

Cross Entropy

박태우

Entropy란?

의외성 (surprisal)

- 어떤 사건이 주는 놀라움의 정도
- 2018년 러시아 월드컵에서 한국이 독일을 이길 확률은 매우 적었다.
- 하지만 이런 사건이 일어난다면 의외성의 수준이 높은 정도이다.

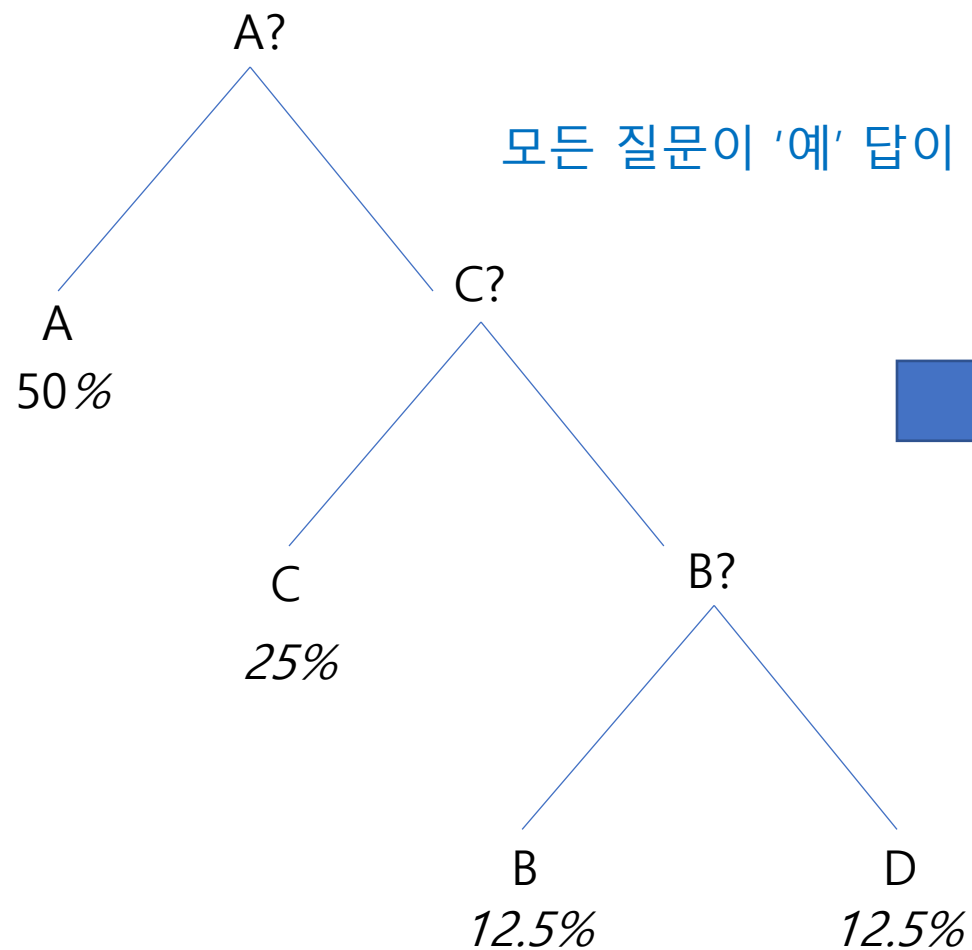
1,2,3 , n 의 사건이 있을때, 사건 i가 발생할 확률이 $p(i)$ 라면 의외성 s 는 이렇게 정의된다.

$$s = -\log p_i$$

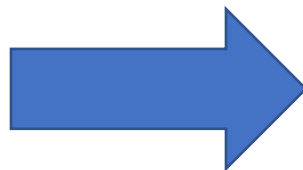
Entropy는 개별 사건이 아니라 전체 사건을 아우르는 현재의 불확실한 상태이므로 기댓값을 사용한다.

$$e = -\sum_{i=0}^n p_i \log p_i$$

Entropy란?



