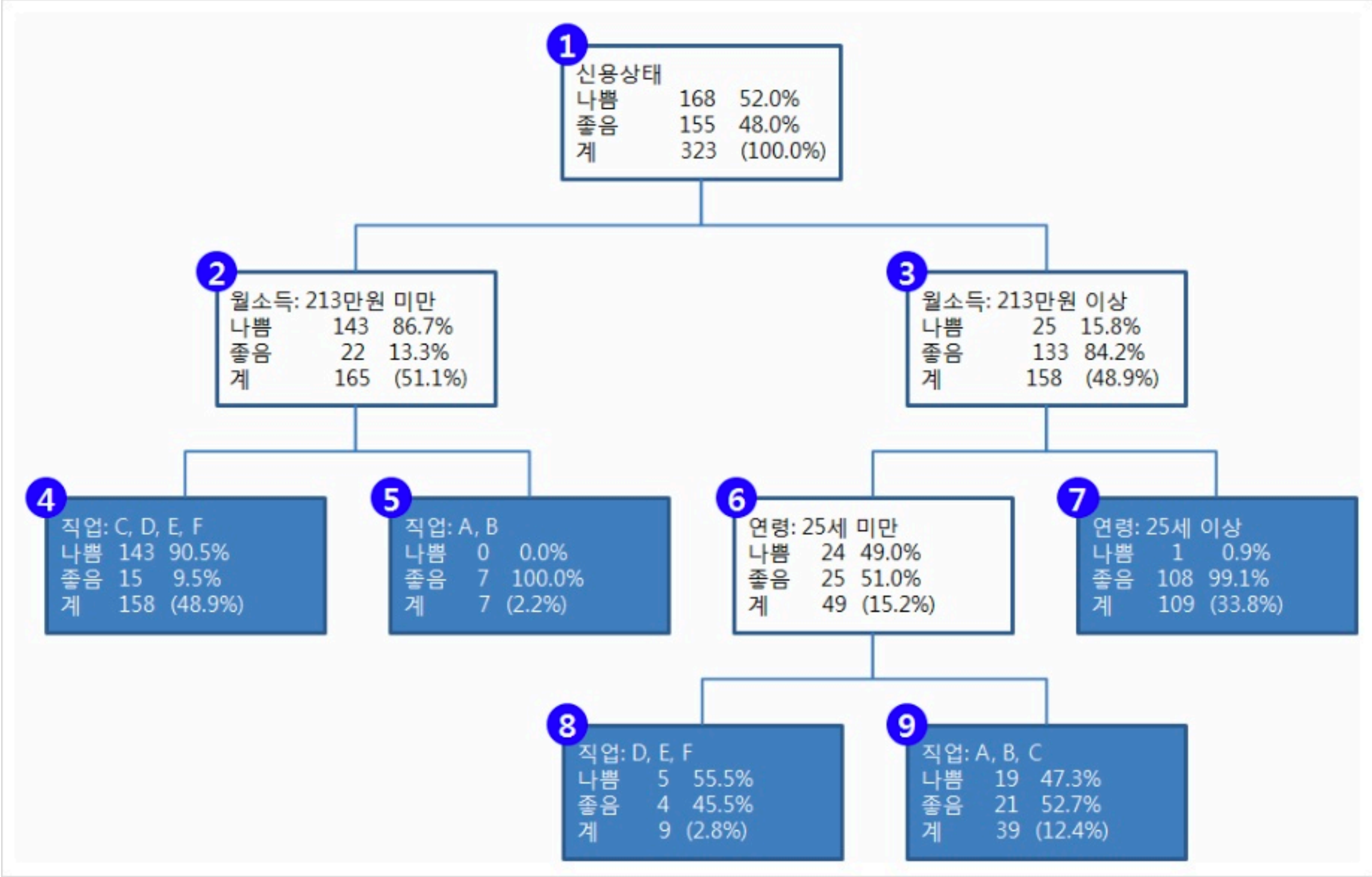
