

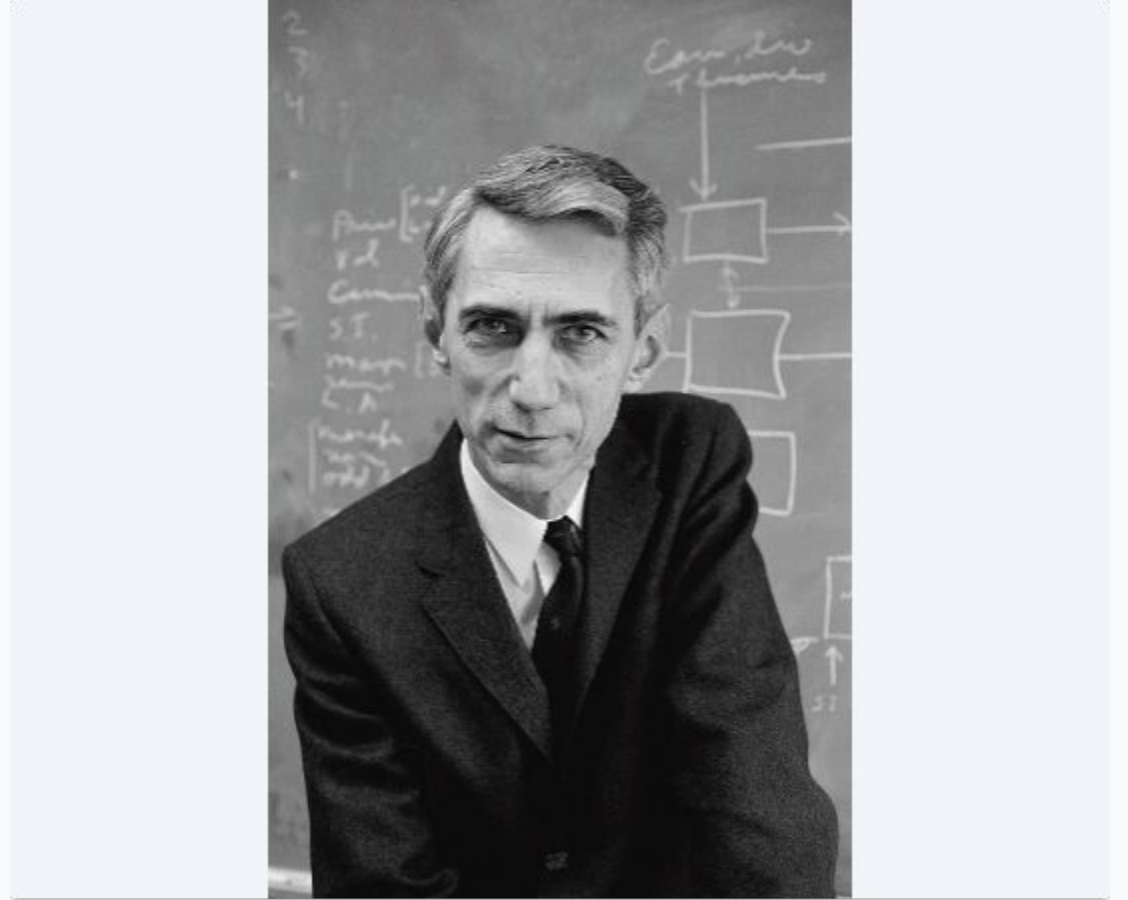
새년의 엔트로피와 AI의 최적화

불확실성을 수학적으로 정복하여

인공지능의 학습을 이끄는 핵심 원리

| 정보 이론의 아버지, 클로드 새넌

- 벨 연구소(Bell Labs)의 수학자이자 전기 공학자로서, 현대 디지털 통신의 기반을 닦았습니다.
- 당시 통신 시스템은 '노이즈(Noise)'로 인해 신호가 왜곡되는 문제가 심각했습니다.
- 새넌은 "정보(Information)"를 어떻게 측정할 것인가?라는 근본적인 질문을 던졌습니다.
- 그는 정보를 '의미'가 아닌 '불확실성의 해소' 관점에서 수학적으로 정의하고자 했습니다.



| 불확실성과 정보량의 관계



확실한 사건

"내일 해가 동쪽에서 뜬다."

