

Homework 11: Decision Tree and Random Forest

Prof. Yao Wang

1. Figure 1 shows a partition of the feature space with two features each in the range of 0 to 1. Draw a decision tree that can realize this partition.

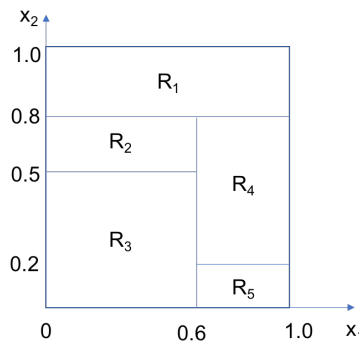


Figure 1: The set of samples to be clustered

2. Figure 2 shows a set of training samples for a binary classification problem. Design a decision tree using the misclassification rate as the loss function. Note that given this simple problem, you should be able to “eyeball” the correct partition threshold for each possible splitting variable, instead of using exhaustive search. That is, for each split, consider both features (x_1 and x_2), but for each feature you should be able to eyeball the proper splitting threshold.

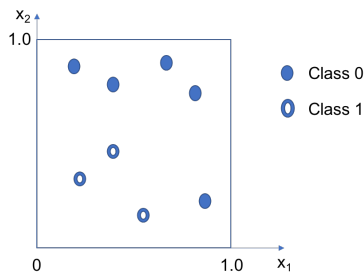


Figure 2: The set of samples to be clustered