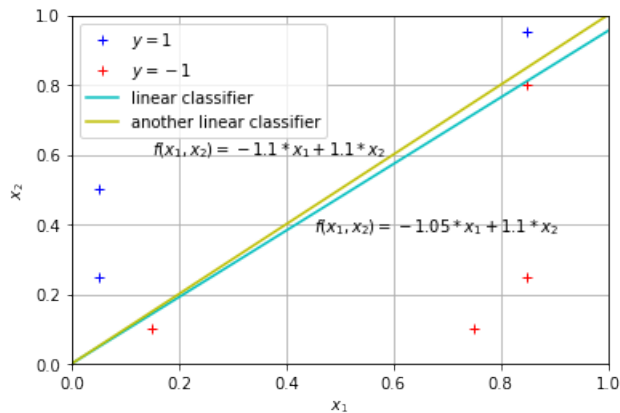# Compsci 571 HW2

Yilin Gao (yg95)
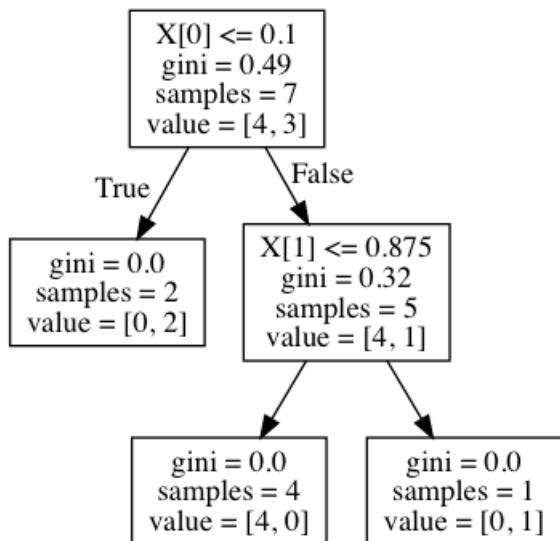
January 26, 2018

1   Classifier for Basketball Courts

(a) When running Perceptron algorithm on the dataset, it takes 3 epochs to converge. The decision boundary is $f(x_1, x_2) = -1.05 * x_2 + 1.1 * x_2$. After it converges, the error rate is 0.

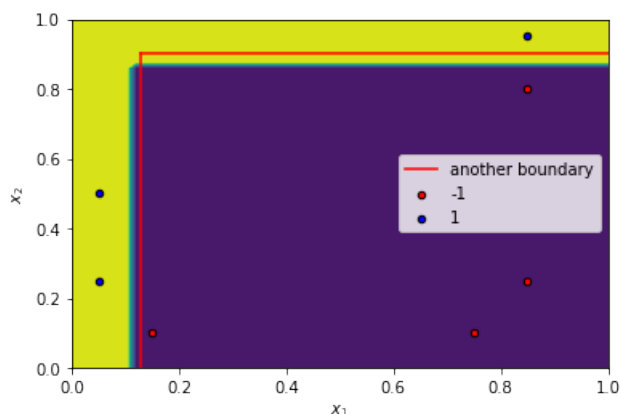The plot of observed data, the decision boundary is:

In the plot there is another linear boundary that goes through the origin to separate the observed data with error $= 0$.

(b) The fully-grown decision tree using Gini index as splitting criterion on the observed data is:

The error rate is also 0.

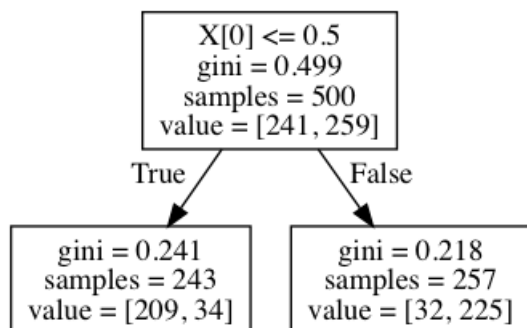The plot of observed data, the decision boundary is:

In the plot there is another decision boundary (in red line) generating the same error rate (0).

(c)

(d)

(e)

(f)

(g)

(h)

2   Variable Importance for Trees and Random Forest

(a)   (i) The decision stump based on the best split is:

The decision stump based on the best surrogate split is:

(ii)

(iii) The mean least-squares error of predictions on the test data of the decision stump based on the best split is 10.

The mean least-squares error of predictions on the test data of the decision stump based on the best split is

(b)   (i)

(ii)

(iii)

(c)   (i)

(ii)