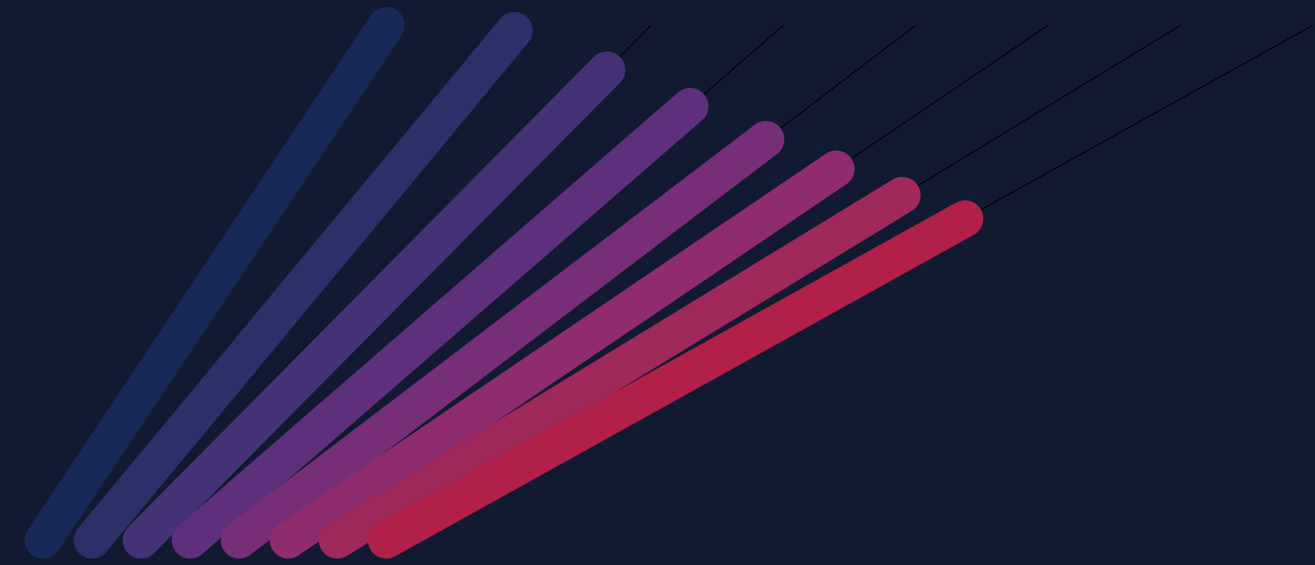




AX 디자인팩토리 부트캠프

청년도약 인재양성 부트캠프 · 온라인 과정

AI는 왜 "모른다"고 하지 않나



1강 · AI는 왜 "모른다"고 하지 않나

과목	차시	길이
생성형 AI와 디지털 리터러시	1 / 5	28분

제작 디벨로켓

도입

오늘 정하는 것

이 편이 끝나면 한 문장에 답할 수 있어야 합니다.

Q 같은 질문에 AI 답이 달라졌을 때, 나는 무엇을 의심해야 하는가

답을 못 하면 AI가 써 준 문장을 그대로 믿게 됩니다.

도입

이 편이 끝나면 보는 것

재현성 테스트 기록표 — 질문 3개 · 답 3개씩 · 판정 칸 (재구성)

이 편 시연 설계(질문 3개·반복 3회) — 집필자 구성 · 확인일 2026-09-04

이번 차시에서만 다루는 것

도구

Antigravity, 프리뷰 무료

- 구글 계정으로 로그인
- 도구 소개는 이 과목에서 이번 한 번뿐

막히면

대처 경로 셋

- 나눠서
- 다른 도구로
- 텍스트로

확률적 생성과 학습 컷오프

이론

7분

소재 없이 - 다음 글자가 어떻게 뽑히는가

AI는 답을 찾지 않는다. 다음에 올 확률이 가장 높은 토 큰을 하나씩 잇는다.

문장 전체가 참인지 거짓인지 따로 판단하는 절차는 없습니
다.

“모릅니다”라고 답하는 것과 답하지 않는 것

모르면 멈추는 사람

확신이 없으면 보류한다

- 아는 것과 모르는 것을 구분한다
- 근거가 없으면 "모릅니다"라고 말한다

확률로 있는 모델

그렇듯하면 멈추지 않는다

- "모르겠습니다"도 확률로 계산된 문장 하나일 뿐
- 자기 확신도를 스스로 모른다

답이 흔들리는 이유 넷

- 1 다음 토큰 예측 — 문장을 통째로 판단하지 않는다
- 2 확률적 생성 — 같은 질문도 매번 같은 문장이 아니다
- 3 학습 컷오프 — 그 시점 이후 바뀐 사실은 모른다
- 4 자기 확신도를 모름 — 확신이 낮다는 신호를 스스로 못 낸다

같은 질문을 세 번 던진다

시연

14분

같은 질문을 반복해 답을 대조한다

재현성 테스트 — 질문 1·2·3 (재구성)

"영양사 국가시험 응시 자격은" 3회 — 이수 과목 수가 16 vs 18로 같림 (정답 18과목 52학점 — 출처는 아래) → "식품기사 실기 시험 과목은" 3회 — 2025년 개편 (식품안전기사·실기 "식품안전관리 실무") 반영 여부로 같림 → "식품 관련 공공기관 채용 공고는 어디서 보나" 3회 — 표준 경로를 말한 답은 하나뿐

동의대·세명대 공지(영양사 응시자격) · 모아바·건국대 안내(식품안전기사 실기) · 잡알리오(공공기관 채용정보시스템) · 확인일 2026-09-04. 화면에서 연 페이지(정적 확인 불가): 국시원·큐넷.

시연 · 판정

아홉 개 답의 판정 (재구성)

질문	완전히 맞음	부분 (방향은 맞음)	지어냄·컷오프로 틀림
응시 자격	2	0	1 (교과목 수 지어냄)
실기 시험 과목	1	0	2 (컷오프 이전 이름)
채용 공고 경로	1	2	0

집필자 구성(정답 대조 근거: 위 demo 강의 대장 출처) · 확인일 2026-09-04

검색 함정 (재구성)

검색 기능을 켜 채 질문 2를 다시 물으면 자격명은 찾아오지만 이름이 비슷한 식품산업기사 페이지를 같이 읽고 그 실기 방식(작업형·약 네 시간)을 식품안전기사의 답인 것처럼 말한다

모아바·건국대 안내(식품안전기사) · 나무위키·gunsys(식품산업기사) · 확인일 2026-09-04. 화면에서 연 페이지(정적 확인 불가): 큐넷 종목 상세(jmCd=1530·2450).

근거가 없는 자리

정리

같은 질문에 답이 달라지는 자리가, 근거가 없는 자리입니다.

다음 차시

그럼 지시를 잘 쓰면 되나요

배웠는지 재는 질문 — 이 답에서 근거를 확인한 문장은 몇 개입니까?

