

Information Theory

Additional mini-project in deeplearning math

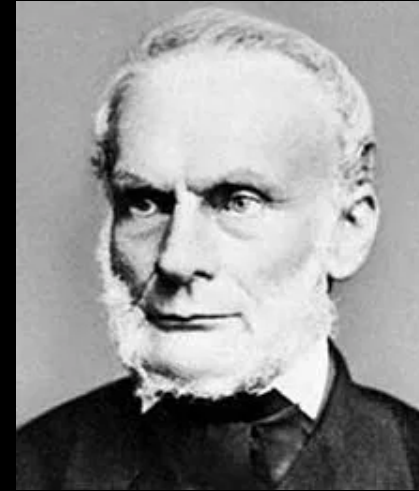
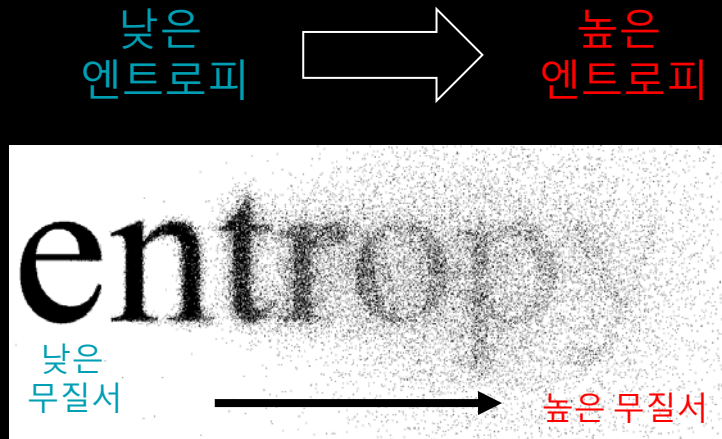
Orientation (오리엔테이션)

소프트웨어 끈대 강의

노기섭 교수

(kafa46@cju.ac.kr)

단어조차 생소한 entropy...



Origng of Entropy: 1865년, Rudolf J. E. Clausius (독일, 물리학자)

Entropy in Physics: Degree of Randomness

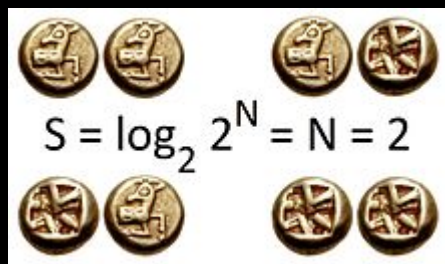
열역학 제2법칙: 고립계에서는 엔트로피가 증가만 하고 감소하지 않는다.

정보 엔트로피의 등장

뭐야 $\pi\pi$ 엔트로피에도 종류가 있어요?

네~~

1948년 Shannon에 의해
발명된 'Information Entropy'라는게 있어요^^



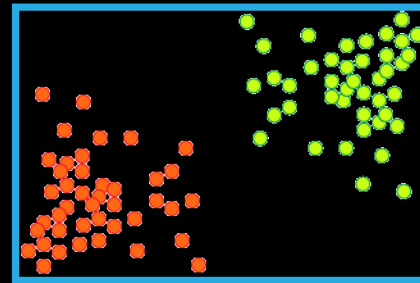
Claude E. Shannon
(미국, 1916~2001)
수학자, 컴퓨터과학자

정보 이론의 창시자
정보 이론의 아버지라고
불림

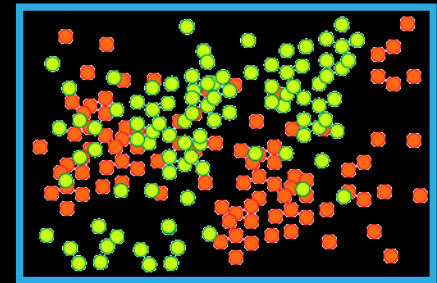
딥러닝에도 엔트로피를 사용하나요?

