

ПРОЕКТИРОВАНИЕ БАЗЫ ДАННЫХ ДЛЯ СОЗДАНИЯ СИСТЕМЫ «ЦИФРОВОЙ ДВОЙНИК ПАЦИЕНТА»

Аннотация. В работе представлены этапы проектирования базы данных для проекта «Телереанимация», направленного на информатизацию отделения реанимации и интенсивной терапии. Описаны основные шаги при проектировании реляционной базы данных.

Ключевые слова: реляционная база данных, СУБД, модели баз данных, цифровой двойник, медицинский датасет, персонифицированная медицина.

Введение. Современная медицина стоит на пороге революционных изменений, обусловленных стремительным развитием информационных технологий и искусственного интеллекта [1].

Целью проекта «Телереанимация» является разработка информационной системы «Цифровой двойник», которая позволит:

- предсказывать дальнейшее состояние пациента при помощи алгоритмов предиктивной аналитики;
- структурировать медицинские показатели пациентов, которые можно было бы использовать в анализе и проверке медицинских гипотез.

Информационная система «Цифровой двойник» — это комплексное решение, включающее в себя: медицинскую базу данных — в нее будут поступать данные о пациентах больницы, с особым акцентом на отделение реанимации и интенсивной терапии. Эти данные могут включать в себя результаты анализов, историю болезни, показатели жизнедеятельности, информацию о принимаемых лекарствах, результаты лабораторных исследований и алгоритмы машинного обучения — обученные на основе накопленных данных. Эти алгоритмы будут способны предсказывать развитие сепсиса и септического шока — опасных состояний, требующих быстрой и точной диагностики.

Анализ больших объемов данных в сочетании с возможностями искусственного интеллекта позволит врачам: выявлять скрытые закономерности и взаимосвязи между различными показателями здоровья, что может привести к открытию новых методов диагностики и лечения [2, 3], персонализировать лечение [4], учитывая индивидуальные особенности каждого пациента и предсказывая его реакцию на терапевтические методы.

Проблема исследования. Целью исследования является проектирование базы данных для информационной системы «Цифровой двойник пациента». Авторами поставлены задачи анализа аналогов, построения концептуальной, логической и физической моделей базы данных.

Материалы и методы. Для проектирования реляционной базы данных будут использоваться методы построения и описания концептуальной и логической моделей в виде схемы в нотации IDEF1X.

Перед началом проектирования проанализированы структуры существующих медицинских датасетов. К рассмотрению предложены три датасета — MIMIC III, MIMIC IV и eICU [5-7].

MIMIC III и MIMIC IV разрабатывались в ходе сотрудничества Медицинского центра Бет Исраэль Диаконесса (BIDMC) и Массачусетского технологического института (MIT) [8].

MIMIC III — более старая версия датасета, данные собирались с 53 423 пациентов старше 16 лет, и писались с двух медицинских информационных систем в период с 2001 по 2012 год [9]. Всего набор данных содержит 26 таблиц.

MIMIC IV — новая версия датасета, которая стала более структурированной, — таблицы были разбиты на модули, например, модуль ОРИТ, Госпитализации, Неотложной помощи и др. При этом интеграция между модулями полностью сохраняется, разделение введено для удобства пользования. Данный датасет содержит уже 45 таблиц [10].

eICU — медицинский датасет с данными пациентов отделения реанимации. Он более узконаправленный, чем предыдущие два набора данных, так как содержит только информацию из отделения реанимации и интенсивной терапии [11]. Датасет рассмотрен лишь как альтернативное решение, так как от него требуется хранение внешней информации, которая появляется вне отделения реанимации.

За основу проектируемой базы данных взята идея о записи регистраций параметров пациента и связи ее с госпитализацией пациента в ОРИТ.

Результаты. Концептуальная схема была построена на основе опросов заказчиков проекта, а также других команд, которые будут использовать разрабатываемую базу данных в своей работе. Сущности разделены на семантические модули — одна сущность может присутствовать в нескольких модулях. Сущности объединены связями в общей схеме, но для удобства восприятия сущности распределены на логические части.

Концептуальная модель базы данных содержит следующие модули:

- модуль госпитализации — содержит сущности, относящиеся к госпитализации пациента. Содержит таблицы: пациент, госпитализация, обстоятельство посещения.
- модуль госпитализации в ОРИТ — содержит сущности, относящиеся к госпитализации пациента в ОРИТ. Содержит таблицы: госпитализации, госпитализации в отделение реанимации и интенсивной терапии и регистрации показателей и иной информации о состоянии пациента.
- модуль показателей — содержит сущности, относящиеся к регистрации показателей состояния пациента при его нахождении в ОРИТ. Содержит таблицы: показатель пациента, регистрация показателя пациента, регистрация основных показателей пациента, единица измерения.
- модуль препаратов содержит сущности, относящиеся к введению препаратов в организм пациента при его нахождении в ОРИТ. Содержит таблицы: препарат, группа препарата, назначение препарата, регистрация введения препарата, лекарственное средство.
- модуль дренажей содержит сущности, относящиеся к дренажам пациента. Содержит таблицы: дренажное оборудование, регистрация дренажей пациента.
- модуль лабораторных исследований содержит сущности, относящиеся к лабораторным исследованиям пациента. Содержит таблицы: профиль лабораторного исследования, лабораторный тест, лабораторное исследование.
- модуль диагнозов содержит сущности, относящиеся к диагнозам пациента за данную госпитализацию. Имеет таблицы: диагноз, код МКБ.

