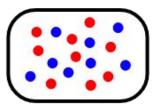# Entropy & Information Gain

- How exactly do we select the "best" feature for partitioning?

- Conceptually, at a very high level, we use "Entropy", which measures the amount of uncertainty for a given dataset

- Partitioning on a good feature should cause the amount of uncertainty to decrease

- "Information Gain" measures this decrease in uncertainty for a given feature

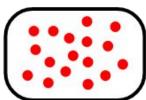- So the "best" feature creates the largest Information Gain
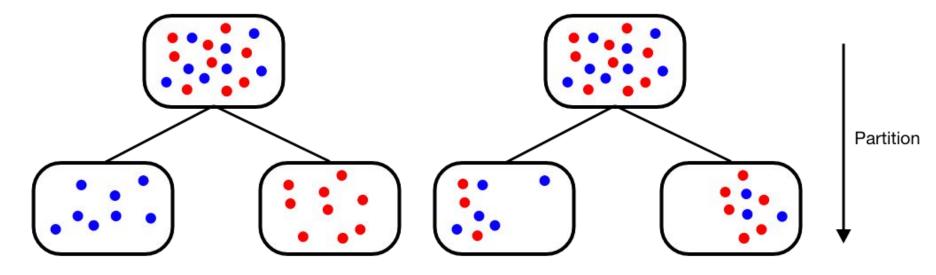
Entropy & Information Gain