크로스 엔트로피(Cross Entropy)

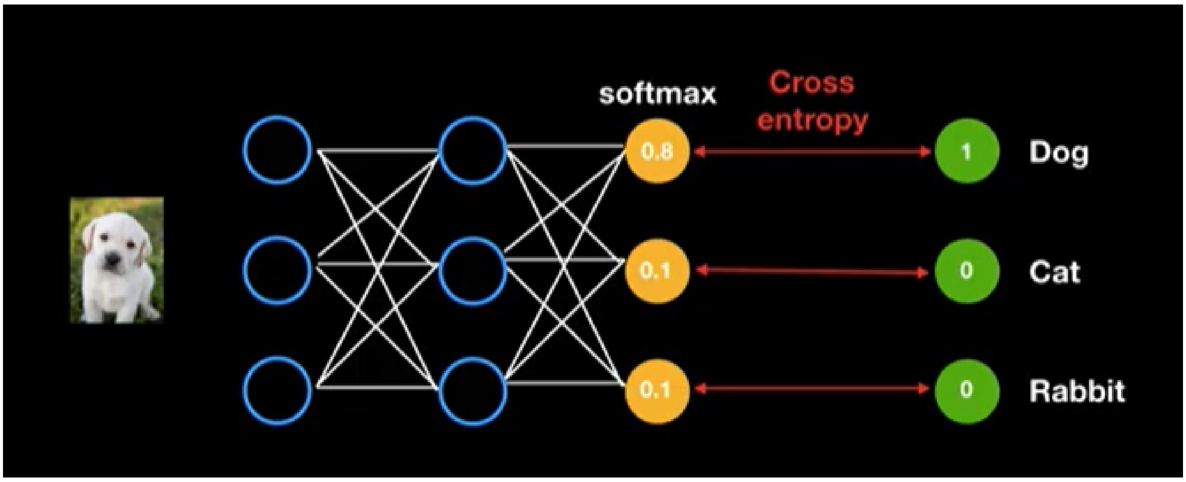
Cross Entropy 정의

- 정의 (in 위키피디아) : 동일한 이벤트 공간의 두 분포 p와 q사이의 교차 엔트로피는 다음과 같이 정의된다

$$H(p, q) = E_p[-log\, q]$$

- 정의 (in wikipedia) : In information theory, the cross entropy between two probability distributions p and q over the same underlying set of events **measures the average number of bits needed to identify** an event drawn from the set if a coding scheme used for the set is optimized for an esimated probablilty distribution q, rather than the true distribution p
- 즉, cross entropy는 두 확률 분포 p와 q를 구분하기 위해 필요한 평균 비트 수를 의미

Cross Entropy as cost function for classfication



- 학습시킬때 model의 예측값이 실제값과 얼마나 유사한지 알고싶기에

Entropy VS. Cross Entropy

amount of info.

$$H(X) = \sum_{i=1}^{n} \log_2 \frac{1}{p_i} \times p_i$$

$$H(P,Q) = \sum_{i=1}^{n} \log_2 \frac{1}{q_i} \times p_i$$

probability distribution

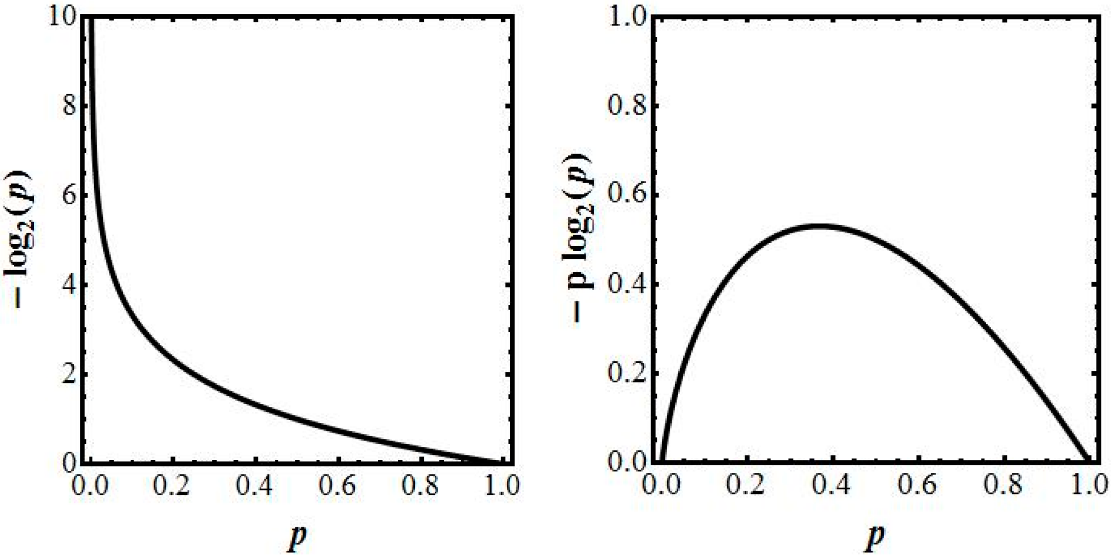
크로스 엔트로피의 특징

1. When Q is totally wrong, cross entropy is infinity.
   - ex : $log_2\frac{1}{0.0} * 1 + log_2\frac{1}{1.0} * 0 + log_2\frac{1}{0.0} * 0 = \infty$
2. When Q is similar to P, cross entropy is similar to entropy.
   - ex : $log_2\frac{1}{0.8} * 1 + log_2\frac{1}{0.1} * 0 + log_2\frac{1}{0.1} * 0 = 0.32$
3. When Q == P, cross entropy is same with entropy
   - $log_2\frac{1}{1.0} * 1 + log_2\frac{1}{0.0} * 0 + log_2\frac{1}{0.0} * 0 = 0.0$

즉, Cross Entropy의 값은 (Shannon) Entropy의 값보다 **크거나 같다.**

참조

- $log_2\frac{1}{a} = -log_2 a$
- $-log_2(p)$와 $-plog_2(p)$ 그래프



(a) **(b)**

출처 : https://www.youtube.com/watch?v=Jt5BS71uVfI&list=PLVNY1HnUlO241gILgQloWAs0xrrkqQfKe&index=53