

Med-PRM: 단계별 가이드라인 검증 프로세스 보상(Process Rewards)을 통한 의료 추론 모델

Jaehoon Yun^{1,4,5*}, Jiwoong Sohn^{1,2*}, Jungwoo Park^{1,4*}, Hyunjae Kim³, Xiangru Tang³, Yanjun Shao³,
Yonghoe Koo⁶, Minhyeok Ko⁵, Qingyu Chen³, Mark Gerstein³, Michael Moor^{2†}, Jaewoo Kang^{1,4†}

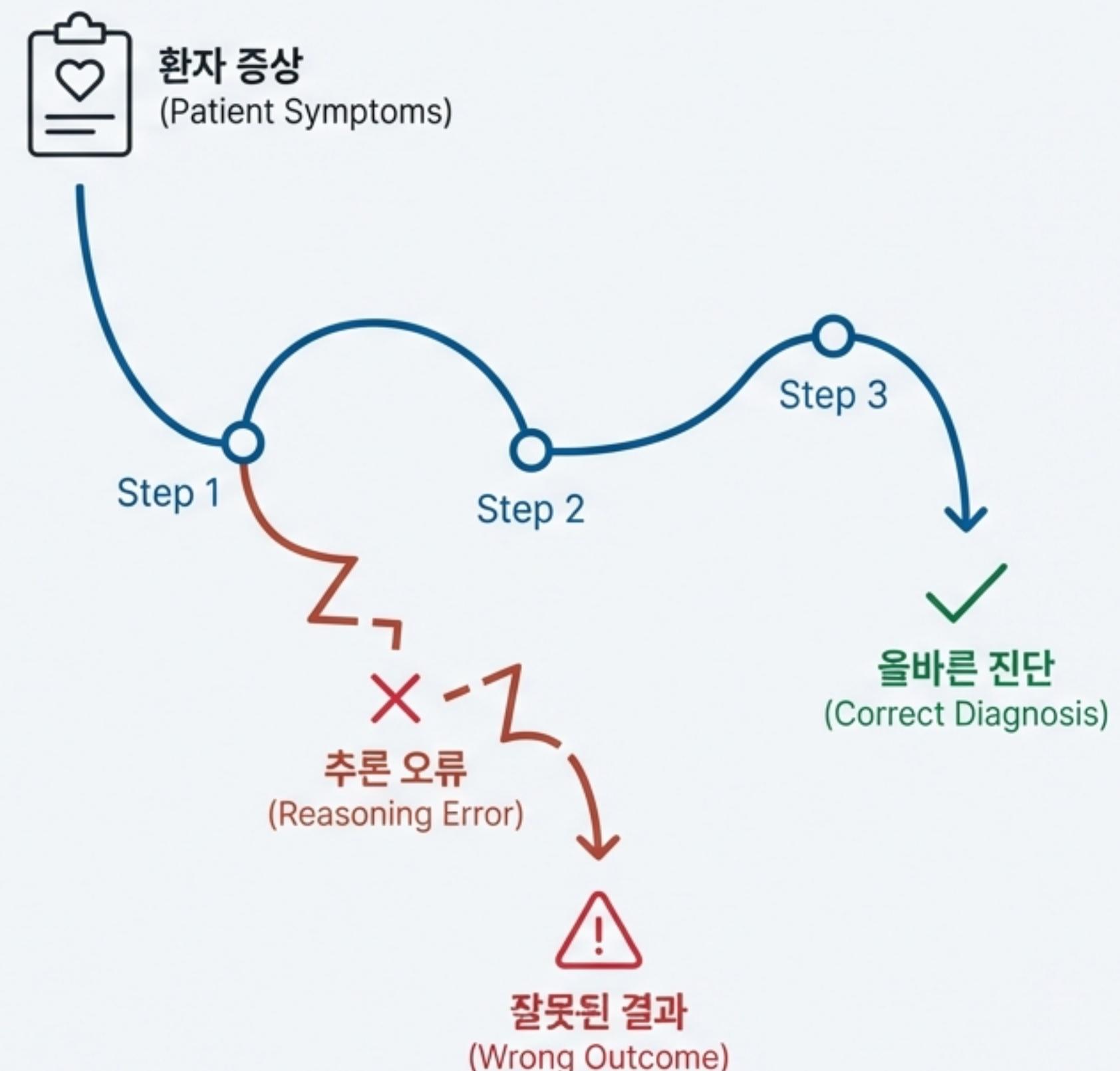
¹고려대학교, ²ETH Zürich, ³Yale University, ⁴AIGEN Sciences,
⁵한양대학교 의과대학, ⁶울산대학교 의과대학

