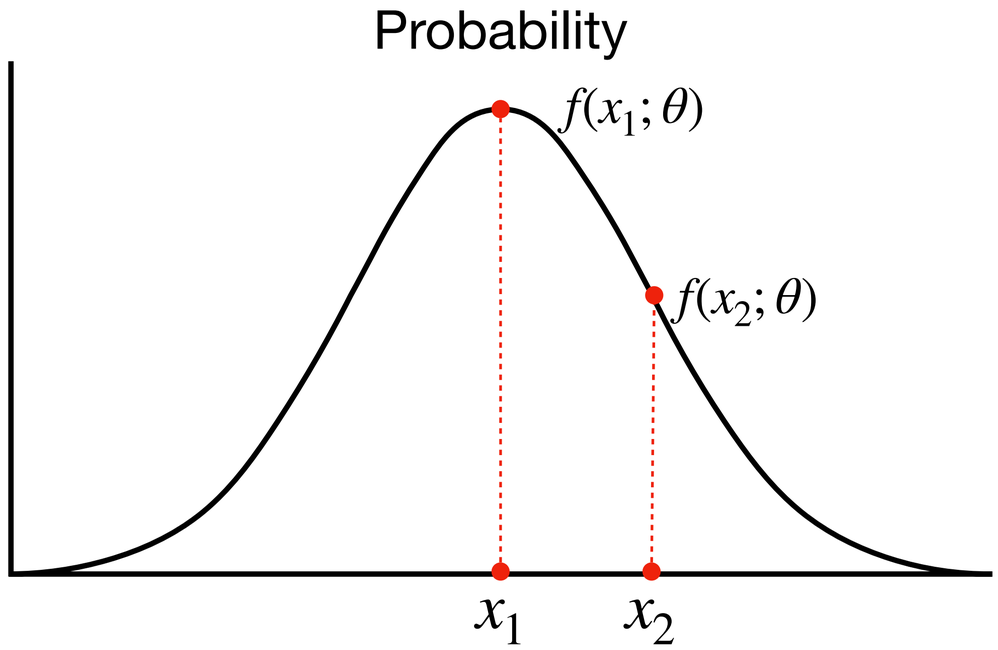
삼성전기 AI전문가 양성과정 - 프로젝트 실습 (비영상)

자연어처리를 위한 Cross Entropy

현청천

2022.02.28

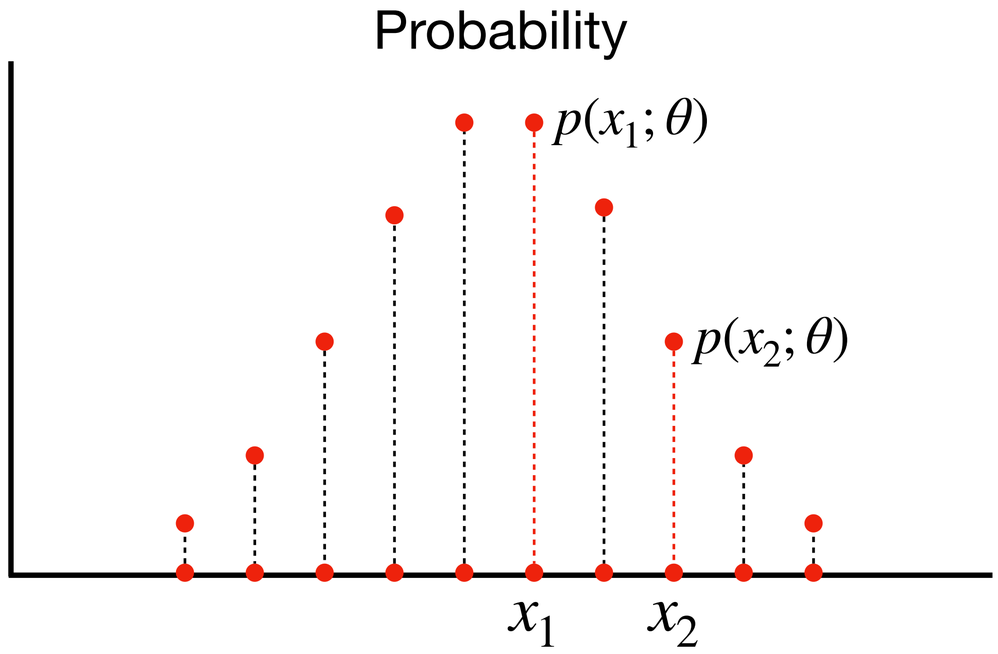
Entropy (연속확률분포)



