



Solar-Open-100B vs GLM-4.5-Air 모델 파생 논쟁의 포렌식

- 날짜: 2026-01-01
- 기획: 페블러스 데이터커뮤니케이션팀
- AI 참조: ChatGPT, Gemini, Claude**
- 인터랙티브: <https://blog.pebblous.ai/>

1. 요약 (Executive Summary)

본 보고서는 대한민국 AI 기업 Upstage의 **Solar-Open-100B** 모델이 중국 Zhipu AI의 **GLM-4.5-Air** 모델에서 파생되었는지에 대한 상반된 기술적 분석을 통합적으로 검토한다.

시오닉 AI(Sionic AI)는 통계적으로 불가능한 수준의 '**LayerNorm 유사도(182 시그마)**'를 근거로 파생 설을 주장하는 반면, 현웅 고(Hyunwoong Ko) 교수 및 커뮤니티 측은 최신 LLM 아키텍처(MoE + RMSNorm)의 '**구조적 수렴성**'을 근거로 독립설을 주장한다.

본 분석의 결론은 "누가 맞느냐"가 아니라, 과학적 판정을 위해 검증의 초점을 **저정보 텐서(LayerNorm)**에서 **고정보 텐서(Attention/MoE)**로 이동시켜야 한다는 것이다. 논쟁은 단 하나의 질문으로 귀결된다.

"시오닉 AI가 발견한 '선택적 보존 패턴'이 모델의 지능을 담당하는 고정보 텐서(Attention, MoE Experts)에서도 재현되는가?"

2. 배경 (Background)

2.1 분석 대상 모델

세 모델 모두 **Mixture of Experts (MoE)** 아키텍처와 **RMSNorm**을 정규화 기법으로 채택하고 있다. 이는 분석 결과에 구조적 편향(Inductive Bias)을 줄 수 있는 중요한 공통점이다.

모델명	개발사	파라미터 (추정)	아키텍처 특징
Solar-Open-	Upstage (한		MoE, 48 Layers, DUS 기법 적용 추

100B	국)	~100B	정 1
GLM-4.5-Air	Zhipu AI (중국)	106B	MoE, 46 Layers, 96 Heads 3
Phi-3.5-MoE	Microsoft (미국)	~42B	MoE, 32 Layers, 대조군 모델

2.2 논쟁의 출처

- 파생설 (Sionic AI): sionic-ai/solar-vs-glm - 통계적 이상치(Outlier) 탐지에 주력 1
- 독립설 (Hyunwoongko): hyunwoongko/solar-vs-glm-vs-phi - 지표의 타당성(Validity) 검증에 주력 4

3. 주장 재구성 및 상호 비교

두 주장은 서로 다른 '현미경'을 사용하고 있다. 시오닉 AI는 **차이**를 강조하고, 현웅 고 교수는 **보편성**을 강조한다.

구분	Sionic AI (파생설)	Hyunwoongko (독립설)
핵심 주장	Solar는 GLM의 파생작이며, 증거는 동일 레이어 LayerNorm의 비정상적 유사도(≈ 0.989) 와 선택적 보존이다.	RMSNorm은 정보량이 낮고 구조적 제약이 강하므로, 높은 Cosine 유사도는 아키텍처적 필연(수렴) 이지 표절의 증거가 아니다.
중심 증거	<ul style="list-style-type: none">• 동일 레이어 간 LayerNorm Cosine ≈ 0.99• 모델 내부(Within-model) 베이스라인 (0.377) 대비 압도적 차이 (182σ)• LN은 보존되고 Attention은 재학습된 "선택적 보존" 패턴	<ul style="list-style-type: none">• 제3의 모델(Phi-3.5-MoE)과 GLM 간에도 0.9+ 유사도 발생• DeepSeek, Mistral 등 타 모델 간에도 LN 유사도 0.99 관측됨• RMSNorm의 벡터 방향성 편향 특성
맹점/한계	<ul style="list-style-type: none">• 결정적 증거가 저정보 파라미터 (LayerNorm) 에 과도하게 의존함.• RMSNorm의 수학적 특성(Scale Invariance)을 고려하지 않음.	<ul style="list-style-type: none">• Sionic이 제시한 Attention/MoE/Embedding의 '0'에 가까운 유사도(재학습 증거)'를 직접 반박하지 않음.• "선택적 보존"이 왜 발생하는지에 대한 설명 부재.

