

АННОТАЦИЯ

диссертационной работы представленной на соискание степени
доктора философии (PhD) по образовательной программе
8D06101 – «Информационные системы (по отраслям)»
Исмухамедовой Айгерим Мэлсатовны

АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ КЛИНИЧЕСКИХ РЕШЕНИЙ

Актуальность темы исследования. С развитием информационных технологий медицинская индустрия переживает значительные трансформации, связанные с цифровизацией больших массивов данных. Эти изменения открывают новые возможности для улучшения качества медицинского обслуживания, однако также порождают вызовы, связанные с необходимостью эффективного анализа и интерпретации сложных и разнообразных данных. В частности, быстрое увеличение объёма доступной медицинской информации требует разработки новых подходов и методов обработки данных, которые могли бы справляться с этими задачами на качественно новом уровне.

Данное исследование вписывается в широкомасштабный проект по цифровизации медицинской отрасли в Республике Казахстан (Kazakhstan 2050, eHealth, Digital Kazakhstan, Электронный паспорт здоровья, реализация концепции Smart City) и касается реализации системы поддержки принятия клинических решений на базе искусственного интеллекта в рамках национальной системы здравоохранения. Такое внедрение инновационных технологий в медицине не только улучшит качество оказания помощи пациентам за счёт более ранней и точной диагностики, но и способствует оптимизации административных процессов, гарантируя надёжное и эффективное управление медицинскими данными.

Цифровая трансформация здравоохранения является процессом внедрения современных цифровых технологий и инноваций в сферу здравоохранения с целью повышения качества медицинских услуг, улучшения доступа к ним, оптимизации управления медицинскими процессами и повышения эффективности работы учреждений здравоохранения. Этот процесс включает в себя использование новых технологий для улучшения диагностики, лечения, мониторинга, а также управления данными и пациентами. Одним из ключевых аспектов цифровой трансформации здравоохранения является использование алгоритмов машинного обучения в анализе медицинских данных, например, для интерпретации изображений, диагностики заболеваний, прогнозирования заболеваний и оптимизации лечения. Алгоритмы машинного обучения могут помогать в обработке больших объёмов данных, что ускоряет процесс принятия решений.

Применение искусственного интеллекта и машинного обучения в обработке данных пациентов с хронической ишемией и в других медицинских областях значительно улучшило результаты в диагностике, терапии и

прогнозировании заболеваний. Джулиан Мутц и Кэтрин Льюис использовали методы машинного обучения для определения биологического возраста на основе данных о психических характеристиках, связанных с ускоренным старением. Искусственный интеллект и машинное обучение играют ключевую роль в разработке новаторских подходов к лечению болезни Паркинсона и ранней диагностике онкологических заболеваний с высокой смертностью.

Теория и практика внедрения информационных технологий в медицинскую сферу, относительно диагностирования и корректировки лечения различных заболеваний рассмотрены в трудах казахстанских ученых М.Т.Ипалакова, М.Е.Мансурова, А.К.Мукашева, Н.П.Сапарходжаев, А.К.Мукашева и других.

Исследования показывают, что большинство моделей машинного обучения достигают точности более 95% в диагностике хронического лимфоцитарного лейкоза, а также обеспечивают 100% точность в дифференциации этого заболевания от других патологий. Методы машинного и глубокого обучения активно применяются для устранения диагностических пробелов в случаях наследственной аритмии. Также важную роль эти технологии играют в выборе оптимальных алгоритмов для лечения рака молочной железы.

Для улучшения результатов исследователи часто используют комбинации из 5-6 моделей машинного обучения, включая базовые методы классификации в рамках глубокого обучения. Индийские специалисты применяют ряд классификаторов машинного обучения, включая NB, KNN, SVM и алгоритмы деревьев решений, такие как ID3 и C4.5, для прогнозирования и диагностики диабета. Коллеги из Пакистана демонстрируют применение шести известных алгоритмов машинного обучения, таких как SVM, KNN, логистическая регрессия, деревья решений, случайный лес и наивный Байес для прогнозирования диабета, достигая точности до 77%.

В рамках данного исследования для исследования процессов принятия клинических решений было выбрана сфера диабетологии. Сахарный диабет, широко известный как диабет, является серьезной глобальной проблемой в области здравоохранения, затрагивающей миллионы людей по всему миру. Это заболевание достигло уровня эпидемии во многих регионах мира, при этом его распространённость продолжает расти. Согласно международным медицинским данным Всемирной организации здравоохранения, примерно 422 миллиона человек на глобальном уровне страдают от этого заболевания, что составляет приблизительно 6,028% от общей численности населения. В контексте Республики Казахстан статистика ВОЗ показывает, что 11,5% населения страдают от сахарного диабета; среди них 11,7% составляют женщины и 11,3% - мужчины. Национальный регистр Республики Казахстан за 2021 год фиксирует 317 597 зарегистрированных случаев сахарного диабета, включая 314 407 взрослых, 2 379 детей до 14 лет и около 811 подростков в возрастной категории от 15 до 17 лет.