의료 AI의 신뢰성: 정답보다 중요한 '올바른 과정'

임상 의사 결정(Clinical Decision Making, CDM)은 환자의 증상 평가, 임상 근거 검색, 진단 및 치료 전략 수립을 포함하는 복잡한 단계 프로세스입니다. AI가 유망한 성능을 보이고 있지만, 의료 분야에서는 추론 과정의 작은 오류 하나가 치명적인 결과로 이어질 수 있습니다.

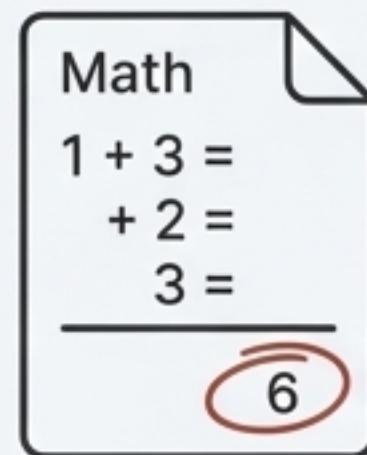
따라서 최종 결과뿐만 아니라, 각 추론 단계가 의학적 기준에 부합하는지 검증하고 피드백하는 메커니즘이 필수적입니다.

안전성, 신뢰성, 표준 치료 절차 준수를 보장하기 위해 단계별 검증은 선택이 아닌 필수입니다.



기존 보상 모델의 한계: 과정은 무시하고 결과만 보거나, 추측에 의존합니다

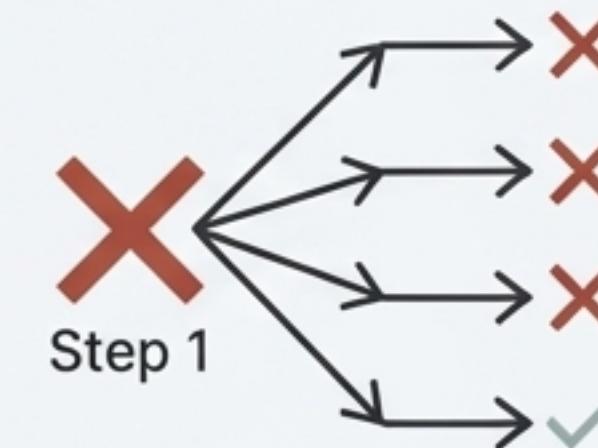
결과 보상 모델 (Outcome Reward Model)



전체 추론 과정의 최종 답이 맞았는지 틀렸는지만 평가합니다.
중간 과정의 옳고 그름은 전혀 고려하지 않습니다.

학생의 풀이 과정은 보지 않고 최종 답만 채점하는 것과 같습니다.

샘플링 기반 프로세스 보상 모델 (MCTS-based PRM)



각 단계에서 무작위로 여러 추론 경로를 생성해보고,
정답에 도달하는 확률을 기반으로 점수를 매깁니다.

정답으로 가는 길이 여러 개일 때, 우연히 시도한 몇 개의 경로가
막혔다고 해서 올바른 첫걸음까지 틀렸다고 판단하는 것과 같습니다.

두 방식 모두 논리적으로 타당한 초기 추론 단계를 부당하게 '오답'으로 판단하는
"False Negative" 문제를 야기합니다.

