

Разработка и интерпретируемые модели машинного обучения для классификации риска псевдокисты поджелудочной железы при остром панкреатите

Источник: Frontiers in Digital Health

Дата публикации: 2023-2024

Оригинал: <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1753529>

SHAP

диагностика

интерпретируемый ИИ

машинное обучение

панкреатит

прогнозирование

Введение

Панкреатические псевдокисты (ППК) представляют собой позднее локальное осложнение острого панкреатита (ОП). Персистирующие ППК несут высокий риск тяжелых исходов. Существующие модели, преимущественно основанные на логистической регрессии, демонстрируют ограниченную прогностическую эффективность и не проходили временную валидацию. Настоящее исследование было направлено на разработку и валидацию интерпретируемой модели машинного обучения с использованием рутинно доступных клинических данных для классификации пациентов с ОП по статусу развития ППК.

Материалы и методы

Мы ретроспективно проанализировали 1 184 пациента с ОП, госпитализированных в третичный центр в период с 2018 по 2023 год. Данные за 2018-2022 годы (n = 979) были случайным образом разделены на обучающую (70%, n = 685) и внутреннюю тестовую (30%, n = 294) выборки,

тогда как когорта 2023 года ($n = 205$) служила независимым набором для временной валидации. Кандидатные предикторы — включая демографические характеристики, клинический анамнез и рутинные лабораторные параметры — были отобраны с помощью однофакторного анализа и дополнительно отобраны с использованием регрессии LASSO для устранения мультиколлинеарности. Были разработаны и сравнены девять алгоритмов машинного обучения: логистическая регрессия, дерево решений, случайный лес, искусственная нейронная сеть, метод опорных векторов, K-ближайших соседей, наивный Байес, XGBoost и LightGBM. Производительность моделей оценивалась с использованием площади под кривой (AUC), точности, чувствительности, специфичности, положительной и отрицательной прогностической ценности. Оптимальная модель была интерпретирована с помощью SHapley Additive exPlanations (SHAP).

Результаты

Регрессия LASSO отобрала семь предикторов: история диабета, история панкреатита, история билиарной хирургии, С-реактивный белок, альбумин, мочевины крови и сывороточный кальций. Модель случайного леса продемонстрировала наилучшую классификационную эффективность, достигнув AUC 0,884 (95% ДИ: 0,827–0,941) на внутренней тестовой выборке и 0,914 (95% ДИ: 0,845–0,983) на выборке для временной валидации. Анализ SHAP выявил сывороточный кальций и С-реактивный белок в качестве наиболее важных предикторов, при этом низкий уровень кальция и повышенный СРБ существенно увеличивали вероятность классификации ППК.

Обсуждение

Мы разработали и временно валидировали интерпретируемые модели машинного обучения для классификации статуса развития ППК с использованием семи рутинно доступных клинических показателей. Модель случайного леса продемонстрировала отличную дискриминацию и обобщаемость, в то время как анализ SHAP обеспечил прозрачные объяснения индивидуальных классификаций. Эти модели могут способствовать раннему выявлению пациентов с ОП высокого риска и направлять проактивное клиническое ведение.

Перевод выполнен: 21.03.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.