$$H(X) = \int_x f(x;\theta) \log \frac{1}{f(x;\theta)} dx$$

확률분포에 대한 정보량의 기댓값

Entropy (이산확률분포)



$$H(X) = \sum_x p(x;\theta) \log \frac{1}{p(x;\theta)}$$

확률분포에 대한 정보량의 기댓값

Entropy (정보량)

내일은 해가 동쪽에서 뜬다

$$p = 0.99999999$$

$$I(p) = \log \frac{1}{p} = \frac{1}{0.99999999} \simeq 0$$

내일은 해가 서쪽에서 뜬다

$$q = 0.00000001$$

$$I(q) = \log \frac{1}{q} = \frac{1}{0.00000001} \simeq 18.42068074$$

Entropy

A	25%	0 0
B	25%	0 1
C	25%	1 0
D	25%	1 1

$\log_2 \frac{1}{0.5} = 1$

$\log_2 \frac{1}{0.25} = 2$

$\log_2 \frac{1}{0.125} = 3$

$0.25 \times 2 + 0.25 \times 2 + 0.25 \times 2 + 0.25 \times 2 = 2$

$$H(x) = \sum_x p_\theta(x) \log_2 \frac{1}{p_\theta(x)}$$

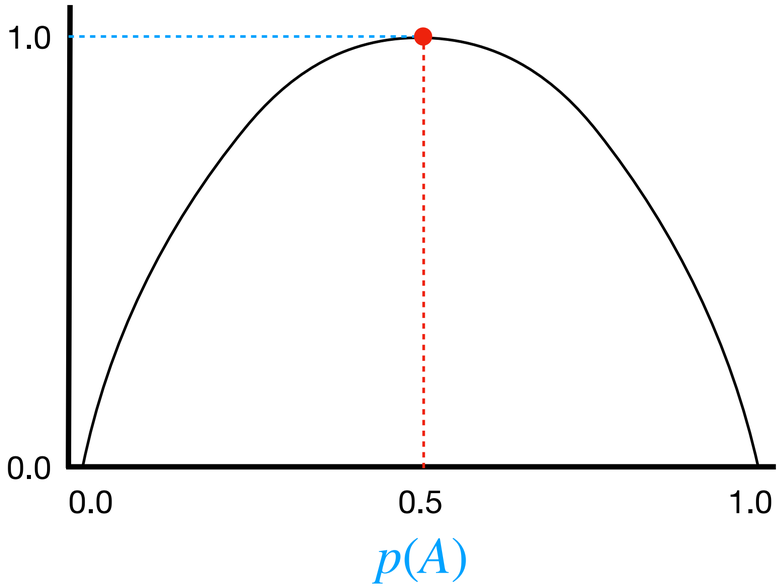
A	50%	0
B	25%	1 0
C	12.5%	1 1 0
D	12.5%	1 1 1

