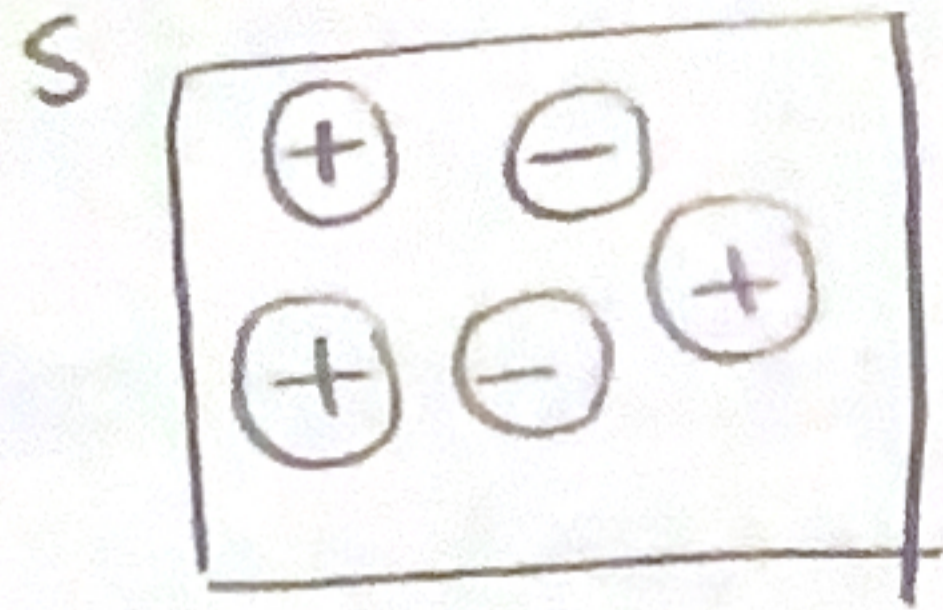


Attribute Selection

Entropy



$$\text{Entropy}(S) = - \underbrace{p_+}_{0.6} \log_2 p_+ - \underbrace{p_-}_{0.4} \log_2 p_-$$

$$\rightarrow -0.6 \log_2 0.6 - 0.4 \log_2 0.4$$

Target attribute can take values more than 2!

$$\text{Then } \rightarrow \text{Entropy}(S) = \sum_{i=1}^n -p_i \log_2 p_i$$

Information Gain

For entropy reduction

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v_i} \frac{|S_{v_i}|}{|S|} \text{Entropy}(S_{v_i})$$

↑
önceki
feature
(parent)

↓
aday
attribute

↳ parent'ta
o değeri
alan example sayısı
parent size