사례 연구: 올바른 추론이 어떻게 '오답'으로 낙인찍히는가

Medical Question & Reasoning

A 46-year-old woman comes to the physician because of a 2-week history of diplopia when reading the newspaper. She also has a 3-month history of amenorrhea, increased sweating, sweating. She has been overweight all hot flashes, and lost 6.8-kg (15-lb) of weight in the past 2 months. Her pulse is 110 bpm/min, and blood

Step 1 & 2: The patient's symptoms...
point towards thyroid disease.

...

Step 5 & 6: ...incorrectly attributes
ocular symptoms to sympathetic
overactivity... leading to the wrong
answer.

평가 비교: 올바른 'Step 1'에 대한 라벨링

MCTS-based Labeling


$$y_{S1}^{SE} = 0.0$$

이 경로에서 생성된 샘플들이 우연히 정답에 도달하지 못했기 때문에, 논리적으로 완벽한 첫 단계임에도 0점을 받았습니다.

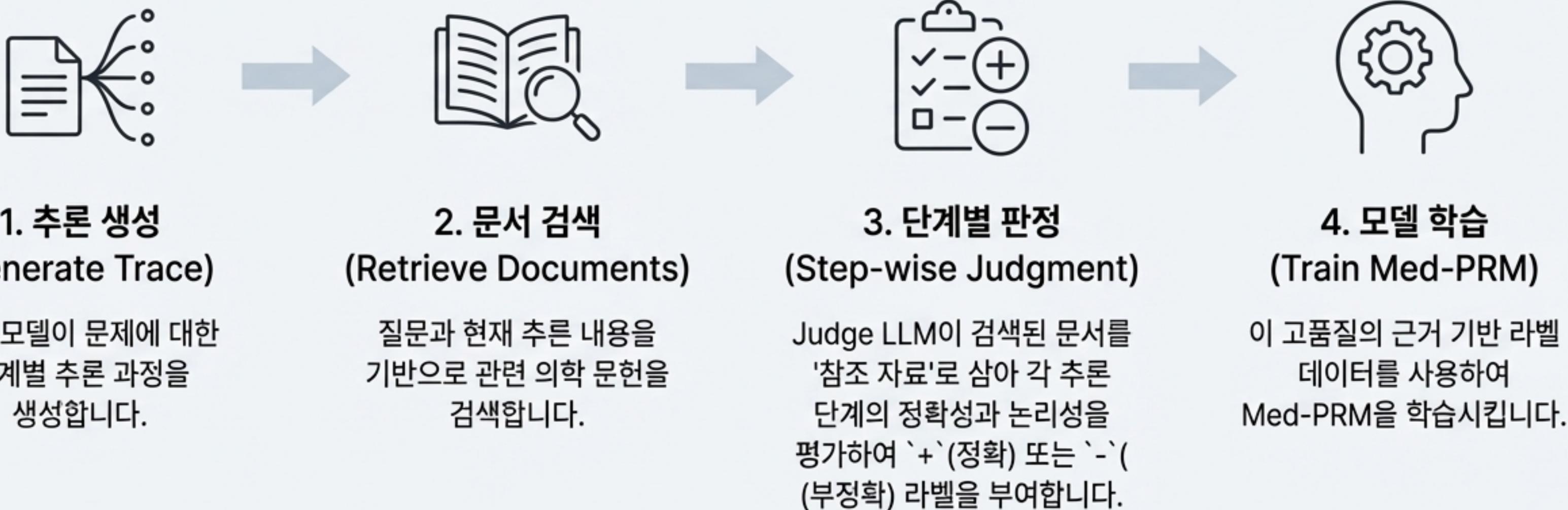
Med-PRM Labeling


$$y_{S1}^{RAG} = 1.0$$

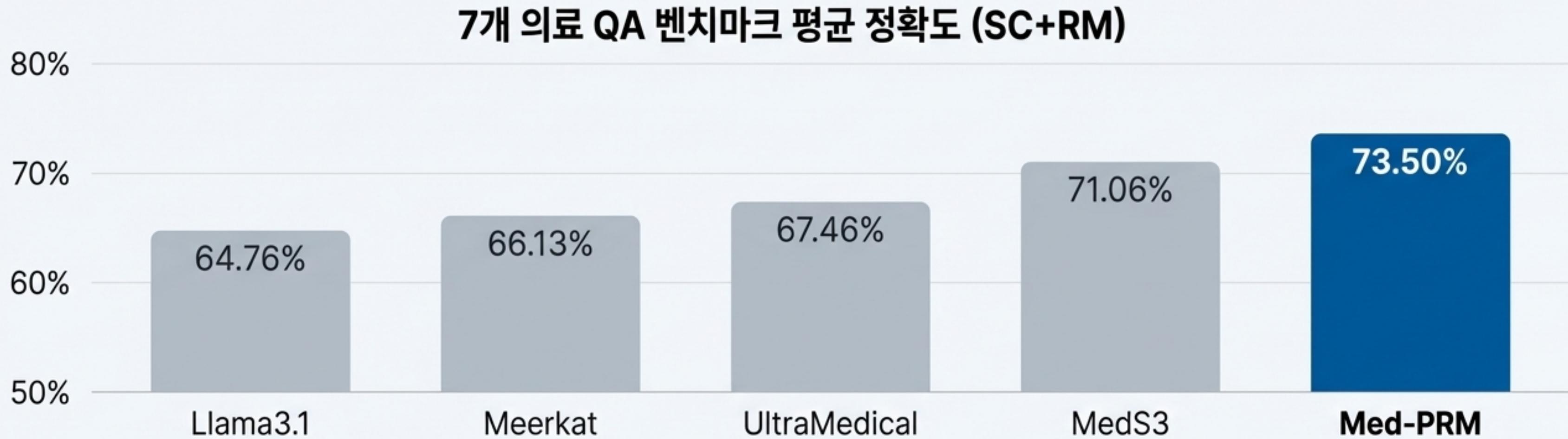
의학 문헌을 근거로 이 단계가 사실에 부합하고 논리적으로 타당함을 검증했습니다.

Med-PRM: 모든 추론 단계를 임상 가이드라인으로 검증합니다

RAG-AS-A-JUDGE: 추론의 각 단계를 평가할 때, 신뢰할 수 있는 외부 의학 지식 베이스(임상 가이드라인, 의학 교과서 등)를 실시간으로 검색(Retrieval)하여 근거 기반으로 판단합니다.



SOTA 달성: 7개 의료 벤치마크에서 기존 모델 성능을 압도합니다



Key Takeaways

- Med-PRM은 7개 벤치마크에서 평균 73.50%의 정확도를 달성하여, 10B 미만 파라미터의 모든 공개 소스 모델 중 최고 성능을 기록했습니다.
- 특히 실제 진단 과정과 유사한 AgentClinic 벤치마크에서는 베이스 모델 대비 최대 12.5%의 성능 향상을 보였습니다.
- 이전 SOTA 프로세스 보상 모델인 MedS3 대비 평균 2.44% 더 높은 성능을 달성했습니다.

Plug-and-Play 검증기: 강력한 정책 모델의 성능을 한 단계 끌어올립니다

Med-PRM은 학습된 정책 모델과 무관하게, 추론 시점에 '검증기(Verifier)'로 작동하여 일관된 성능 향상을 제공합니다.

Model: Meerkat-8B (강력한 의료 특화 모델)