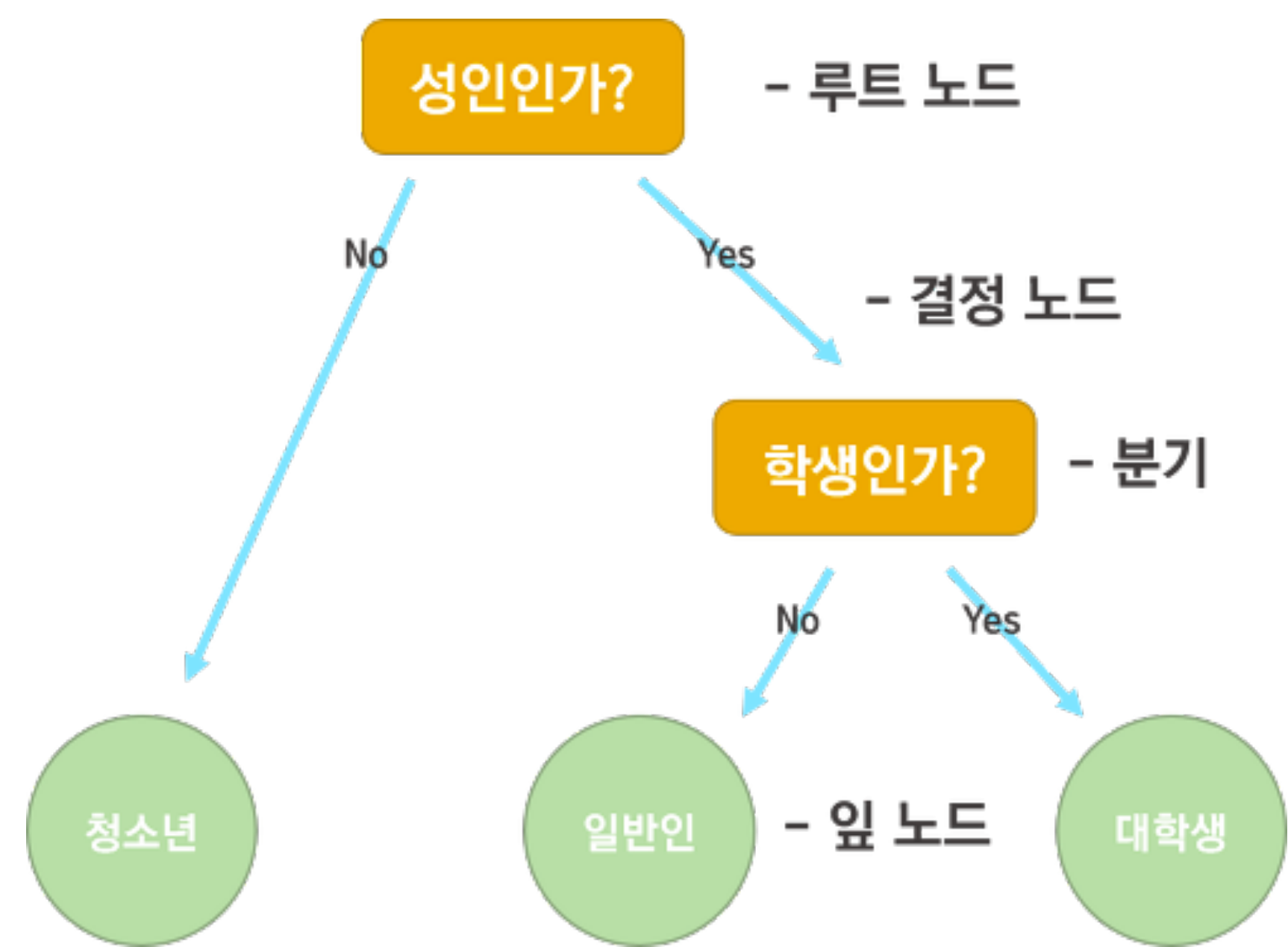
