

Расшифровка клинических корреляций высокого порядка: фреймворк на основе большой языковой модели, управляемой знаниями, для специализированных медицинских решений

Источник: Frontiers in AI — Medicine

Оригинал: <https://www.frontiersin.org/articles/10.3389/frai.2026.1772518>

LLM

LangChain

клинические корреляции

медицинские решения

торакальная хирургия

Актуальность

Специализированные вопросы по торакальной хирургии требуют интеграции многофакторных клинических взаимосвязей в рамках текста, однако большие языковые модели (LLM) общего назначения могут показывать недостаточно высокие результаты на подобных экзаменационных тестах.

Методы

Мы разработали агента DK-LLM, встроив курированные знания из медицинских учебников в архитектуру на основе LangChain для поддержки предметно-ориентированных рассуждений. Модель была протестирована на наборе из 56 экзаменационных вопросов по торакальной хирургии в условиях ограниченного текстового режима без доступа к интернету и внешним инструментам, а также сравнивалась с базовыми LLM общего назначения, тремя торакальными хирургами и тремя неспециалистами-инженерами. Оценивались экзаменационные баллы и структура ошибок.

Результаты

Конфигурация DK-LLM с дополненными знаниями показала преимущество в 11,8 балла по экзаменационному результату по сравнению с версией без локальной базы знаний. Коммерческие LLM-агенты превзошли базовые модели с открытым исходным кодом и участников-неспециалистов на данном наборе вопросов, тогда как опытные торакальные хирурги в целом достигли наивысших результатов; в абляционном анализе удаление локальной базы знаний снизило экзаменационный балл на 11,8 процентных пункта.

Заключение

Встраивание предметно-специфических знаний в LLM может улучшить результаты на специализированных экзаменационных вопросах по торакальной хирургии в рамках данного текстового бенчмарка. Однако настоящая оценка на 56 вопросах не устанавливает клиническую эквивалентность, диагностическую точность в практике, мультимодальную компетентность или готовность к реальной поддержке клинических решений.

Перевод выполнен: 04.09.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.