BCE: Binary Cross Entropy

CE: Cross Entropy



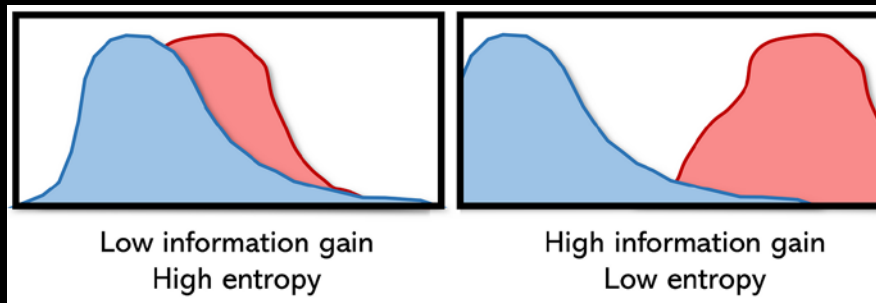
Low Entropy



High Entropy

이미지 출처:
<https://www.javatpoint.com/entropy-in-machine-learning>

KL Divergence Loss: Kullback-Leibler Divergence Loss

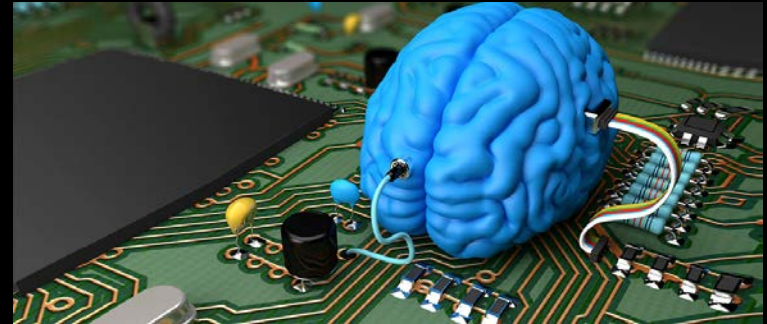
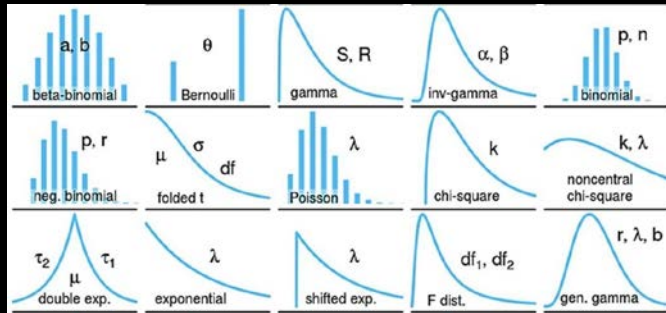
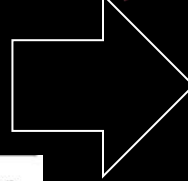


이미지 출처:
<https://towardsdatascience.com/understanding-entropy-the-golden-measurement-of-machine-learning-4ea97c663dc3>

딥러닝이 학습하는 방법

데이터로부터 패턴을 배운다

데이터로부터 확률 분포를 배운다

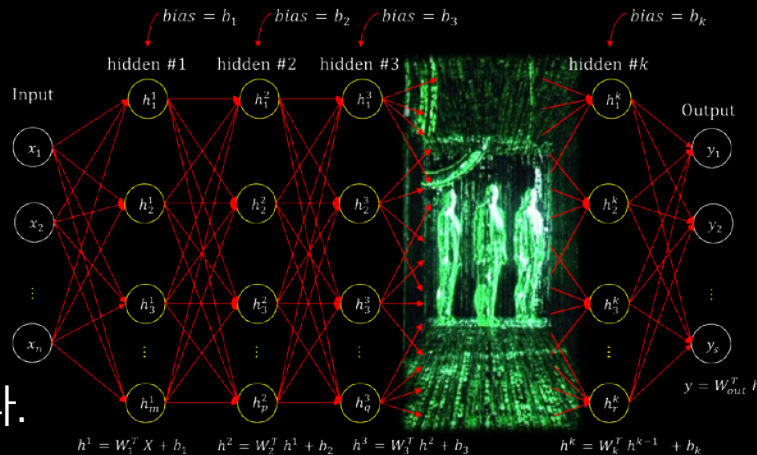


이미지 출처:
<https://knowledge.wharton.upenn.edu/article/democratization-ai-means-tech-innovation/>