너무나 당연한 사실은 우리에게 새로운 정보를 주지 않습니다. **불확실성이 낮으면 정보량도 낮습니다.**



불확실한 사건

"오늘 오후 2시에 번개가 칠 것이다."

예측하기 어려운 사건이 발생했을 때 우리는 크게 놀랍니다. **불확실성이 높을수록 정보량은 큼니다.**

| 엔트로피 (Entropy)의 수학적 정의

새넨은 불확실성의 평균값(기대값)을 **엔트로피**라고 정의했습니다.
이는 메시지를 표현하는 데 필요한 최소한의 비트(bit) 수와도
연결됩니다.

$$H(X) = - \sum_{i=1}^n P(x_i) \log_2 P(x_i)$$

(P(x): 사건 x가 일어날 확률)



| 왜 하필 '엔트로피'인가?

”

"그걸 **엔트로피**라고 부르게. 첫째, 그 식은
물리학의 엔트로피 식과 같네. 둘째, 아무도
엔트로피가 뭔지 모르니 논쟁하면 자네가 무조건
이길 걸세."

— 폰 노이만 (John von Neumann)이 새넌에게

| AI를 위한 확장: 크로스 엔트로피

엔트로피 (Entropy)

하나의 확률 분포 P 가 가진 **내재적인 불확실성**

"데이터 자체가 얼마나 무질서한가?"

크로스 엔트로피 (Cross-Entropy)

두 확률 분포 P (정답)와 Q (예측) 사이의 **차이 (Difference)**

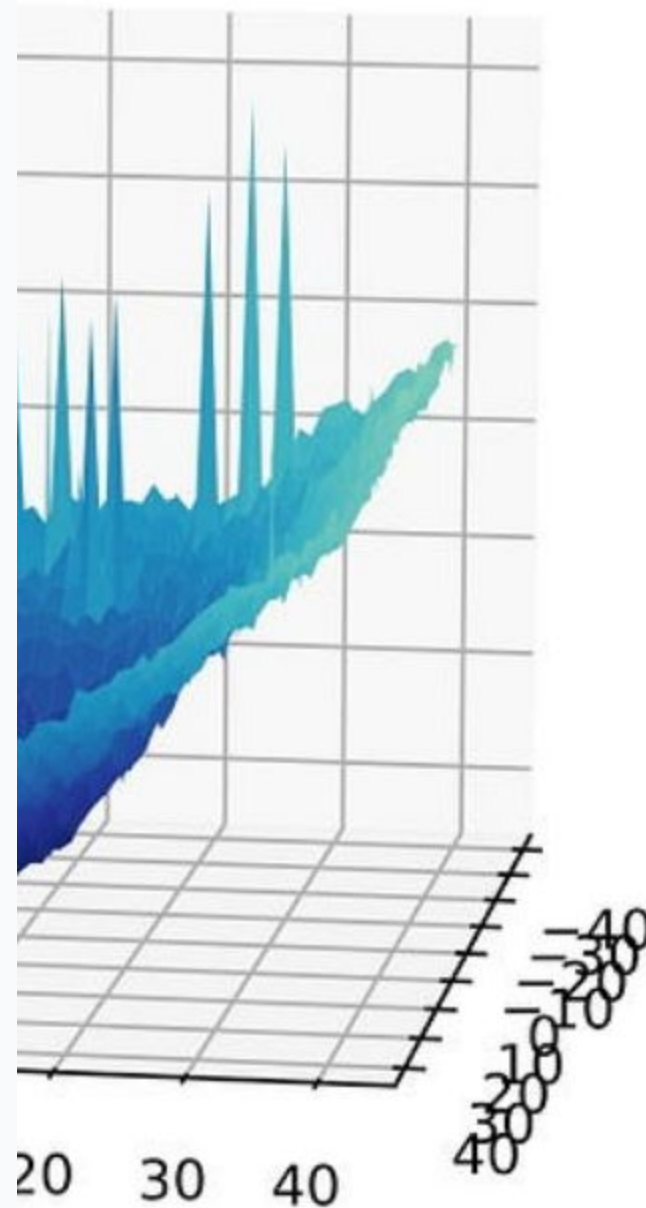
"AI의 예측이 정답과 얼마나 다른가?"