Base + Self-Consistency: **MedQA 76.04%**

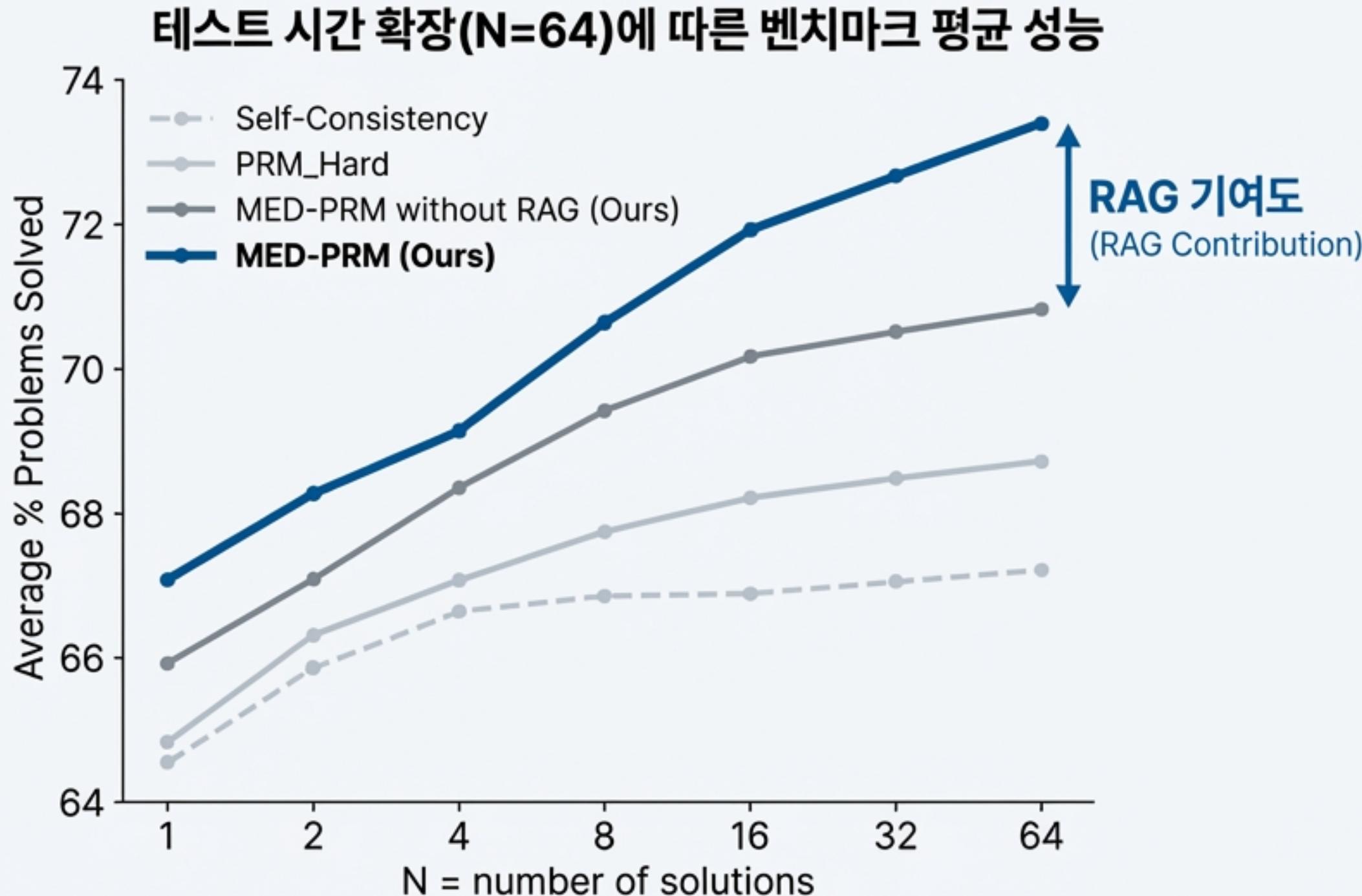
Base + Self-Consistency + Med-PRM Verifier: **MedQA 80.35%**

+13.70%
from base model

8B급 모델 최초로 MedQA 벤치마크 80% 정확도 돌파.

이 결과는 Med-PRM이 다양한 모델에 적용 가능한 범용적인 검증기로서의 가치를 증명합니다.

성능의 핵심 동력: 검색 증강(RAG)의 결정적 역할



Key Takeaways

- 1. LLM-as-a-Judge의 우수성**
RAG가 없는 Med-PRM조차도 기존 샘플링 기반 PRM(PRML Hard)보다 우수한 성능을 보입니다. 이는 LLM을 직접 평가자로 사용하는 것의 효과를 보여줍니다.
- 2. RAG의 압도적 기여**
RAG를 추가했을 때 성능이 가장 큰 폭으로 향상됩니다. 이는 외부 의학 지식을 통한 근거 기반 검증이 Med-PRM 성능의 핵심임을 증명합니다.

