DAY 15 · 2부 여기까지 온 길

크기만 키웠는데 왜 새 재주가 생겼을까 — 2020 GPT-3

하루 5분 AI 수업 · 2026년 9월 21일

뼈대는 그대로, 크기와 읽은 글만 키웠더니 예시만 봐도 따라 하는 재주가 생겼어.

한 줄로 말하면

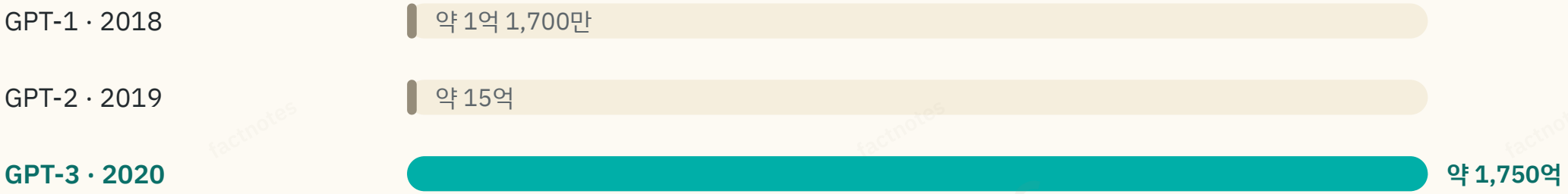
1. 어제 배운 트랜스포머 뼈대를 바꾸지 않고 그냥 크게 키웠더니, 2020년에 숫자 **1,750억 개**짜리 모델이 나왔어. 이름이 GPT-3야.

2. 놀라운 건 크기가 아니라 **배우는 방식**이었어. 다시 훈련시키지 않고 질문 안에 예시 두세 개만 끼워 넣어도 바로 따라 했거든.

3. 그래서 논문 제목도 **'언어 모델은 예시 몇 개만 보면 된다'**였어. 이때부터 '뭐라고 시키느냐', 즉 프롬프트가 중요해졌지.

그림 한 장으로

3년 동안 GPT가 커진 정도



앞의 두 막대는 같은 자로 재면 거의 보이지도 않아. 그만큼 한 번에 뛴 거야.

핵심: 뼈대는 그대로 두고 크기와 읽은 글만 키웠는데 새 재주가 나타났다

자릿수 감각 — GPT는 얼마나 커진 걸까

약 1억 1,700만

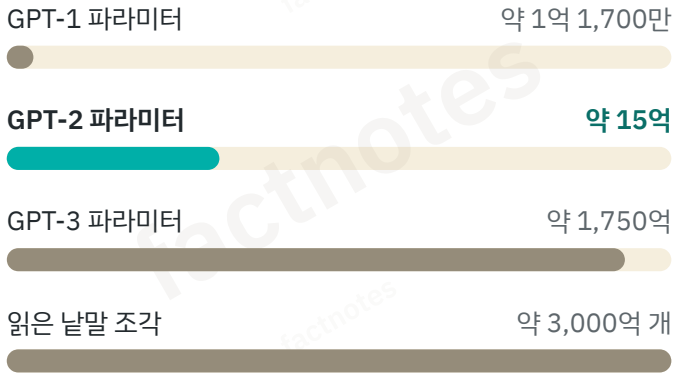
GPT-1 파라미터



이때는 이 정도가 '큰 모델'이었어.

약 15억

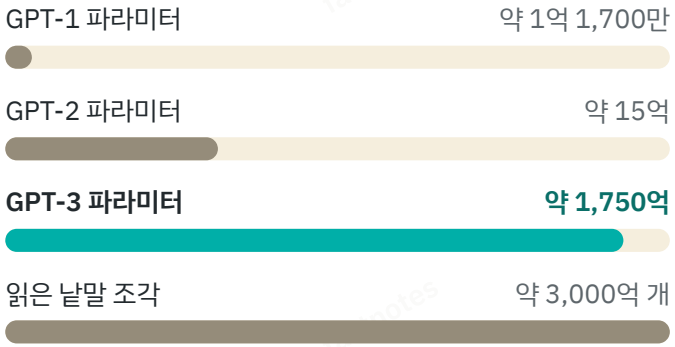
GPT-2 파라미터



1년 만에 자릿수가 한 칸 넘게 뛰었어.

약 1,750억

GPT-3 파라미터



GPT-2의 100배가 넘어. 뼈대는 그대로인데 크기만 이만큼 키운 거야.

손잡이를 움직여 '10의 몇 제곱쯤일까'를 먼저 어림해 봐. 정답 확인을 누르면 실제 숫자가 나와. 맞히는 게 목적이 아니라...

웹 수업에서 직접 해 볼 수 있어요 · next-token.net/learn/day-15/

확인해보기

세 문제만 풀어 보자

Q1 GPT-3가 그 전 모델들과 가장 달랐던 점은?

① 인터넷을 실시간으로 검색했다
② 예시 몇 개만 보여주면 새 일을 바로 따라 했다
③ 사람이 규칙을 하나하나 넣어 줬다
④ 혼자서 전원을 켜고 껐다

Q2 GPT-3가 가진 숫자(파라미터)는 약 몇 개일까?

① 약 1억 7천만
② 약 17억
③ 약 1,750억
④ 약 175억

Q3 질문 안에 예시를 끼워 넣어 시키는 것을 뭐라고 부를까?

① 업데이트
② 다운로드
③ 새로고침
④ 프롬프트

왜 그런지까지 알고 가자

Q1

②

다시 훈련시키지 않아도 질문 안 예시만 보고 따라 했어. 이걸 '예시 몇 개만 보면 된다'고 해서 논문 제목이 그렇게 붙었지.

Q2

③

1,750억 개야. GPT-2가 약 15억이었으니 100배가 넘게 커진 셈이지.

Q3

④

모델에게 건네는 말 전체를 프롬프트라고 불러. GPT-3부터 이걸 어떻게 쓰느냐가 결과를 크게 바꿨어.

뼈대는 그대로, 크기와 읽은 글만 키웠더니 예시만 봐도 따라 하는 재주가 생겼어.



내일은 2022년 ChatGPT — 모델보다 '대화창'이 세상을 바꾼 이야기.