

Раннее прогнозирование риска диабета 2 типа с помощью объяснимого машинного обучения в двухэтапном подходе

Источник: Frontiers in Digital Health

Оригинал: <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1743619>

диагностика

машинное обучение

мобильное здравоохранение

объяснимый ИИ

скрининг

эндокринология

Фон

Сахарный диабет — хроническое заболевание, характеризующееся повышенным уровнем глюкозы в крови. Без раннего выявления и надлежащего управления оно может привести к серьёзным осложнениям и увеличить расходы на здравоохранение. Его глобальная распространённость растёт, при этом многие случаи остаются невыявленными. В данном исследовании мы разработали объяснимую модель машинного обучения, используя двухэтапный подход для прогнозирования диабета.

Методы

Пять моделей машинного обучения (ML), включая многослойный перцептрон (Multi-Layer Perceptron), метод опорных векторов (Support Vector Machine), метод k-ближайших соседей (K-Nearest Neighbor), экстремальный градиентный бустинг (Extreme Gradient Boosting, XGBoost) и наивный байесовский классификатор (Naïve Bayes), были обучены и оценены с использованием двухэтапного подхода. На первом этапе использовался открытый набор данных, содержащий 520 образцов, и для отбора признаков применялись объяснения Шепли (Shapley Additive exPlanations, SHAP) и веса

MLP. На втором этапе те же модели были обучены и оценены с использованием набора данных из 270 943 образцов, собранных в Руанде. SHAP дополнительно использовался для объяснения вывода модели.

Результаты

На первом этапе модель многослойного перцептрона показала наилучшие результаты на открытом наборе данных с точностью 95,19%. Методы отбора признаков выявили 10 наиболее влиятельных предикторов, связанных с риском диабета, включая те, которые рекомендуются поставщиками услуг по уходу за диабетом в Руанде. На втором этапе модель XGB превзошла другие модели, достигнув точности 97,14%.

Заключение

В данном исследовании представлен двухэтажный объяснимый фреймворк машинного обучения для систематического скрининга сахарного диабета 2 типа. Первый этап оценивает риск на основе сообщаемых симптомов, тогда как второй этап включает демографические, антропометрические данные и показатели жизненно важных функций для уточнённой оценки риска. Интеграция этих моделей в мобильное приложение mUzima может повысить возможности работников здравоохранения на уровне общин по выявлению и направлению лиц с высоким риском. Благодаря возможности раннего и точного выявления предложенный подход имеет потенциал снизить количество невыявленного диабета и поддержать улучшение управления заболеванием.

Перевод выполнен: 28.03.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.