임상 전문가와의 높은 정렬성: 어려운 문제일수록 차이가 명확해집니다

Metric: 모델 생성 라벨과 인간 전문가 주석 간의 피어슨 상관관계(Pearson Correlation)

문제 난이도	Med-PRM	Soft Label (MCTS)	Hard Label (MCTS)
쉬운 문제 (Easy)	0.74	0.64	0.70
어려운 문제 (Hard)	0.71	0.34 (▼)	0.31 (▼)

- Med-PRM은 문제 난이도와 상관없이 전문가 판단과 일관되게 높은 상관관계(**0.71 이상**)를 유지합니다.
- 반면, 기존 MCTS 기반 라벨링 방식은 어려운 문제에서 상관관계가 급격히 하락(0.34, 0.31)하며 신뢰도가 떨어집니다.

Med-PRM은 복잡하고 어려운 추론 시나리오에서도 전문가 수준의 일관되고 신뢰도 높은 평가를 제공합니다.

사례 연구: Med-PRM은 어떻게 미묘한 임상적 오류를 포착하는가

양수과다증(Polyhydramnios)이 의심되는 환자 사례. 모델은 후부 요도 판막(posterior urethral valve)이 원인이라고 잘못 추론합니다.

Conventional PRM (No Evidence)

"후부 요도 판막은 소변 생성 증가와 양수과다증을 유발할 수 있다."

$$r_{s5} = 0.86$$

외부 근거 없이는 이 오류를 감지하기 어려워
높은 점수를 부여합니다.

Med-PRM (With Retrieved Evidence)



Retrieved Document Snippet

...posterior urethral valve... can block the outflow of urine... a major concern for oligohydramnios (양수과소증).

"후부 요도 판막은 소변 생성 증가와 양수과다증을 유발할 수 있다."

$$r_{s5} = 0.22 \leftarrow$$

검색된 문헌에서 '양수과소증'을 유발한다는 명확한 근거를 찾아
오류를 정확히 식별하고 낮은 점수를 부여합니다.

Med-PRM은 검색된 문헌을 바탕으로 미묘하지만 결정적인 임상적 오류를 정확히 포착하여 추론의 질을 정밀하게 평가합니다.

Med-PRM: 신뢰할 수 있는 의료 AI를 위한 새로운 패러다임



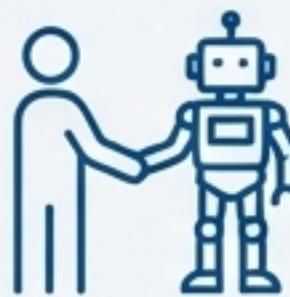
근거 기반의 정밀한 평가 (Evidence-Grounded Assessment)

RAG-AS-A-JUDGE를 통해 추론의 모든 단계를 의학적 근거에 기반하여 평가하는 새로운 프레임워크를 제시합니다.



최고 수준의 성능과 범용성 (State-of-the-Art Performance & Versatility)

7개 벤치마크에서 SOTA를 달성하고, 어떤 모델에도 적용 가능한 강력한 Plug-and-Play 검증기로서의 가치를 입증했습니다.



전문가 수준의 신뢰도 (Expert-Aligned Reliability)

인간 임상 전문가의 판단과 높은 상관관계를 보이며, 기존 방법론의 논리적, 사실적 한계를 극복했습니다.

Med-PRM은 단순한 성능 향상을 넘어, 의료 AI 시스템의 정확성, 투명성,
그리고 신뢰성을 높이는 핵심적인 접근법을 제공합니다.

Code & Data



Med-PRM.github.io

우리의 코드, 데이터, 그리고 모델은 위 주소에서 모두 공개되어 있습니다.

Corresponding Authors

Michael Moor, PhD (ETH Zürich)

Jaewoo Kang, PhD (Korea University)

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF), the Ministry of Health & Welfare, the Ministry of Science and ICT, and the Institute of Information & Communications Technology Planning & Evaluation (IITP).