Learning from Dataset

데이터 x 를
숫자로 만들고
입력 계층에
전달한다.

모든 파라미터를
랜덤하게 설정한다.



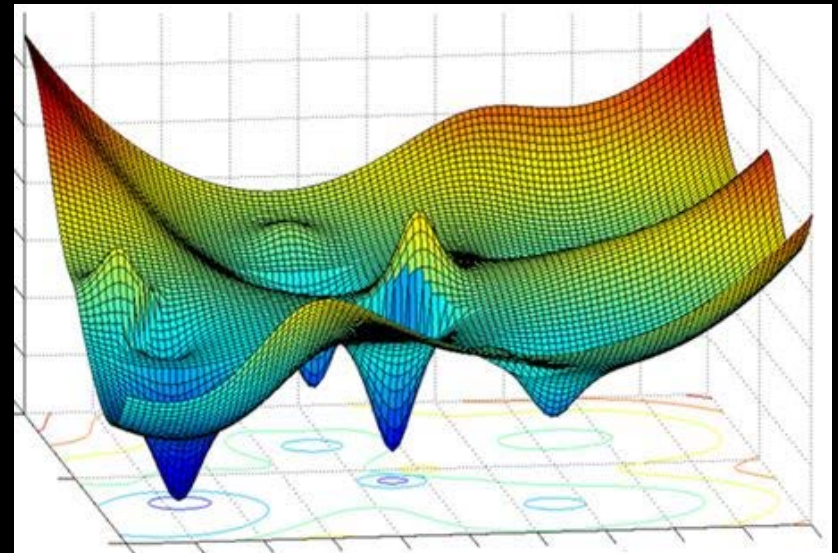
y 가 나올 확률을
계산한다.

얼마나 틀렸는지
계산한다.

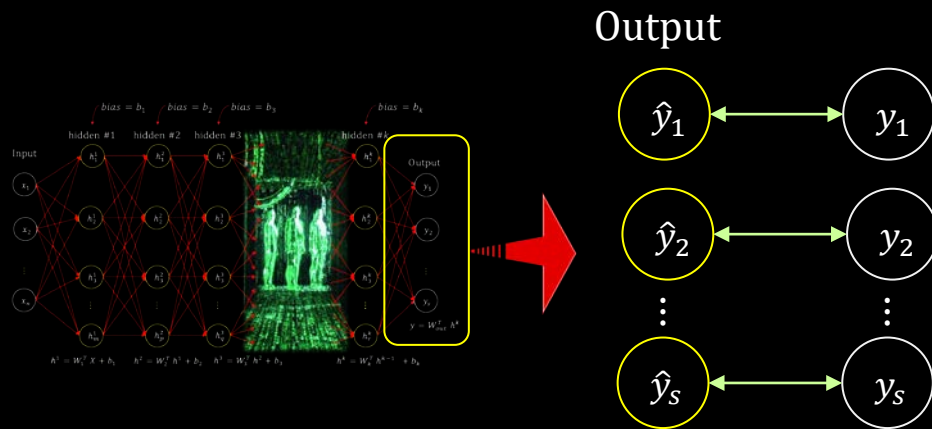
틀린 정도를 최소화할 수
있도록 모든 파라미터를
업데이트 한다.



너무 많은 파라미터.....,
무한대에 가까운
파라미터 조합 때문에
정확히 알 수는 없지만...
파라미터 세팅에 따라서
얼마나 틀렸는지(Loss)에 대한
확률 분포가 존재할 것.



