Автоматизированная оценка навыков деэскалации конфликтов в здравоохранении с использованием виртуальной реальности на базе разговорного ИИ

Источник: Frontiers in Digital Health

Оригинал: https://www.frontiersin.org/articles/10.3389/fdgth.2026.1938760

NLP | автоматическая оценка | виртуальная реальность | глубокое обучение | коммуникация | медицинское образование

Введение

Среды обучения цифровому здравоохранению все чаще требуют масштабируемых методов оценки коммуникативных навыков в реалистичных интерактивных сценариях. В данной статье представлено сравнительное исследование подходов к автоматизированной оценке в рамках симуляции виртуальной реальности на базе диалогового искусственного интеллекта, предназначенной для обучения навыкам вербальной деэскалации конфликтов.

Методы

Всего 20 студентов-медсестер прошли обучение в рамках симуляции, где их устные ответы фиксировались и преобразовывались в текст с помощью системы распознавания речи (speech-to-text), после чего оценивались по 11 компетенциям деэскалации. Оценка сравнивала три дообученные модели глубокого обучения (BERT, RoBERTa и XLNet) с модифицированной версией RATAS — структурированной на основе рубрик системой, использующей большие языковые модели (LLM). Было проведено два эксперимента: один с

использованием контролируемых валидационных наборов данных по 100 элементов для каждого раздела, и другой с использованием реальных ответов, собранных в ходе пользовательского исследования.

Результаты

На валидационных наборах данных модели глубокого обучения показали наилучшие результаты: RoBERTa достигла средней точности 0,9927, за ней следовали BERT с точностью 0,9900 и XLNet с точностью 0,9790, тогда как RATAS показала результат 0,8554. На наборе данных с ответами студентов производительность всех подходов снизилась: RoBERTa показала точность 0,8590, RATAS — 0,8181, BERT — 0,8000, а XLNet — 0,7681.

Обсуждение

Эти результаты свидетельствуют о том, что модели на основе трансформеров демонстрируют наилучшие показатели в контролируемых условиях, тогда как RATAS продемонстрировала стабильность при выставлении оценок на различных типах наборов данных. Эти предварительные выводы подтверждают целесообразность интеграции подходов к автоматизированной оценке на основе глубокого обучения и рубрик, использующих большие языковые модели, для обучения коммуникации и деэскалации конфликтов в медицинском образовании.

Перевод выполнен: 02.11.2026 | ai4med.ru

Машинный перевод. Рекомендуем сверять с оригиналом при клиническом использовании.