Шведские ученые исследуют применение машинного обучения в

профилактических программах по борьбе с диабетом 2 типа, выявляя ключевые рисковые факторы развития этого заболевания.

Таким образом, актуальность темы заключается в том, что машинное обучение и искусственный интеллект становятся ключевыми технологиями в разработке методик и подходов в системах поддержки принятия медицинских решений. Государственные программы служат платформой для внедрения и тестирования этих передовых технологий, и чем обширнее база данных о пациентах, тем выше точность анализа и быстрее реализация новых уникальных программных продуктов и решений в области диагностики и прогнозирования заболеваний.

Актуальность диссертационной работы также подтверждается тем, что исследование выполнено в рамках НИР по договору №321/23-25 от 03.08.2023 по теме АР19679525 «Программный комплекс диагностики клиничко-гематологических синдромов для электронного паспорта здоровья», выполняемому в рамках бюджетной программы «Грантовое финансирование научных исследований», а также НИР по договору №128/ЖГ 5-24-26 от 02.06.2024 по теме АР22683316 «Применение алгоритмов машинного обучения для систем поддержки принятия врачебных решений» выполняемому в рамках бюджетной программы «Грантовое финансирование молодых ученых по проекту «Жас ғалым» на 2023-2025 годы».

Объектом исследования является система поддержки принятия клинических решений.

Предметом исследования алгоритмическое обеспечение интеллектуальной системы поддержки принятия клинических решений в эндокринологии и диабетологии.

Идея работы - применение технологий машинного обучения и искусственного интеллекта для задач интеллектуальной поддержки процессов принятия клинических решений в эндокринологии и диабетологии, позволяющих внести вклад в реализацию глобальную стратегию Всемирной организации здравоохранения по электронному здравоохранению на 2020-2025 годы.

Цель исследования заключается в разработке алгоритмов интеллектуальной поддержки принятия клинических решений на основе алгоритмов машинного обучения.

Для достижения поставленной цели необходимо выполнить следующие **исследования** и решить **основные задачи**:

- исследование процессов диагностики в информационно-клинической области и особенностей систем поддержки принятия клинических решений;
- изучение вопросов эффективности использования и проблематики применения технологии машинного обучения для систем поддержки принятия клинических решений;
- разработка концептуальной модели процесса поддержки принятия клинических решений на основе методики EDA;
- разработка гибридного алгоритма поддержки клинических решений на основе андерсемплинга и автоматической оптимизации параметров;

- разработка алгоритма применения метода Grid Search для задач поддержки принятия клинических решений на основе модели CNN;
- разработка алгоритма ансамблирования архитектур нейронных сетей для задач поддержки клинических решений;
- разработка информационной модели интеллектуальной системы поддержки принятия клинических решений;
- реализация методики EDA на клинических данных эндокринологии и диабетологии;
- проведение экспериментального исследования алгоритма поддержки клинических решений эндокринологии на основе технологии андерсэмплинга;
- оценка эффективности алгоритма прогнозирования диабета на основе модели глубокой нейронные сети с оптимизированными гиперпараметрами;
- оценка точности реализации алгоритма ансамблирования архитектур нейронных сетей LSTM и RNN для задач поддержки клинических решений;
- разработка архитектуры системы поддержки принятия клинических решений.

Научная новизна диссертационного исследования заключается в том, что впервые для повышения эффективности процессов поддержки принятия клинических решений в эндокринологии и диабетологии предложен комплекс алгоритмов, интегрирующий технологию андерсэмплинга и ансамблирования архитектур нейронных сетей LSTM и RNN.

Основные научные положения, выносимые на защиту:

- *гибридный алгоритм* поддержки клинических решений на основе андерсэмплинга и автоматической оптимизация параметров;
- *алгоритм* применения метода Grid Search для задач поддержки принятия клинических решений на основе модели CNN;
- *алгоритм* ансамблирования архитектур нейронных сетей для задач поддержки клинических решений.

Практическая значимость результатов диссертационного исследования заключается в применимости предложенного комплекса алгоритмов, интегрирующий технологию андерсэмплинга и ансамблирования архитектур нейронных сетей LSTM и RNN для решения задач информационно-аналитической поддержки принятия управленческих решений при сопровождении технологических бизнес-процессов в эндокринологии и диабетологии.

Научно-обоснованные теоретические и экспериментальные результаты диссертационной работы использованы в научном проекте по теме «Программный комплекс диагностики клиничко-гематологических синдромов для электронного паспорта здоровья», выполняемого в рамках бюджетной программы «Грантовое финансирование научных исследований», о чем свидетельствует справка об участие в проекте.

Стоит отметить, что экспериментальные результаты работы использованы в научном проекте по теме «Применение алгоритмов машинного обучения для систем поддержки принятия врачебных решений», выполняемого в рамках конкурса на грантовое финансирование исследований

молодых ученых по проекту «Жас ғалым» на 2024-2026 годы, о чем свидетельствует справка об участие в проекте.

Получены свидетельства о государственной регистрации прав на объекты авторского права №49449 от 04.09.2024 «Алгоритм поддержки клинических решений на основе технологии андерсэмплинга», так же №4737 от 01.08.2019 «Программный модуль диагностирования клинико-гематологических синдромов» и №41784 от 05.01.2024 «База данных дифференциального диагностирования клинико-гематологических синдромов на основе алгоритма морфологической классификации».