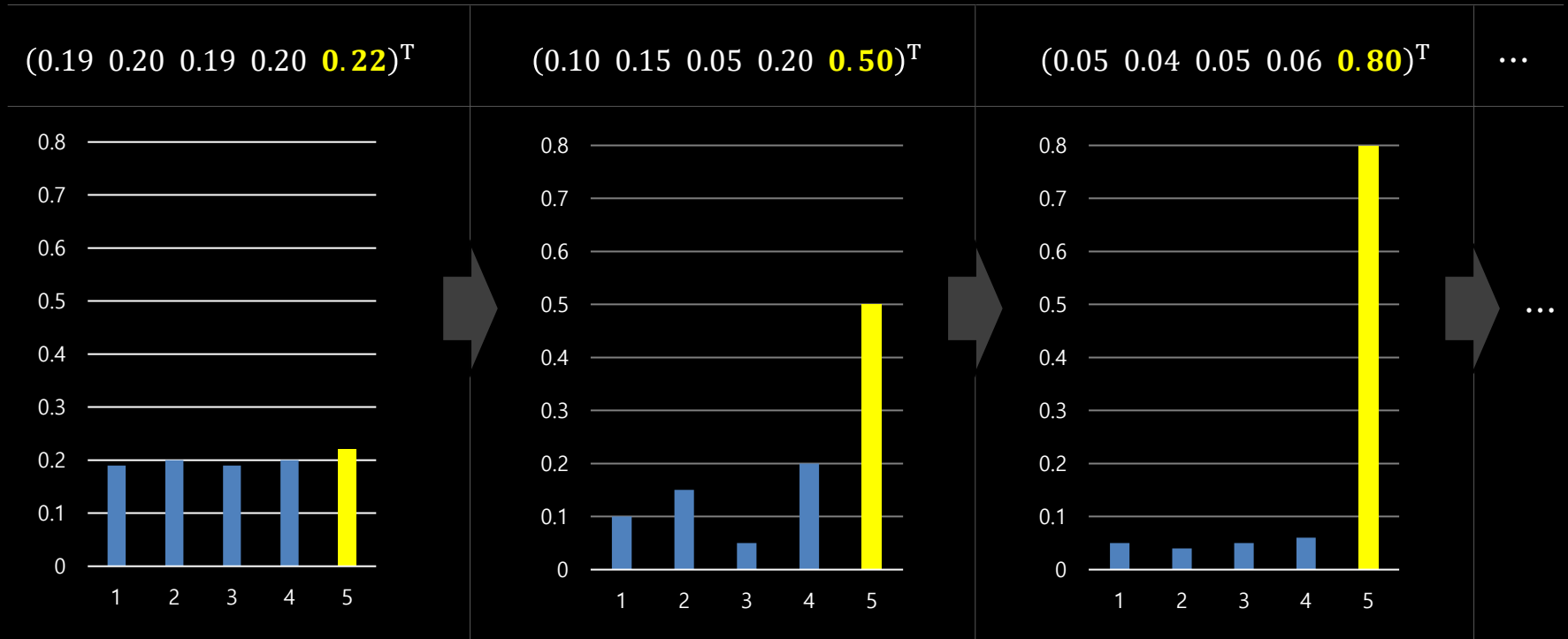
딥러닝 학습을 정보이론 관점으로 이해하기



$$L(\theta) = \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

$$\sum_{i=1}^5 \left(\begin{pmatrix} 0.19 \\ 0.20 \\ 0.19 \\ 0.20 \\ 0.22 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right)^2 = 0.19^2 + 0.20^2 + 0.19^2 + 0.20^2 + (-0.78)^2$$

정답 (Correct Answer) is indicated by a red arrow pointing to the value 1 in the target vector.



