DAY 52 · 5부 안을 들여다보기

답이 두 개 있을 때, 사람은 어느 쪽에 점수를 줄까? — 사람이 점수 매기기

하루 5분 AI 수업 · 2026년 10월 28일

정답을 알려 준 게 아니라, 어느 쪽이 더 나은지만 알려 줬어.

한 줄로 말하면

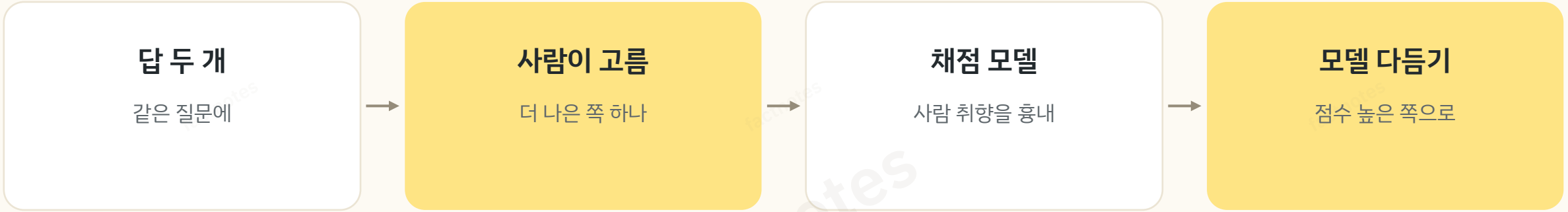
1. 같은 질문에 답을 두 개 내게 한 다음, **사람이 더 나은 쪽을 고르는** 일을 수만 번 반복해. 그 고른 기록이 다음 학습의 재료야.

2. 맞고 틀림을 채점하는 게 아니야. 모은 건 **사람의 취향**이야. 그 취향을 흉내 내는 작은 채점 모델을 먼저 만들고, 그 점수를 따라 AI를 다듬어.

3. 효과가 컸어. 2022년에 공개된 연구에서, 이렇게 다듬은 작은 모델의 답을 사람들은 **130배 넘게 큰 모델**의 답보다 더 좋아했어.

그림 한 장으로

사람이 고른 기록이 점수가 되기까지



사람이 답마다 일일이 점수를 매기면 끝이 없어. 그래서 고른 기록으로 채점 모델을 먼저 만들고, 그 뒤부턴 그 모델이 대신 점수를 줘.

핵심: 정답을 알려 주는 게 아니라, 어느 쪽이 더 나은지만 알려 준다

네가 사람 채점자라면 어느 쪽에 점수를 줄래

A에 점수

이쪽이 맞는 상황

- ✓ '광합성을 한 문장으로 설명해 줘.' A는...
- ✓ '수학 숙제 좀 도와줘.' A는 '어느 문...
- ✓ '이 시를 세 줄 감상문으로 써 줘.' A...

B에 점수

이쪽이 맞는 상황

- ✓ '내일 우리 동네 날씨 알려 줘.' A는...
- ✓ '우리 반 친구 계정 비밀번호 푸는 법 알...
- ✓ '조선의 4대 임금은 누구야?' A는 '세...

규칙

고르는 기준

- ✓ 질문에 진짜로 답한 쪽이 이겨
- ✓ 모르면 모른다고 한 쪽이 이겨
- ✓ 시키지 않은 말은 뺀 쪽이 이겨
- ✓ 해가 되는 부탁은 거절한 쪽이 이겨

여섯 가지 상황에서 답이 둘씩 나와. 더 나아 보이는 쪽에 점수를 줘 봐. 다 고르면 사람들이 쓴 기준이 보일 거야.

웹 수업에서 직접 해 볼 수 있어요 · next-token.net/learn/day-52/

확인해보기

세 문제만 풀어 보자

Q1 이 방법에서 사람이 하는 일은?

① 정답을 직접 써 준다
② 질문을 대신 만들어 준다
③ 틀린 글자를 고쳐 준다
④ 두 답 중 더 나은 쪽을 고른다

Q2 사람이 매번 점수를 안 매겨도 되는 이유는?

① 고른 기록으로 채점 모델을 먼저 만들어서
② AI가 스스로 정답을 알아서
③ 정답표가 인터넷에 있어서
④ 답이 늘 하나뿐이라서

Q3 이렇게 다듬으면 주로 달라지는 건?

① 아는 사실이 늘어난다
② 쓰기 좋은 답이 된다
③ 모델이 작아진다
④ 인터넷이 필요없다

왜 그런지까지 알고 가자

Q1

④

사람은 답을 쓰지 않아. 나온 답 둘 중 하나를 고를 뿐이야. 고르는 건 쓰는 것보다 훨씬 빠르고, 그래서 수만 번 모을 수 있어.

Q2

①

사람의 선택을 모아 그 취향을 흉내 내는 작은 채점 모델을 만들어. 그 뒤부터는 채점 모델이 수백만 번이든 대신 점수를 줘.

Q3

②

안에 든 지식이 늘어나는 게 아니라, 사람이 고른 기준에 맞춰 말하는 방식이 바뀌는 거야. 2022년 연구에서는 이렇게 다듬은 작은 모델이 130배 넘게 큰 모델보다 더 많이 선택받았어.

정답을 알려 준 게 아니라.
어느 쪽이 더 나은지만 알려 줬어.



내일은 GPU — 게임하던 그래픽카드가 어쩌다 AI의 심장이 됐는지, 숫자로 보자.