4. 상호 검증이 필요한 핵심 질문

논쟁을 해결하기 위해서는 각 진영의 주장을 교차 검증해야 한다.

- 파생설 검증 질문:** "선택적 보존 패턴(LN은 같고 나머지는 다름)이, RMSNorm의 수학적 특성을 제거하고 고정 정보 텐서(Attention Weight)를 비교했을 때도 유지되는가?"
- 독립설 검증 질문:** "RMSNorm을 배제했을 때, Solar와 GLM 사이의 거리가 Solar와 Phi 사이의 거리보다 통계적으로 유의미하게 가까운가?"

이 두 질문은 사실상 동일한 실험을 요구한다.

5. 두 주장을 결합해야 하는 이유: 과학적 포렌식

- Sionic AI**는 '패턴 탐지'에 강점이 있으나, 그 패턴이 '도둑질의 흔적'인지 'RMSNorm의 특성'인지 구분하지 못했다.
- Hyunwoongko**는 '지표 비판'에 강점이 있으나, 비판을 넘어선 결정적 반증(고정 정보 텐서 비교)을 제시하지 않았다.

결합 검증 프레임워크:

Sionic의 '선택적 보존' 가설을 유지하되, 대상 변수를 RMSNorm에서 Attention / MoE / FFN으로 변경하여 Solar-GLM-Phi 3자 비교를 수행한다.

- 결과 A (파생설 강화):** 고정 정보 텐서에서도 Solar-GLM 유사도가 타 모델 대비 유의미하게 높게 측정됨.
- 결과 B (독립설 확정):** 고정 정보 텐서에서는 모든 모델 간 유사도가 낮거나 비슷하게 측정됨.

6. 각 주장의 세부 데이터 분석

6.1 Sionic AI의 파생설: "182 시그마의 충격"

시오닉 AI의 분석은 텐서 유형별로 극단적인 유사도 차이를 보여준다.

표 1: Sionic AI가 제시한 텐서별 유사도 1

텐서 유형	Cosine 유사도	시오닉 측 해석	비고
input_layernorm	0.949	원본 보존	구조적 뼈대
post_attention_layernorm	0.986	원본 보존	구조적 뼈대

k_proj / v_proj	0.001	재학습됨	Attention (지능)
mlp.gate	0.004	재학습됨	MoE Router (지능)
embed_tokens	0.002	재학습됨	어휘 확장 1

논리 구조: 독립적으로 학습된 모델이라면 우연히 0.99 유사도가 나올 수 없다($P\text{-value} < 10^{(-1000)}$). 따라서 뼈대(LN)를 가져다 놓고 근육(Attn/MLP)만 갈아끼운 것이다.

6.2 Hyunwoongko의 독립설: "RMSNorm의 함정"

현웅 고 교수는 제3의 모델인 Phi-3.5-MoE를 도입하여 이 '유사도'가 허상임을 증명했다.

표 2: 10번째 레이어 LayerNorm Cosine 유사도 매트릭스 (재구성) 4

구분	Solar-Open	GLM-4.5-Air	Phi-3.5-MoE
Solar-Open	1.000	0.9+	0.9+
GLM-4.5-Air	0.9+	1.000	0.9+
Phi-3.5-MoE	0.9+	0.9+	1.000

핵심 관찰: 서로 전혀 다른 모델(Phi vs GLM)끼리도 LayerNorm 유사도는 0.9를 넘는다.

원인 분석: RMSNorm은 학습이 진행될수록 특정 값(Scale)으로 수렴하며, 벡터의 방향성이 유사해지는 '구조적 수렴(Convergence)' 현상이 발생한다. 따라서 LN 유사도는 모델의 '지문'이 될 수 없다.

7. 결론: 단일 검증 질문으로의 수렴

현재 상황은 한쪽은 현미경을 들고 "패턴이 보인다"고 외치고, 다른 한쪽은 "그 현미경 렌즈가 왜곡되었다"고 말하는 형국이다. 따라서 렌즈(분석 대상 텐서)를 바꿔서 다시 관찰해야 한다.

결론: 이 논쟁의 최종 판정은 "선택적 보존 패턴이 고정보 텐서(Attention/MoE)에서도 재현되는가?" 라는 질문에 달려 있다. LayerNorm은 기각되었다. 이제 모델의 진짜 '뇌'인 Attention 가중치를 비교해야 한다.

