DAY 01–10 · 1부 AI가 뭔지

1부 총정리 — AI가 뭔지

같은 열흘을, 왜 그런지까지 · 설명 한 겹 더

시험 없음 — 읽고 넘어가는 복습 자료

한 조각씩 고르고, 붙이고, 또 고른다

1. AI는 문장을 통째로 지어내지 않아. 다음에 올 조각 하나만 고르고, 그걸 붙인 다음 또 고른다.

2. 이 고르기를 수백 번 반복한 결과가 우리가 읽는 긴 답이야.

3. 그래서 AI는 어딘가에 있는 답을 찾아오는 게 아니라, 매번 새로 만들어 내는 거야.

왜 하필 한 조각씩일까

1. 학습할 때 시킨 일이 딱 그거였기 때문이야 — 문장 일부를 가리고 다음에 올 조각을 맞히기.

2. 잘 맞히는 법을 배운 기계는, 답할 때도 배운 그 방식으로 움직인다.

3. 답이 한 글자씩 흘러나오는 것처럼 보이는 것도 이 때문이야. 실제로 그렇게 만들어지고 있으니까.

그 조각의 이름, 토큰

1. AI가 고르는 그 조각에 이름이 있어. 토큰이라고 불러.

2. 토큰은 글자도 단어도 아니야 — 자주 붙어 다니는 덩어리는 통째로, 낯선 덩어리는 잘게 자른 조각이야.

3. '오늘 급식 뭐야?'가 사람 눈엔 세 어절이지만, AI 눈엔 다섯 조각쯤 된다.

토큰 수는 곧 시간이고 돈이다

1. 자르는 규칙을 토크나이저라고 해. 모델마다 규칙이 달라서 같은 문장도 토큰 수가 다르게 세어져.

2. 토큰이 많아지면 계산할 것도 그만큼 늘어나 — 느려지고 비싸진다.

3. 한국어는 영어보다 같은 뜻을 담는 데 토큰이 더 많이 드는 경우가 흔해.

확률표와 온도 손잡이

1. 다음 조각을 고를 때 AI는 후보마다 점수를 매기고, 그 점수대로 제비를 뽑아.

2. 온도는 그 제비뽑기를 얼마나 모험적으로 할지 정하는 손잡이야.

3. 낮추면 1등만 계속 뽑히고, 올리면 아래 후보도 자주 튀어나온다.

언제 낮추고 언제 올릴까

1. 같은 질문에 매번 다른 답이 나오는 건 기분이 달라서가 아니라 제비뽑기라서야.

2. 정답이 하나인 일(계산·요약·코드)에는 낮추는 게 낫고, 아이디어를 여러 개 받고 싶을 땐 올리는 게 나아.

3. 온도를 올린다고 똑똑해지지 않아 — 다양해질 뿐이고, 엉뚱해질 위험도 같이 커진다.

학습은 조금씩 고치기의 반복

1. AI의 학습은 깨달음이 아니라 '조금씩 고치기'의 반복이야.
2. 문장 일부를 가리고 다음 조각을 맞혀 보게 한 다음, 틀린 만큼 속의 숫자를 살짝 고쳐.
3. 이걸 어마어마하게 반복하면 문법도 상식도 그 숫자들 안에 녹아든다.

숙제 채점에 가까운 일

1. 틀린 문제마다 '이만큼 틀렸다'는 점수를 내고, 그 방향으로 조금 움직이는 것에 가까워.

2. 한 번에 크게 고치지 않아. 조금씩, 대신 아주 여러 번.

3. 그래서 학습에는 엄청난 시간과 전기가 들고, 한 번 끝나면 그 뒤로 숫자는 고정된다.

고쳐지는 숫자의 개수, 파라미터

1. 학습으로 고쳐지는 그 숫자들의 개수가 파라미터야.

2. 많을수록 담을 수 있는 패턴의 그릇이 커진다.

3. 요즘 큰 모델은 수백억에서 1조 개 단위야 — 세어 볼 수 있는 숫자가 아니야.

이 숫자가 얼마나 큰가

1. 1초에 하나씩 센다면 1천억 개를 세는 데 3천 년이 넘게 걸려.

2. 사람 뇌의 시냅스가 약 100조 개니까, 아직 거기엔 못 미친다.

3. 그런데 숫자가 크다고 다 똑똑한 건 아니야 — 무엇을 얼마나 학습했는지가 같이 따라와야 해.

