От специализированных сетей к фундаментальным моделям в лечении диабета: критический обзор возможностей, пробелов в оценке и границ безопасности

Источник: Frontiers in AI — Medicine

Дата публикации: 2025-12-31

Оригинал: https://www.frontiersin.org/articles/10.3389/frai.2026.1983619

NLP | безопасность ИИ | диабет | диагностика | клинические испытания | фундаментальные модели | электронные медкарты

Искусственный интеллект в лечении диабета исторически опирался на специфические для задач методы обучения с учителем, такие как специализированные нейронные сети для оценки ретинопатии, прогнозирования заболеваний почек и прогнозирования уровня глюкозы. Появление фундаментальных моделей (FMs: крупных моделей, предварительно обученных на обширных наборах данных, как правило, с использованием самообучения, и адаптированных для множества последующих задач) вытесняет эту парадигму в различных модальностях, связанных с диабетом: модели офтальмологического профиля, такие как RETFound, предварительно обученные на 1,6 миллиона ретинальных фотографий, показали превосходство над специализированными сетями при оценке ретинопатии и прогнозировании системных заболеваний, особенно в условиях дефицита размеченных данных; трансформерные модели временных рядов заменяют конвейеры на основе долгой краткосрочной памяти (LSTM) для прогнозирования показателей непрерывного мониторинга глюкозы (CGM); клинические большие языковые модели (LLM), такие как Med-PaLM и GPT-4, демонстрируют уровень экспертов в задачах ответов на медицинские вопросы согласно бенчмаркам; а генеративные модели

обеспечивают создание синтетических данных CGM и электронных медицинских карт (EHR). В данном обзоре представлена критическая синтезирующая оценка этого перехода, структурированная по модальностям данных. Мы (i) формализуем таксономию из трех эпох: правила оценки риска (эпоха I), специфическое для задач глубокое обучение (эпоха II, 2016–2020 гг.) и фундаментальные модели (эпоха III, с 2021 г.), и сопоставляем литературу по диабету с этой структурой; (ii) проводим обзор архитектур фундаментальных моделей, целей предварительного обучения и стратегий адаптации в области ретинальной визуализации, CGM, продольных данных EHR и мультимодальных данных; (iii) выявляем кризис оценки: загрязнение бенчмарков при оценке медицинских LLM, повсеместное отсутствие в рассмотренной литературе внешней и проспективной валидации для ретинальных и CGM фундаментальных моделей, некорректную калибровку при сдвиге распределения данных и, насколько нам известно и в рамках рассмотренной литературы (с января 2016 г. по декабрь 2025 г.; негативные результаты чувствительны ко времени), полное отсутствие каких-либо рандомизированных оценок применения фундаментальной модели в рабочем процессе лечения диабета; и (iv) анализируем риски безопасности: галлюцинированные рекомендации по дозировке инсулина, предвзятость автоматизации, унаследованную алгоритмическую несправедливость и риски нарушения конфиденциальности в процессе адаптации. Мы приходим к выводу, что ключевой нерешенной проблемой больше не является обучение представлений, а обоснованное, калиброванное и каузально информированное развертывание, требующее оценки, устойчивой к загрязнению данных, проспективных испытаний и гибридных архитектур, сочетающих представления фундаментальных моделей с оценкой эффекта лечения.

Перевод выполнен: 31.10.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.