

Многомодельное прогнозирование на основе машинного обучения плохого самочувствия, генерализованного тревожного расстройства и бессонницы среди взрослых в неформальных поселениях Индии: к объяснимому цифровому скринингу психического здоровья

Источник: Frontiers in Digital Health

Оригинал: <https://www.frontiersin.org/articles/10.3389/fdgth.2026.1935205>

диагностика

машинное обучение

объяснимый ИИ

психическое здоровье

цифровое здравоохранение

эпидемиология

Введение

Уровень заболеваемости психическими расстройствами среди населения, проживающего в неформальных городских поселениях, формируется под влиянием сложного взаимодействия поведенческих факторов риска, таких как время, проводимое перед экранами, продолжительность рабочего дня и продолжительность сна, а также социально-демографической уязвимости.

Методы

Мы проанализировали данные перекрестного исследования, полученные от 1600 взрослых участников, набранных из 16 кластеров в зоне полевой практики Центра подготовки по городскому здравоохранению медицинского колледжа в Пуне, Индия. Целью исследования было разработка и сравнение десяти алгоритмов контролируемого машинного обучения для прогнозирования низкого уровня благополучия (по шкале WHO-5), умеренной или тяжелой генерализованной тревожной расстройства (по шкале GAD-7) и

умеренной или тяжелой бессонницы (по шкале ISI). Внутренняя согласованность всех трех инструментов была подтверждена на основе ответов на уровне отдельных пунктов. Поскольку участники были сгруппированы в 16 кластеров, основной анализ использовал кросс-валидацию с оставлением одного кластера (LOCO-CV), при которой каждый кластер поочередно выделялся в качестве независимого тестового набора; в качестве вторичного анализа сохранялось обычное разделение на уровне участников 80/20 с 10-кратной кросс-валидацией. Дискриминационная способность, калибровка и клиническая полезность моделей оценивались с использованием площади под кривой (AUC), чувствительности, специфичности, предиктивных значений, показателя Брайера, анализа кривых принятия решений и теста Делонга, причем для каждого показателя рассчитывались 95% доверительные интервалы с использованием кластерной бутстрап-оценки. Для интерпретации лучшей по производительности модели использовался метод SHAP, а также была дополнительно оценена согласованность атрибуции признаков в рамках множества Рашимона (Rashomon set), состоящего из нескольких ближайших по качеству топовых моделей.

Результаты

Шкалы WHO-5, GAD-7 и ISI показали приемлемую до отличной внутренней согласованность в данной выборке (альфа Кронбаха 0,730–0,938; омега Макдональда 0,815–0,953). Алгоритм LightGBM достиг наивысшего первичного (LOCO-CV) значения AUC для всех трех исходов: 0,977 (95% ДИ 0,970–0,983) для низкого уровня благополучия, 0,987 (95% ДИ 0,984–0,992) для тревожного расстройства и 0,988 (95% ДИ 0,982–0,993) для бессонницы. Несколько социально-демографических предикторов показали необычно сильные однофакторные ассоциации с исходами, что было связано с повторяющимся междукластерным паттерном в выборочной совокупности; такая структура означает, что даже валидация, учитывающая кластерную структуру, может оставаться излишне оптимистичной по сравнению с genuinely внешней популяцией, что мы подробно обсуждаем в разделе ограничений. Рейтинги на основе SHAP были высоко согласованы (Спирман ρ = 0,84–0,95) между тремя лучшими по производительности статистически близкими по качеству моделями. Анализ чувствительности с использованием взвешивания классов улучшил положительную прогностическую ценность для низкого уровня благополучия с 0,478 до 0,600 при минимальных затратах на снижение AUC.

Заключение

Ансамбли градиентного бустинга, в частности LightGBM, продемонстрировали сильную и внутренне согласованную дискриминационную способность для этих трех исходов, оцененных с помощью надежных инструментов; однако, учитывая выявленную в данной выборке внутрикластерную однородность, внешняя валидация на независимой популяции необходима перед любыми утверждениями о обобщаемой клинической полезности.

Перевод выполнен: 03.10.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.