자신 있게 지어내는 일 — 환각

1. AI가 없는 사실을 태연하게 지어내는 걸 환각이라고 불러.

2. 거짓말하려는 게 아니야. 그럴듯한 다음 조각을 고르는 구조에서 당연히 따라오는 일이야.

3. AI는 '나는 이걸 모른다'를 표시하는 장치를 따로 갖고 있지 않아.

특히 잘 틀리는 자리

1. 학습 데이터에 드물었던 것, 아주 최근 일, 그리고 숫자·날짜·이름 같은 구체적인 사실에서 잘 틀려.

2. 말투는 아는 것과 모르는 것이 똑같아. 자신 있어 보인다고 사실인 게 아니야.

3. 출처를 물어보고 그 출처를 직접 열어보는 습관이 가장 싸고 확실한 대비야.

기억이 아니라 다시 읽기

1. AI는 대화를 저장해 두지 않아. 매번 지금까지의 대화를 통째로 다시 읽고 답해.

2. 한 번에 읽을 수 있는 분량에 한계가 있는데, 이걸 컨텍스트 길이라고 불러.

3. 대화가 길어지면 앞부분이 창밖으로 밀려난다.

길어도 문제, 가운데도 문제

1. '아까 말했잖아'가 안 통할 때가 있어. 그 말이 이미 창밖으로 나갔을 수 있거든.

2. 길이가 남아 있어도 문제야 — 가운데 있는 내용이 앞뒤보다 덜 반영되는 경향이 확인됐어.

3. 정말 중요한 정보는 대화 뒷부분에서 한 번 더 짚어주는 게 가장 확실해.

질문을 바꾸면 확률표가 바뀐다

1. 프롬프트는 AI에게 건네는 입력 문장 전체야.

2. 문장이 달라지면 다음 조각의 확률표 자체가 달라져 — 그래서 답이 확 바뀐다.

3. AI를 바꾸는 게 아니라 질문을 바꾸는 거야.

잘 통하는 세 가지

1. 역할을 주고, 조건을 적고, 원하는 형식을 말하기 — 이 셋이 가장 잘 통해.

2. '좋은 글 써줘'보다 '중학생이 읽을 300자 안내문, 예시 한 개 포함'이 훨씬 쓸모 있는 답을 끌어낸다.

3. 속이는 기술이 아니야. 내가 원하는 걸 더 또렷하게 말하는 것뿐이야.

크기는 파라미터 수다

1. 모델이 크다는 건 5일차에 본 파라미터가 많다는 뜻이야.

2. 큰 모델은 아는 게 많고 어려운 문제를 버티지만, 느리고 비싸고 서버에서만 돈다.

3. 작은 모델은 내 기기 안에 들어가서 빠르고, 인터넷이 없어도 돌아.

어느 일을 어느 쪽에 맡길까

1. 급하고 자주 하는 가벼운 일은 작은 쪽, 오래 생각하고 많이 읽어야 하는 일은 큰 쪽.

2. 큰 모델이 만든 답을 교재 삼아 작은 모델을 가르치는 방법도 있어.

3. 그러면 작은 모델이 제 덩치보다 잘하게 되지만, 큰 모델을 그대로 대신하진 못해.

알맹이가 어디에 있느냐

1. 닫힌 모델은 알맹이가 회사 서버 안에만 있어. 우리는 창구로 말만 걸 수 있어.

2. 열린 모델은 그 알맹이, 곧 가중치를 통째로 내려받아 내 컴퓨터에서 돌릴 수 있어.

3. 묻는 자리는 같아도, 모델이 사는 자리가 다른 거야.

'열렸다'를 네 조각으로

1. '열렸다'는 한 덩어리가 아니야. 가중치·돌리는 코드·학습 데이터·고쳐 쓸 권리, 네 조각으로 쪼개 봐야 해.

2. 가중치와 코드는 열어도 학습 데이터는 안 여는 쪽이 대부분이야.

3. 그래서 열렸다는 말만 듣지 말고, 어느 조각이 열렸는지를 따져 물어야 한다.

열흘은 하나의 이야기였다.
다음 조각을 고르는 기계.



2부에서는 이 물건이 어디서 왔는지 — 1958년, 스스로 배우는 인공 뉴런부터 거슬러 올라간다.