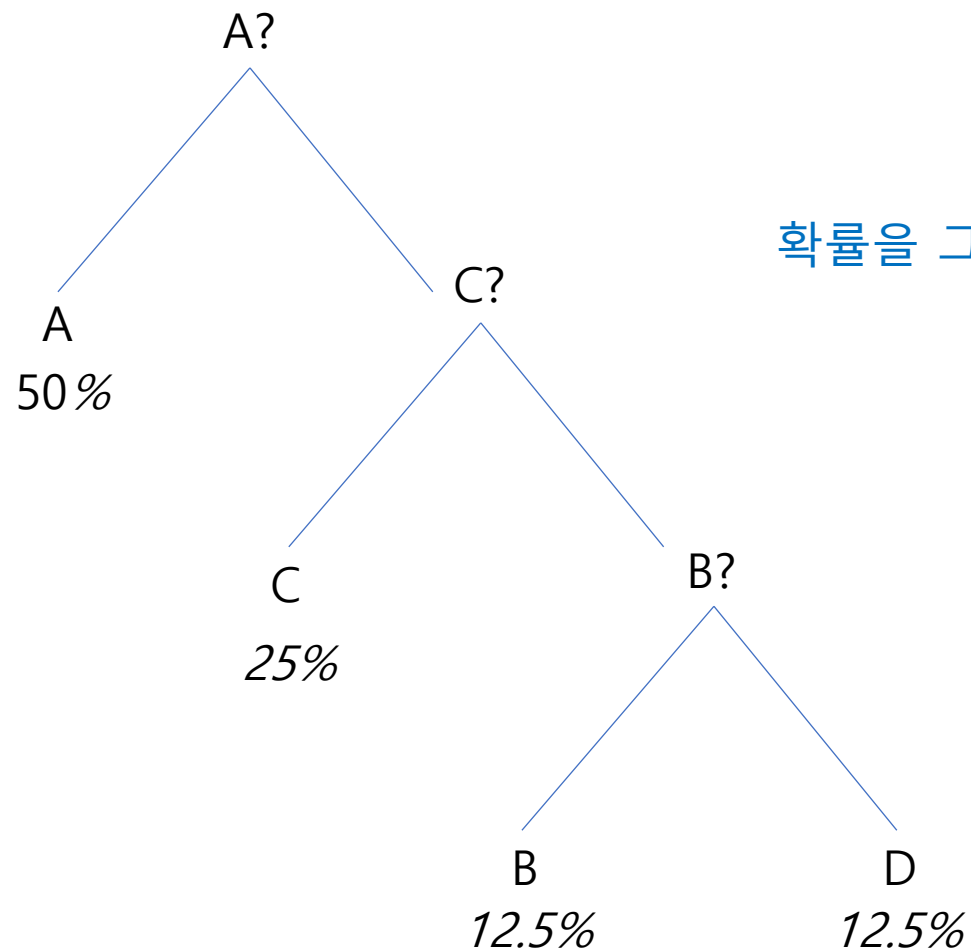
모든 질문이 '예' 답이 50% 일 때



$$\begin{array}{r} 0.5 \times 1 \\ + 0.25 \times 2 \\ + 0.125 \times 3 \\ + 0.125 \times 3 \\ \hline 1.75 \end{array}$$

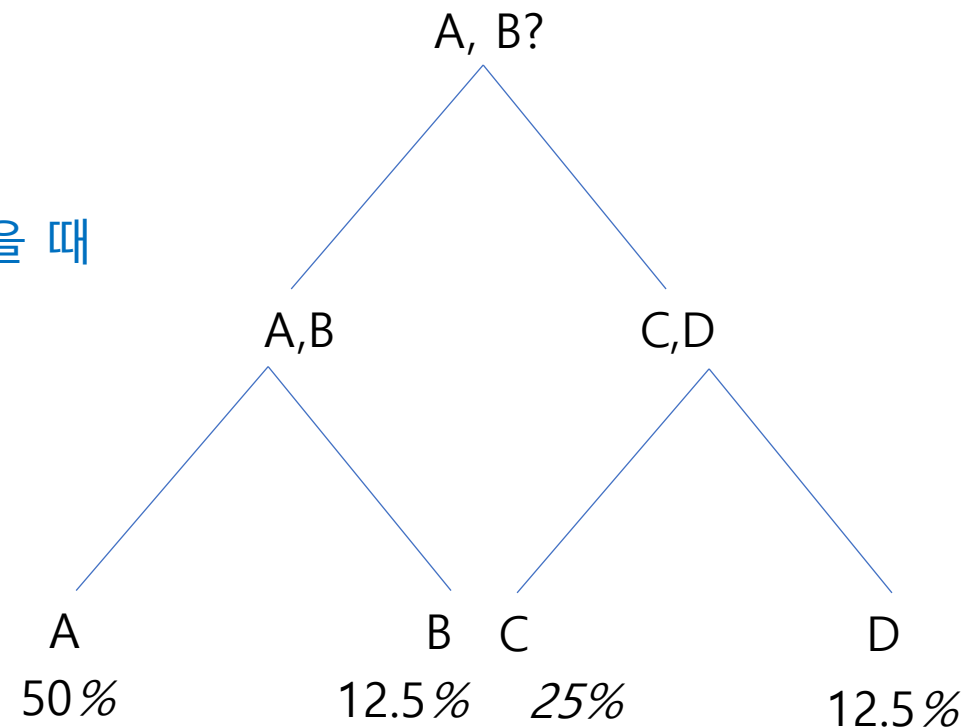
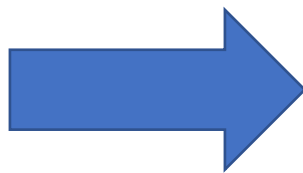
Cross Entropy이란?

- 확률 분포로 된 어떤 문제 p 에 대해 확률 분포로 된 어떤 전략 q 를 사용할 때의 질문 개수의 기댓값



Entropy의 값은 1.75

확률을 그대로 다른 트리에 적용했을 때



Entropy의 값은 2

Cross Entropy이란?

$p = [0.5, 0.125, 0.25, 0.125]$

$q = [0.25, 0.25, 0.25, 0.25]$

q에 대해서 학습하고 있을 때 p의 확률 분포에 가까워질수록 cross entropy의 값은 작아진다.

$$\begin{aligned} H(p, q) &= \sum_i p_i \log_2 \frac{1}{q_i} \\ &= - \sum_i p_i \log_2 q_i \end{aligned}$$

Binary Cross Entropy이란?

확률의 값이 0 또는 1의 값을 가질 때를 가정한다.

즉 class가 2개일 때 (logistic regression)를 뜻함.

$$\begin{aligned} H(p, q) &= \sum_i p_i \log_2 \frac{1}{q_i} \\ &= - \sum_i p_i \log_2 q_i \end{aligned} \quad \xrightarrow{C=2} \quad -(p \log q + (1-p) \log(1-q))$$

https://hyunw.kim/blog/2017/10/26/Cross_Entropy.html

참고 링크

https://hyunw.kim/blog/2017/10/26/Cross_Entropy.html