예측 손실(loss) 감소와 무질서도(randomness) 변화

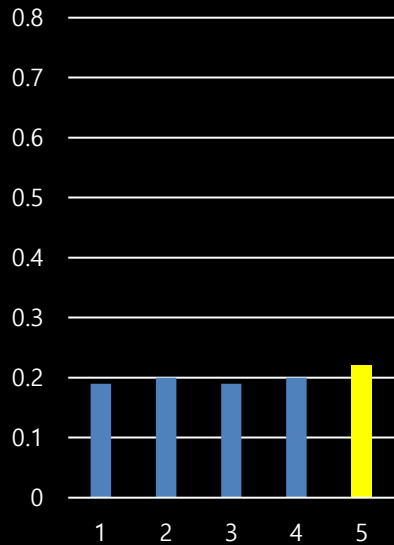
학 습 진 행

높은
엔트로피

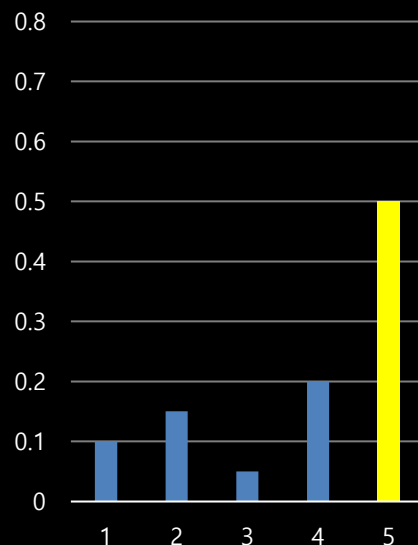
무질서도(randomness) → 계속 감소.....

낮은
엔트로피

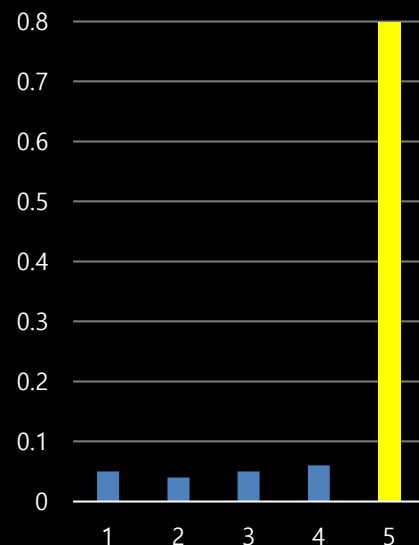
$(0.19 \ 0.20 \ 0.19 \ 0.20 \ 0.22)^T$



$(0.10 \ 0.15 \ 0.05 \ 0.20 \ 0.50)^T$



$(0.05 \ 0.04 \ 0.05 \ 0.06 \ 0.80)^T$



...

...

Motivation & Course Objectives

딥러닝과 어떤 관계가 있을 것 같아요 ^^.

교수님~~ 좀 더 자세히 알고 싶어요!



Concept of Information

What is the Information?

Deep learning & entropy

Entropy loss, KL divergence

Course Structure
→ see next slide

Course Overview

Topic	Contents
01. Orientation 오리엔테이션	Motivations & Course introduction 동기부여, 과정 소개
02. Information 정보	What is the information? Concept & definition 정보란 무엇인가? 개념과 정의
03. Information Entropy 정보 엔트로피	Concepts, notation, and operations on information entropy 정보 엔트로피의 개념, 표기, 연산
04. Entropy in Deep learning 딥러닝에서의 엔트로피	How to apply the information entropy into Deep learning? 어떻게 정보 엔트로피를 딥러닝에 적용하는가?
05. Entropy Loss 엔트로피 손실	Loss function using entropy, BCE, and cross entropy 엔트로피를 이용한 손실 함수, BCE, 크로스 엔트로피
06. KL Divergence KL 발산	Concept & definition of KL divergence KL 발산의 개념과 정의
07. Summary & Closing 요약 및 마무리	Summary & closing on this project, 'Information Theory' 정보 이론 요약 및 마무리



수고하셨습니다 ..^^..