Разработанная в диссертационной работе модули по методам управления большими данными, включая их сбор, хранение и обработку, на основе работы с медицинскими данными в контексте СППКР была успешно применена на базе производства ТОО «ЮвентаМед», о чем свидетельствует акт внедрения.

Разработанная в диссертационной работе методы управление IT проектами и этические аспекты использования данных были внедрены в учебный процесс ОП 7M04104 «IT Менеджмент», КАСУ на 2024-2025 года в виде дисциплинах «Design and Implementation of Software System» и «Digital business modeling», о чем свидетельствует акт внедрения.

Результаты диссертационного исследования были внедрены в учебный процесс в 2023-2024 учебном году в следующих курсах лекционных и практических занятий ОП «Математическое и компьютерное моделирование», ВКТУ им. Д.Серикбаева, а именно в дисциплины «Моделирование биологических процессов» и «Основы нейронных сетей», о чем свидетельствует акт внедрения.

Методы исследования. В работе используется методы машинного и глубокого обучения, методы статистического анализа, методы обработки больших данных, теория принятия решений, а также методы экспериментального исследования и моделирования архитектуры системы.

Апробация результатов диссертационного исследования. Основные результаты диссертационной работы докладывались на научных семинарах кафедры «Информационные системы» ВКГТУ им. Д. Серикбаева и на следующих международных научно-практических конференциях: «ADVANCED SCIENCE» (Пенза, Россия, 2017 г.); «4th International Conference on Computer and Technology Applications» (Стамбул, Турция, 2018 г.); «4th International Conference on Engineering and MIS» (Стамбул, Турция, 2018 г.); «IV Международная научно-техническая конференция студентов, магистрантов и молодых ученых» (Усть-Каменогорск, Казахстан, 2018 г.); «XVIII международный научно-исследовательский конкурс «Лучшая научная статья 2018»» (Пенза, Россия, 2018 г.); «Computational and Information Technologies in Science, Engineering and Education: 9th International Conference, CITech 2018» (Усть-Каменогорск, Казахстан, 2018 г.); «Application of Information and Communication Technologies-AICT 2018» (Алматы, Казахстан, 2018 г.); «5th International Conference on Engineering and MIS» (Нур-Султан, Казахстан, 2019 г.); «12th IEEE International Conference «Application of

Information and Communication Technologies - AICT2019»»» (Баку, Азербайджан, 2019 г.); «VI Международная научно-техническая конференция студентов, магистрантов и молодых ученых «Творчество молодых - инновационному развитию Казахстана»» (Усть-Каменогорск, Казахстан, 2020 г.); «2021 International Young Engineers Forum (YEF-ECE)» (2021 г.); «7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)» (Анкара, Турция, 2023 г.).

Личный вклад автора. Постановка проблемы, формализация всех рассмотренных задач, поиск методов и алгоритмов их решения, а также приведенные в диссертации научные и практические результаты, их анализ, формирование итоговых выводов осуществлены лично автором диссертации.

Публикации по теме диссертационного исследования. По теме диссертации опубликовано 28 научных работы, из них 8 в научных журналах, рекомендованных Комитетом по контролю в сфере науки и высшего образования МОН РК; 12 в трудах международных конференций; 3 публикации индексируются в базе данных SCOPUS; 1 монография.

Публикации, опубликованные в рамках научного исследования в научных изданиях Scopus и Web of Science были процитированы 49 раз.

Результаты экспериментального исследования алгоритмов интеллектуальной поддержки систем принятия клинических решений были описаны в статье на тему «Integrating machine learning in electronic health passport based on WHO study and healthcare resources», опубликованной в журнале «Informatics in Medicine Unlocked», имеющий в Scopus показатель процентиля по CiteScore равный 86 по направлению «Computer Science».

Структура и объем диссертационной работы. Диссертация состоит из введения, трех глав, заключения, библиографии и приложений.

В первом разделе исследованы существующие подходы к диагностике, выявлены ключевые проблемы и преимущества применения систем поддержки принятия клинических решений.

Второй раздел сосредоточился на алгоритмах и моделях, поддерживающих принятие клинических решений. Разработанные концептуальная модель и алгоритмы, такие как метод андерсэмплинга и техника Grid Search для нейронных сетей, продемонстрировали свою практическую значимость и возможность интеграции в существующие клинические системы. Ансамблирование различных архитектур нейронных сетей показало свою способность повышать точность диагностических прогнозов, что имеет ключевое значение для улучшения качества медицинского обслуживания. Экспериментальное исследование, проведенное в третьем разделе, подтвердило эффективность предложенных алгоритмов на клинических данных. Оценка точности моделей LSTM и RNN показала, что эти технологии способны значительно улучшить результаты диагностики и поддержки клинических решений.

В заключении диссертационной работы показаны основные результаты, сделаны выводы по диссертационным исследованиям, таким образом подтверждена научная новизна и практическая значимость исследования.