

N / A / N / O / D / E / G / R / E / E

지도학습 알고리즘

07. Information Theory

Logistic regression

- ✓ Negative logarithm of the likelihood, which gives the cross-entropy error function

$$E(\mathbf{w}) = -\log p(\mathbf{t} | \mathbf{w}) = -\sum_{n=1}^N \{t_n \log y_n + (1 - t_n) \log(1 - y_n)\}$$

- ✓ Taking the gradient of the error function, we obtain

$$\nabla E(\mathbf{w}) = \sum_{n=1}^N (y_n - t_n) \mathbf{x}_n$$

Entropy

- ✓ Discrete random variable

X

- ✓ Probability mass function

$p(x)$

- ✓ Entropy = 평균 정보량

$$H(p) = \mathbb{E}[-\log p] = -\sum p(x) \log p(x)$$

Entropy와 불확실성, 그리고 정보량

- ☑ Entropy가 최대인 확률 분포
 - Discrete: uniform distribution
 - Continuous: Gaussian distribution



Cross-entropy and relative entropy

- ✓ Cross-entropy

$$H(p, q) = -\mathbf{E}_p[\log q]$$

- ✓ Relative entropy

$$\mathcal{D}_{\text{KL}}(p||q) = \mathbf{E}_p \left[\log \frac{p}{q} \right]$$

- ✓ 두 분포가 같을 때, relative entropy = 0

$$\mathcal{D}_{\text{KL}}(p||q) = 0 \iff p(x) = q(x)$$

Cross-entropy and relative entropy

- ✓ Entropy, cross-entropy and relative entropy

$$H(p, q) = H(p) + \mathcal{D}_{\text{KL}}(p||q)$$