AI는 불확실성을 연산하며 배운다

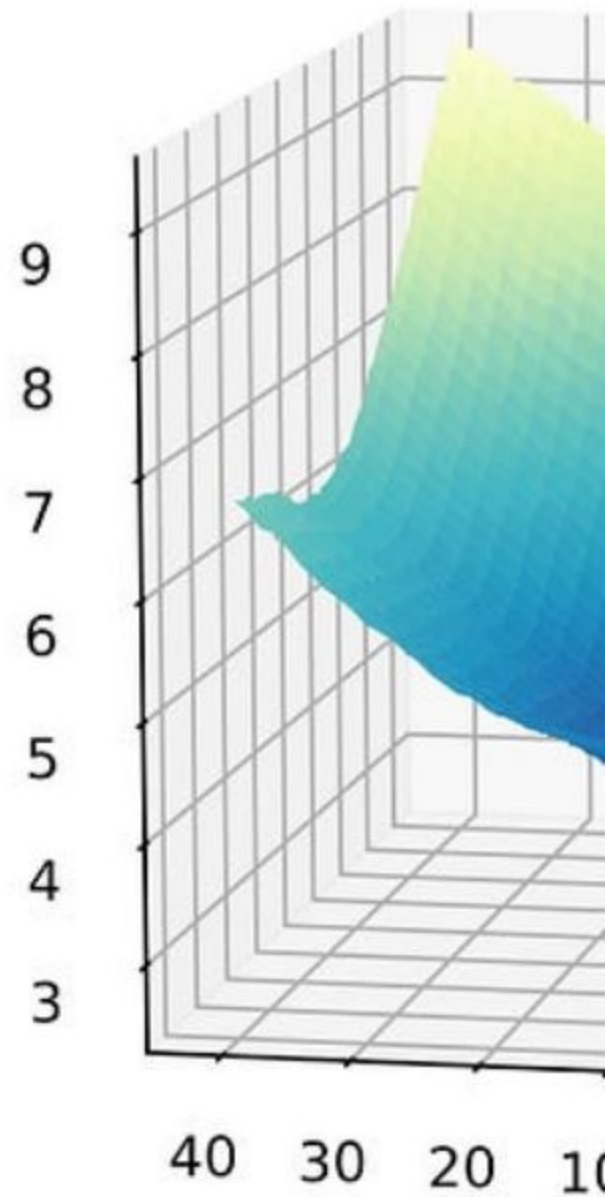
AI 모델의 학습 목표는 **크로스 엔트로피(손실, Loss)**를 **최소화**하는 것입니다.

예측 분포 Q 가 실제 정답 분포 P 와 가까질수록, 크로스 엔트로피 값은 낮아집니다.

이 '차이'를 줄이는 방향으로 파라미터를 조정하는 과정이 바로 **경사 하강법(Gradient Descent)**이며, 이는 곧 불확실성을 줄여가는 여정입니다.



andscape



Learned

| 요약: 새년에서 AI까지

1. 새년의 통찰

정보는 측정 가능한 물리량이며, 그 본질은 '불확실성의 해소'이다.

2. 엔트로피의 도입

불확실성을 $-\sum p \log p$ 라는 수식으로 **정량화**하여 연산 가능하게 만들었다.

3. AI의 최적화

AI는 **크로스 엔트로피**를 지표 삼아, 정답과의 오차(불확실성)를 0으로 줄여가는 최적의 해를 찾는다.

| Image Sources



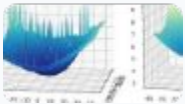
<https://images.fineartamerica.com/images/artworkimages/mediumlarge/2/mathematician-claude-e-shannon-alfred-eisenstaedt.jpg>

Source: fineartamerica.com



https://img.freepik.com/premium-photo/digital-data-stream-blue-numbers-code-flowing-dark-space-abstract-background-technology-digital-information-concept-with-binary-number-black-screen_1271306-5690.jpg

Source: www.freepik.com



https://miro.medium.com/v2/resize:fit:1200/1*JDZSbovagGC3aquY9oJqwQ.jpeg

Source: medium.com