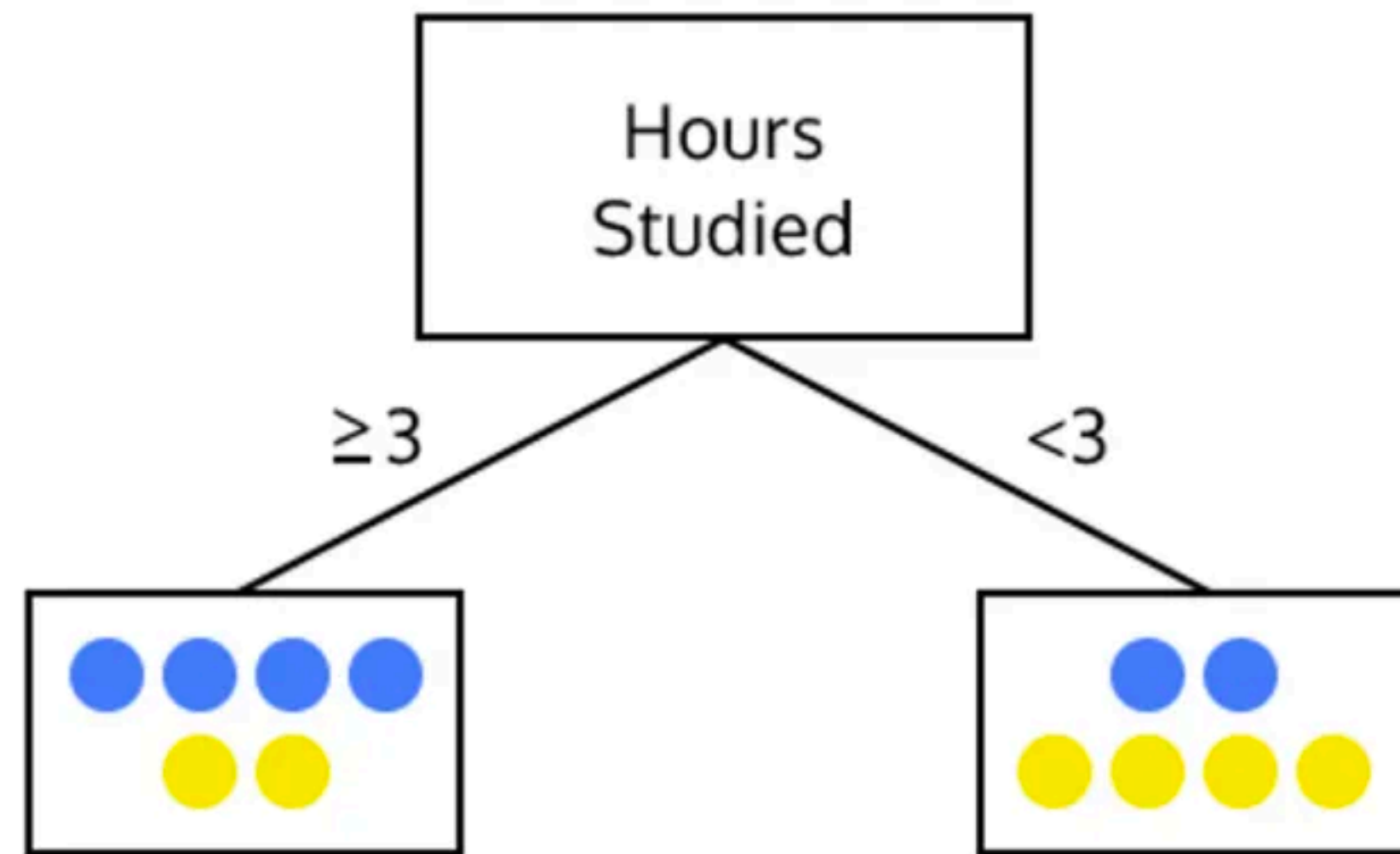