Модуль истории болезни содержит сущности, относящиеся к истории болезни пациента, включая назначения и результаты фармакотерапии, и сопутствующие заболевания у пациента. Содержит таблицы: сопутствующее заболевание, регистрация результата осмотра в истории болезни, результат фармакотерапии.

На основе концептуальной схемы построена логическая схема, которая представлена на рис. 1. Логическая модель позволяет увидеть полную структуру базы данных и спроектировать физическую реализацию для любой системы управления базами данных. Благодаря моделированию логического уровня сформировалась четкая структура базы данных.

Заключение. Для проекта «Телереанимация» проведена работа по проектированию базы данных для системы «Цифровой двойник» под нужды заказчика. Для глубокого понимания предметной области и потребностей пользователей проведена исследовательская работа. Анализ персональных медицинских данных позволил изучить структуру медицинской информации, особенности ее хранения и обработки. Кроме того, проведено интервью со студентами медицинского университета и заказчиками проекта, чтобы учесть их мнение и пожелания при проектировании базы данных.

Результатом работы стали две модели базы данных: концептуальная модель — на этом этапе были определены основные сущности, их атрибуты и взаимосвязи между ними, логическая модель, в которой определены первичные и внешние ключи и ограничения целостности данных. Разработанная база данных должна стать надежным фундаментом для системы «Цифровой двойник пациента» и позволить реализовать все ее возможности.

СПИСОК ЛИТЕРАТУРЫ

1. Современный передовой уровень искусственного интеллекта для умной медицины / О.Ю. Колесниченко, А.В. Мартынов, В.В. Пулит [и др.]. — Текст: электронный // Ремедиум. — 2019. — № 4. — URL: <https://cyberleninka.ru/article/n/sovremennyu-peredovoy-uroven-iskusstvennogo-intellekta-dlya-umnoy-meditsiny> (дата обращения: 18.04.2024).
2. A Machine-Learning Approach for Dynamic Prediction of Sepsis-Induced Coagulopathy in Critically Ill Patients With Sepsis *Frontiers in medicine* / Qin-Yu, Zhao, Liu, Le-Ping, Luo, Jing-Chao [et al.]. — doi:10.3389/fmed.2020.637434. — Text: electronic // PubMed Central. — 21 Jan. 2021. — Vol. 7. 637434. — URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7859637/> (date of the application: 18.04.2024).
3. Бодрин К.А., Красноперова А.А. Использование технологий машинного обучения в медицине / К.А. Бодрин, А.А. Красноперова. — Текст: электронный // Теория и практика современной науки. — 2018. — № 10 (40). — URL: <https://cyberleninka.ru/article/n/ispolzovanie-tehnologiy-mashinnogo-obucheniya-v-meditsine> (дата обращения: 20.04.2024).
4. Наймушина А.Н., Чекардовская И.А., Бакановская Л.Н. Системный анализ и обработка экспертной информации для определения степени значимости факторов риска / А.Н. Наймушина, И.А. Чекардовская, Л.Н. Бакановская. — Текст: электронный // Информационные технологии и системы. — Ханты-Мансийск, 12-16 марта 2019 г. — С. 20-25. URL: <https://www.elibrary.ru/item.asp?id=38541051> (дата обращения: 19.04.2024).
5. MIMIC-III documentation. — Text: electronic // Medical Information Mart for Intensive Care: [site]. — URL: <https://mimic.mit.edu/docs/iii/> (date of the application: 10.03.2024).
6. MIMIC-IV documentation. — Text: electronic // Medical Information Mart for Intensive Care: [site]. — URL: <https://mimic.mit.edu/docs/iv/> (date of the application: 10.03.2024).
7. eICU Collaborative Research Database. — Text: electronic // Medical Information Mart for Intensive Care: [site]. — URL: <https://eicu-crd.mit.edu/about/eicu> (date of the application: 10.03.2024).

8. MIMIC-IV, a freely accessible electronic health record dataset / A.E.W., Johnson, Lucas, Bulgarelli, Lu, Shen [et al.]. — Text: electronic // Sci Data. — 2023. — 10, 1. — URL: <https://doi.org/10.1038/s41597-022-01899-x> (date of the application: 27.03.2024).
9. MIMIC-III, a freely accessible critical care database / A.E.W., Johnson, T., Polland, Lu, Shen [et al.]. — Text: electronic // Sci Data. — 2016. — 3, 160035. — URL: <https://doi.org/10.1038/sdata.2016.35> (date of the application: 27.03.2024).
10. What's new in MIMIC-IV? — Text: electronic // Medical Information Mart for Intensive Care: [site]. — URL: <https://mimic.mit.edu/docs/iv/about/whatsnew> (date of the application: 23.04.2024).
11. Чернышева Т.Ю. Обзор аналогов и первоисточников медицинских данных для разработки структуры датасета «цифровой двойник» / Т.Ю. Чернышева, Л.Н. Бакановская, В.Д. Кравец, А.И. Насибуллин. — Текст: электронный // Современные тенденции и инновации в науке и производстве: XIII международ. науч.-практ. конф. 24 апр. 2024 г. — Междуреченск, 2024. (принята к печати)