8. 업스테이지(Upstage)의 대응 전략 제언

업스테이지는 현재 "From Scratch(처음부터 학습)"를 주장하며 학습 로그(WandB) 공개를 예고했다. 그러나 이것만으로는 부족할 수 있다. 기술적 의혹은 기술적 데이터로 해소해야 한다.

8.1 전략적 권고: "결정적 실험"의 선제 수행

업스테이지는 수동적인 해명에 그치지 말고, 위에서 도출된 '고정보 텐서 비교 실험'을 직접 수행하여 공개해야 한다.

실행 방안:

1. **비교군 설정:** Solar-Open-100B, GLM-4.5-Air, Phi-3.5-MoE, DeepSeek-V3, Mixtral 8×7B.
2. **분석 대상:** q_proj, k_proj, v_proj, o_proj, gate_proj, up_proj, down_proj. (RMSNorm 제외)
3. **예상 결과 및 대응:**
 - **유사도 낮음 (독립설 입증):** "보시다시피 모델의 지능을 담당하는 99%의 파라미터는 완전히 다르다. LN 유사도는 업계 공통 현상일 뿐이다."라고 선언.
 - **유사도 높음 (위기):** 만약 고정보 텐서에서도 GLM과 특이하게 높은 유사도가 나온다면, 이는 '참조'를 인정하고 데이터셋의 유사성이나 베이스 모델(Llama 등)의 공유 가능성을 설명해야 한다.

8.2 투명성 제고

단순히 "베끼지 않았다"는 주장보다, **개발 타임라인**과 **데이터 파이프라인**을 공개하는 것이 효과적이다. 특히 GLM-4.5-Air 공개 시점 이전에 Solar-Open-100B의 해당 레이어 구조가 확정되었다는 증거(Git Commit Log 등)가 있다면 논란은 즉시 종결될 것이다.

9. 맺음말: 이 논쟁이 남기는 교훈

이번 'Solar vs GLM' 사태는 LLM 오픈 웨이트(Open Weights) 시대에 **모델 포렌식(Model Forensics)**이 얼마나 중요한지를 보여주는 첫 번째 대형 사례다.

- **분석가:** 단일 지표(특히 LayerNorm 같은 저정보 파라미터)에 의존한 성급한 결론은 위험하다.
- **개발사:** 모델 개발 과정의 투명성(WandB, Checkpoints)은 이제 선택이 아닌 '보합'이다.
- **커뮤니티:** "누가 맞나"보다 **"어떤 실험이 과학적으로 타당한가"**를 묻는 것이 건강한 생태계를 만든다.

결국, 이 논쟁의 승자는 가장 목소리가 큰 쪽이 아니라, **"가장 재현 가능한 실험 결과"**를 제시하는 쪽이 될 것이다.

면책조항: 본 보고서는 공개된 두 GitHub 리포지토리(sionic-ai, hyunwoongko)의 데이터를 기반으로 작성되었으며, 특정 주장의 진위 여부를 최종적으로 확정하지 않습니다. 본 보고서는 과학적 검증을 위한 방법론을 제안하는 데 목적이 있습니다.

Works cited

1. Solar-Open-100B vs GLM-4.5-Air: 가중치 파생 분석 - GitHub, accessed January 1, 2026, <https://github.com/sionic-ai/solar-vs-glm>
2. Solar 10.7B: Comparing Its Performance to Other Notable LLMs - Analytics Vidhya, accessed January 1, 2026, <https://www.analyticsvidhya.com/blog/2024/01/solar-10-7b-comparing-its-performance-to-other-notable-llms/>
3. Why GLM-4.5 and GLM-4.5 Air Could Kill DeepSeek's Open Source AI Monopoly - Medium, accessed January 1, 2026, <https://medium.com/towards-agi/why-glm-4-5-and-glm-4-5-air-could-kill-deepseeks-open-source-ai-monopoly-dc68c1370c24>
4. Upstage Denies Copying Chinese Model, Vows Public Verification, accessed January 1, 2026, <https://www.chosun.com/english/industry-en/2026/01/01/PN5Q5LK2CVH3FNBR5KAO3GSIEE/>
5. Upstage Solar-Open-100B Public Validation - LocalLLaMA - Reddit, accessed January 1, 2026, https://www.reddit.com/r/LocalLLaMA/comments/1q0zst6/upstage_solaropen100b_public_validation/

Pebblous

Pebblous Makes Data Tangible

contact@pebblous.ai