

Объяснимый ИИ в здравоохранении: систематический обзор случаев применения XAI в визуализации, диагностике и реабилитации

Источник: Frontiers in AI — Medicine

Дата публикации: 2025

Оригинал: <https://www.frontiersin.org/articles/10.3389/frai.2026.1749527>

диагностика

клиническое принятие решений

машинное обучение

медицинская визуализация

объяснимый ИИ

реабилитация

Фон

Объяснимый искусственный интеллект (XAI — explainable artificial intelligence) используется в здравоохранении для повышения прозрачности и клинической применимости результатов машинного обучения. Это важно, поскольку многие модели машинного обучения работают как «чёрный ящик», что может скрывать предвзятость и снижать доверие к модели. XAI решает эту проблему, показывая, какие признаки или области изображения повлияли на результат — как для отдельного пациента, так и в рамках набора данных.

Цели

Наша цель — предоставить чёткий систематический обзор того, как XAI применяется в здравоохранении. Мы суммируем основные методы XAI, данные и модели, с которыми они сочетаются, а также то, как эти объяснения поддерживают клиническое понимание в области медицинской визуализации, диагностики и реабилитации.

Методы

Мы провели систематический обзор с нарративным синтезом (2020–2025) 36 эмпирических исследований по трём направлениям — Визуализация ($n = 10$), Диагностика ($n = 16$) и Реабилитация ($n = 10$), которые были идентифицированы через PubMed/MEDLINE, IEEE Xplore и Google Scholar в соответствии с рекомендациями PRISMA 2020. Мы включили исследовательские работы, применявшие XAI в трёх указанных направлениях. Мы исключили обзорные статьи и работы с мнением авторов. Количество отобранных записей: идентифицировано 1 481; удалены дубликаты 647; другие удаления 187; отобрано 647; исключено 532; запрошено отчётов 115; не получено 31; проанализировано 84; исключено по полному тексту 48; включено 36. Из каждого исследования мы извлекли модели машинного обучения, методы XAI, дизайн исследования, методологии и источник/набор данных. Мета-анализ не проводился из-за гетерогенности.

Результаты

В рамках 36 исследований SHAP использовался в 21 исследовании, Grad-CAM — в $\sim 12/36$, и LIME — в $\sim 11/36$. Выявилась чёткая связь метода и модальности: в области визуализации преобладали методы saliентности/тепловых карт, особенно Grad-CAM, для пространственных доказательств. Диагностика и реабилитация были доминируемы инструментами атрибуции признаков, такими как SHAP и LIME, для глобальных и индивидуализированных объяснений. Многие работы сочетали ≥ 2 объяснителя для перекрёстной проверки интерпретаций, а именно SHAP+LIME и Grad-CAM + LIME.

Заключение

Недавние разработки XAI в здравоохранении демонстрируют последовательное соответствие метода и модальности и часто сочетают два или более метода, что помогает перевести непрозрачные предсказания в рассуждения, ориентированные на клиницистов. Для обеспечения надёжного внедрения будущие работы должны сочетать эти практики со стандартизированной отчётностью XAI, оценками верности/стабильности и внешней, межучрежденческой валидацией.

Перевод выполнен: 01.04.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.