$0.5 \times 1 + 0.25 \times 2 + 0.125 \times 3 + 0.125 \times 3 = 1.75$

Entropy

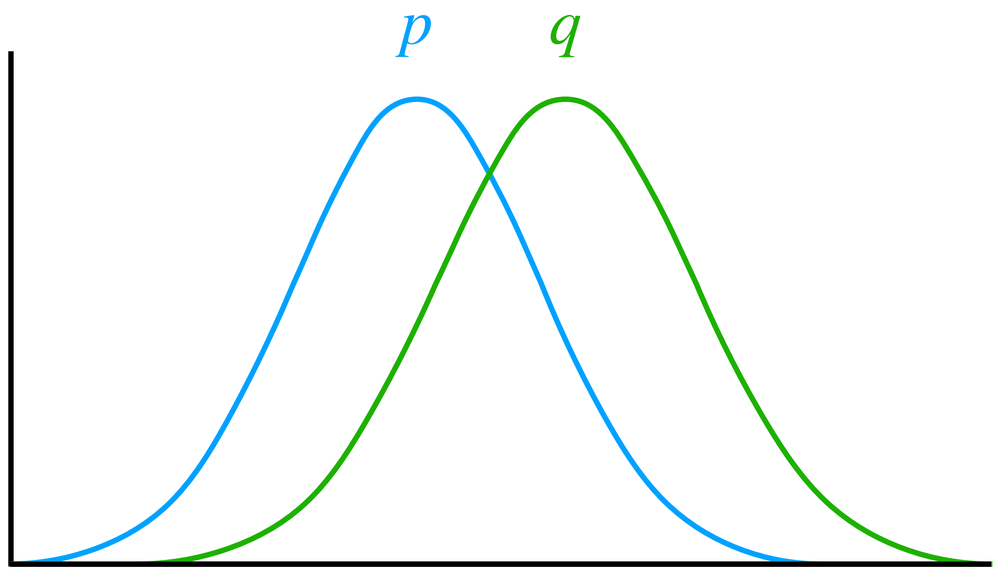
A, B 두 글자가 발생하는 경우

A 발생 확률에 따른 Entropy



확률분포의 불확실성이 증가하면 Entropy가 증가

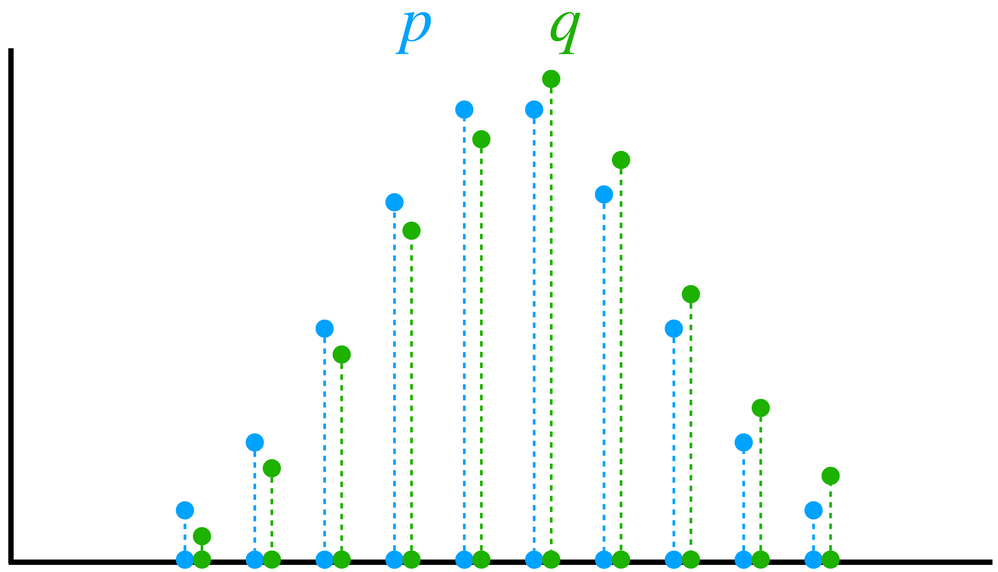
Cross Entropy (연속확률분포)



$$H(p,q) = \int_x p(x) \log \frac{1}{q(x)} dx$$

확률분포 p에 대한 확률분포 q의 정보량의 기댓값

Cross Entropy (이산확률분포)