# 의사결정나무(Decision Tree)

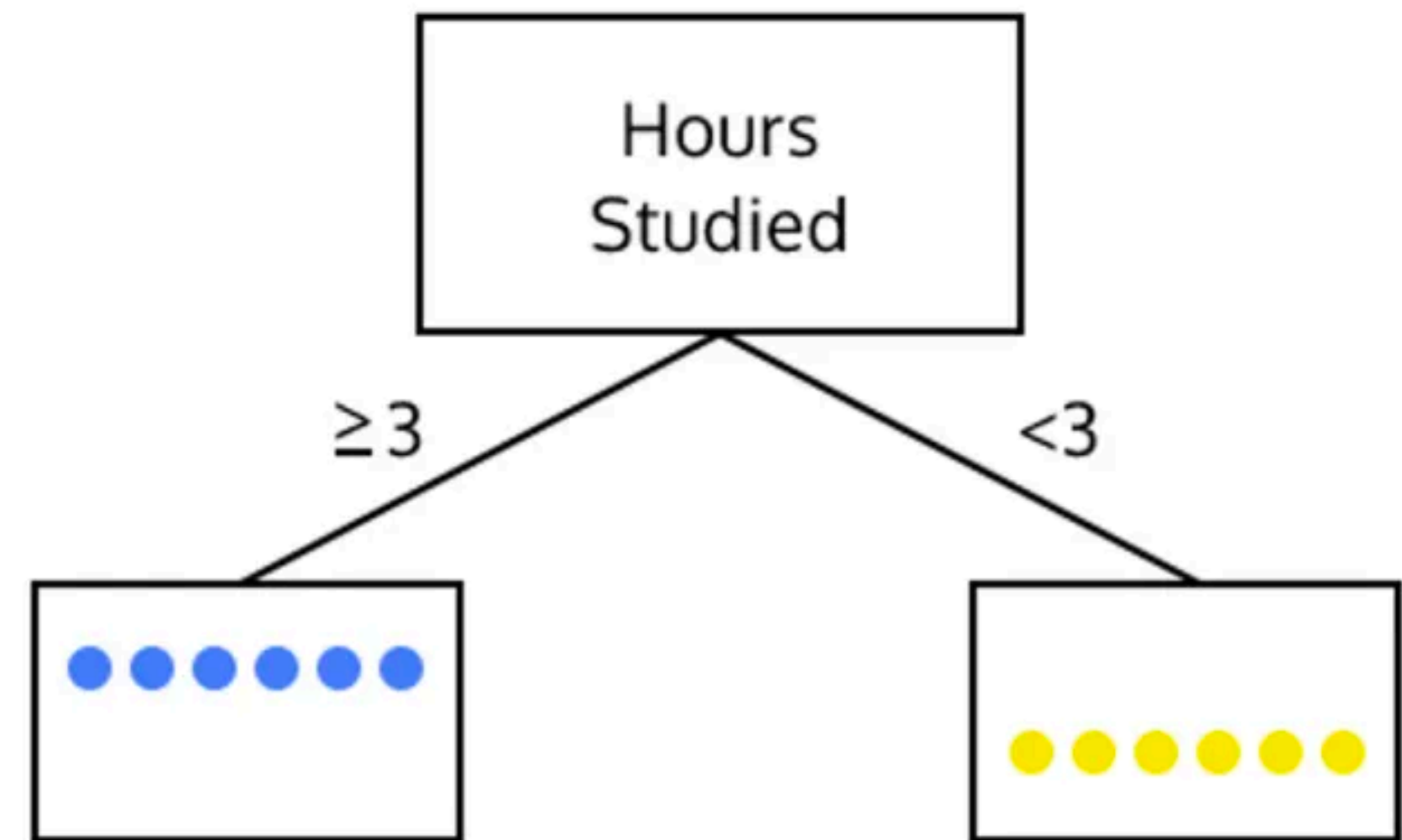


## 지니 불순도 (Gini Impurity)



$$1 - \sum_{j=1}^J (p_j(1 - p_j))$$

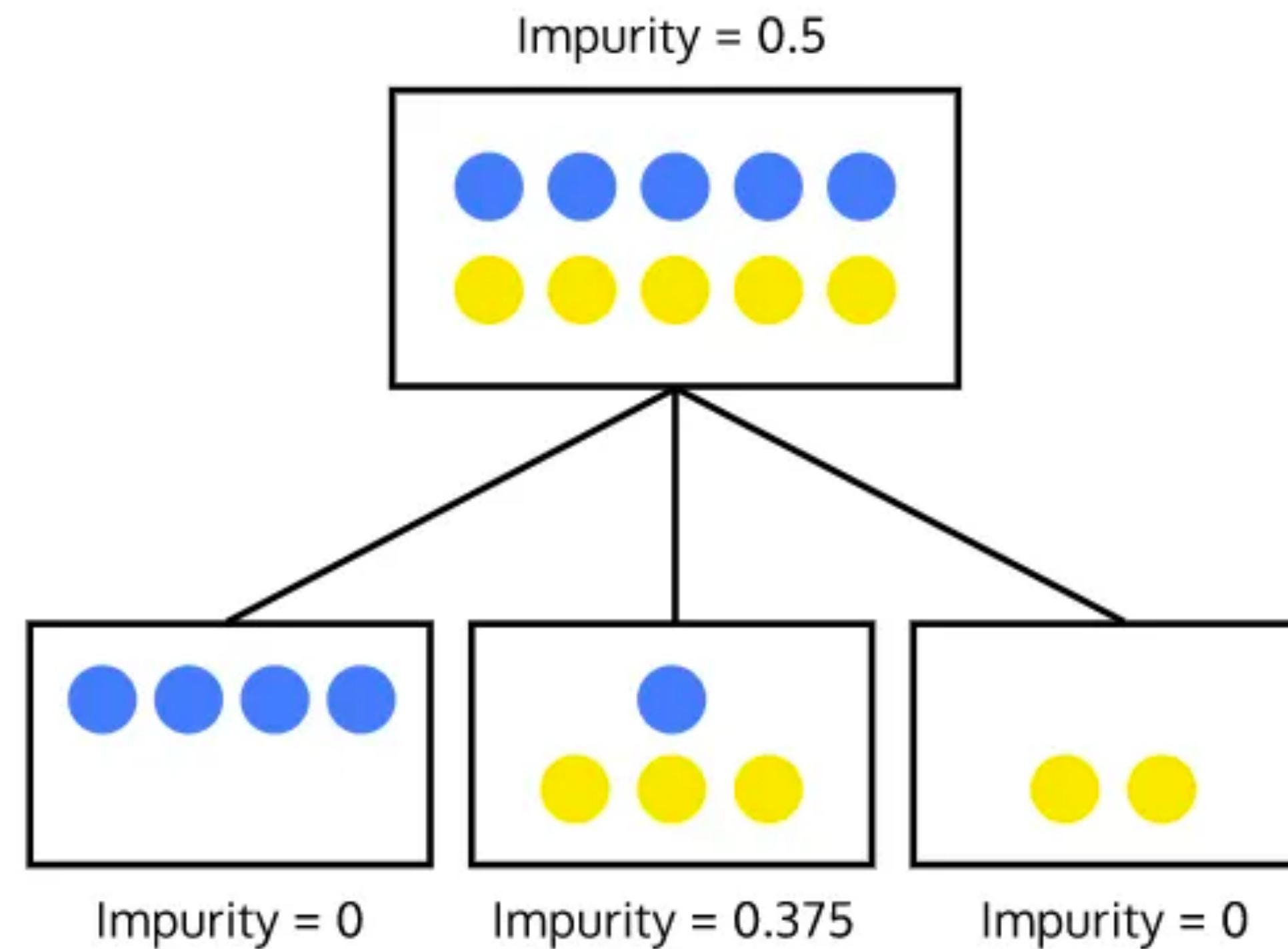
