

5. Suppose we produce ten bootstrapped samples from a data set containing red and green classes. We then apply a classification tree to each bootstrapped sample and, for a specific value of X , produce 10 estimates of $P(\text{Class is Red}|X)$:

0.1, 0.15, 0.2, 0.2, 0.55, 0.6, 0.6, 0.65, 0.7, and 0.75.

There are two common ways to combine these results together into a single class prediction. One is the majority vote approach discussed in this chapter. The second approach is to classify based on the average probability. In this example, what is the final classification under each of these two approaches?

Answer

① Majority vote approach

- Not Red
- Not Red
- Not Red
- Not Red

} 4

• Red

• Red

• Red

• Red

• Red

• Red

} 6

Majority Vote



Class is Red



② Average Probability

$$\begin{aligned} \text{Avg. Prob} = & 0.1 + 0.15 + 0.2 + 0.2 + 0.55 \\ & + 0.6 + 0.6 + 0.65 + 0.7 + 0.75 \\ & \hline & 10 \end{aligned}$$

$$\therefore \text{Avg. Prob} = \frac{4.5}{10} = 0.45 < 0.5$$

\therefore Prediction: Class is Not Red