$$H(p,q) = \sum_{x} p(x) \log \frac{1}{q(x)}$$

확률분포 p에 대한 확률분포 q의 정보량의 기댓값

Cross Entropy

p

A	50%	0
B	25%	1 0
C	12.5%	1 1 0
D	12.5%	1 1 1

$\log_2 \frac{1}{0.5} = 1$

$\log_2 \frac{1}{0.25} = 2$

$\log_2 \frac{1}{0.125} = 3$

q

A	25%	0 0
B	25%	0 1
C	25%	1 0
D	25%	1 1

$$H(p,q) = \sum_x p(x) \log_2 \frac{1}{q(x)}$$

$$0.5 \times 2 + 0.25 \times 2 + 0.125 \times 2 + 0.125 \times 2 = 2$$

Cross Entropy

p

A	50%	0
B	25%	1 0
C	12.5%	1 1 0
D	12.5%	1 1 1

$\log_2 \frac{1}{0.5} = 1$

$\log_2 \frac{1}{0.25} = 2$

$\log_2 \frac{1}{0.125} = 3$

q

A	50%	0
B	25%	1 0
C	12.5%	1 1 0
D	12.5%	1 1 1

$$H(p,q) = \sum_x p(x) \log_2 \frac{1}{q(x)}$$

$$0.5 \times 1 + 0.25 \times 2 + 0.125 \times 3 + 0.125 \times 3 = 1.75$$

Cross Entropy

A, B 두 글자가 발생하는 경우

A 발생 확률에 따른 두 확률분포의

Cross Entropy

$q(A)$

$p(A)$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.33	0.36	0.44	0.55	0.69	0.88	1.12	1.47	2.08
0.2	0.54	0.5	0.53	0.59	0.69	0.84	1.03	1.33	1.86
0.3	0.76	0.64	0.61	0.63	0.69	0.79	0.95	1.19	1.64
0.4	0.98	0.78	0.70	0.67	0.69	0.75	0.87	1.05	1.42
0.5	1.20	0.92	0.78	0.71	0.69	0.71	0.78	0.92	1.20
0.6	1.42	1.05	0.87	0.75	0.69	0.67	0.7	0.78	0.98
0.7	1.64	1.19	0.95	0.79	0.69	0.63	0.61	0.64	0.76
0.8	1.86	1.33	1.03	0.84	0.69	0.59	0.53	0.5	0.54
0.9	2.08	1.47	1.12	0.88	0.69	0.55	0.44	0.36	0.33

Cross Entropy Loss

Cross Entropy loss $$CE = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\ \log\frac{1}{\hat{y}_{ij}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\ \log\hat{y}_{ij}$$

정답확률분포와 예측확률분포의 Cross Entropy의 평균

Cross Entropy

y	
A	4개
B	3개
C	2개
D	1개

	A	B	C	D
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0
4	1	0	0	0
5	0	1	0	0
6	0	1	0	0
7	0	1	0	0
8	0	0	1	0
9	0	0	1	0
10	0	0	0	1

Cross Entropy

$\hat{y}$

A	25%
B	25%
C	25%
D	25%

	A	B	C	D
1	0.25	0.25	0.25	0.25
2	0.25	0.25	0.25	0.25
3	0.25	0.25	0.25	0.25
4	0.25	0.25	0.25	0.25
5	0.25	0.25	0.25	0.25
6	0.25	0.25	0.25	0.25
7	0.25	0.25	0.25	0.25
8	0.25	0.25	0.25	0.25
9	0.25	0.25	0.25	0.25
10	0.25	0.25	0.25	0.25

$$CE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij} \log \hat{y}_{ij} = 1.3862943611198906$$

Cross Entropy

$\hat{y}$

A	40%
B	30%
C	20%
D	10%

	A	B	C	D
1	0.40	0.30	0.20	0.10
2	0.40	0.30	0.20	0.10
3	0.40	0.30	0.20	0.10
4	0.40	0.30	0.20	0.10
5	0.40	0.30	0.20	0.10
6	0.40	0.30	0.20	0.10
7	0.40	0.30	0.20	0.10
8	0.40	0.30	0.20	0.10
9	0.40	0.30	0.20	0.10
10	0.40	0.30	0.20	0.10

$$CE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij} \log \hat{y}_{ij} = 1.2798542258336674$$

Cross Entropy Loss vs Negative Log Likelihood

$$CE = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\ \log \hat{y}_{ij}$$

$$CE = -\frac{1}{N}\sum_{i=1}^{N} \log \hat{y}_i$$

$$NLL = -\sum_{i=1}^{n} \log \hat{y}_i$$

Minimize cross entropy loss

Minimize negative log likelihood

Maximize likelihood

Cross Entropy Loss (NMIST)

$$p(y \mid x; \theta)$$

y

5	0	0	0	0	0	1	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0

$\hat{y}$

0.04	0.03	0.02	0.1	0.01	0.7	0.01	0.02	0.03	0.04
0.6	0.01	0.2	0.02	0.03	0.01	0.04	0.02	0.04	0.03
0.01	0.04	0.03	0.02	0.75	0.03	0.05	0.02	0.04	0.01
0.03	0.65	0.01	0.04	0.02	0.01	0.03	0.02	0.15	0.04

감사합니다.