**VS**



Ex) 총 4개 데이터 세트에서 레이블 A가 3개, B가 1개 포함

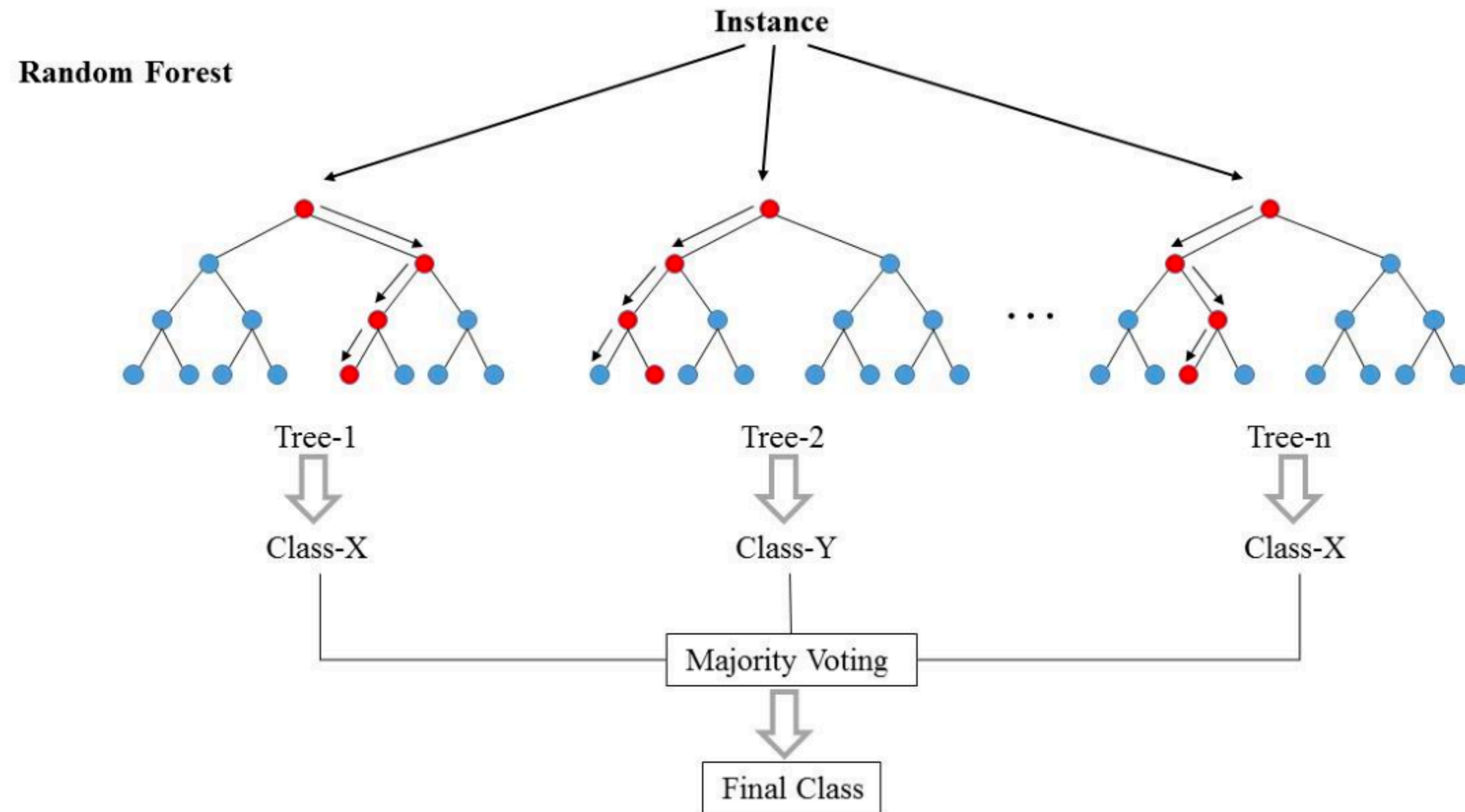
$$1 - \left(\frac{3}{4}\right)^2 - \left(\frac{1}{4}\right)^2 = 0.375$$

## 정보 획득량(Information Gain)



$$\text{Information Gain} = 0.5 - (0 + 0.375 + 0) = 0.125$$

# 랜덤 포레스트(Random Forest)





# 랜덤 포레스트(Random Forest)

