

УТВЕРЖДАЮ

Проректор по инновационной деятельности
федерального государственного бюджетного
образовательного учреждения высшего
образования «Уфимский университет
науки и технологий»,
кандидат технических наук, доцент

« 28 »

Г.К. Агеев

2024 г.



ОТЗЫВ

ведущей организации на диссертационную работу

Гришиной Любови Сергеевны на тему:

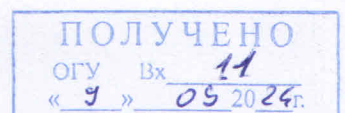
«Методы и алгоритмы интеллектуальной поддержки принятия решений в
медицинской практике на основе обработки естественных языков»,
представленной на соискание ученой степени кандидата технических наук по
специальности 2.3.1. Системный анализ, управление и обработка
информации, статистика

1. Актуальность темы диссертации

Проблемам оказания своевременных, эффективных и качественных медицинских услуг в нашей стране и во всем мире уделяется огромное внимание, а исследования, направленные на поддержку принятия врачебных решений для оптимизации процессов медицинских учреждений, персонализации лечения и сокращения временных затрат, играют ключевую роль в данном вопросе. С каждым годом количество доступной медицинской информации растет экспоненциально. Электронные медицинские карты и данные результатов обследований создают огромные массивы информации, которые требуют эффективных методов анализа и интерпретации.

Использование технологий для мониторинга состояния здоровья пациентов в реальном времени позволяет выявлять заболевания на ранних стадиях и заниматься их профилактикой. Исследования в этой области объединяют знания из медицины, информатики и инженерии, что приводит к инновационным решениям и улучшению качества медицинской помощи.

Важно отметить, что автоматизация рутинных задач позволяет медицинским работникам сосредоточиться на более важных аспектах своей



деятельности, что в свою очередь может повысить общую эффективность системы здравоохранения.

Прогресс в области искусственного интеллекта и машинного обучения открывает новые возможности для разработки и внедрения интеллектуальных систем в медицинскую практику. В этой связи, внедрение систем поддержки принятия врачебных решений на основе технологий искусственного интеллекта – это важный этап цифровой трансформации здравоохранения. Интеллектуальные системы способны выявлять закономерности и аномалии в данных, что может привести к более раннему и точному диагностированию заболеваний, снижая риски для пациентов. При этом, методы обработки естественных языков являются важным инструментом для анализа неструктурированных данных, а также для создания более интуитивных и эффективных интерфейсов взаимодействия между врачами и информационными системами.

Таким образом, диссертация Гришиной Любови Сергеевны на тему «Методы и алгоритмы интеллектуальной поддержки принятия решений в медицинской практике на основе обработки естественных языков» направлена на решение значимой для оптимизации процессов в здравоохранении задачи, что определяет актуальность настоящей работы.

2. Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации, их достоверность и новизна

Обоснованность исследований достигнута за счёт правильности выбора исходных предпосылок и анализа состояния изучаемого вопроса, а также принятия рациональной методологии проведения аналитических исследований и вычислительных экспериментов, основанных на использовании признанных методов.

Достоверность научных положений, выводов и рекомендаций, приведенных в диссертации, обеспечивается использованием апробированных методов исследования, обоснованными достаточным объемом исходных данных, адекватным планированием эксперимента, сходимостью результатов оценки качества построенных интеллектуальных моделей с данными известных работ других исследователей.

Теоретические исследования выполнены на основе анализа научной и нормативно-технической литературы, содержания правовой базы, исходных данных из множества реальных деперсонализированных данных электронных медицинских карт, полученных из базы данных медицинских организаций Оренбургской области. Научные положения теоретической части исследования, выводы и рекомендации, базируются на теории системного анализа и принятия решений, алгоритмах обработки информации, методах

машинного обучения, нейросетевых технологиях, методах обработки естественного языка и методах глубокого обучения.

Научная новизна исследований состоит:

– в разработке концептуальной модели анализа клинических данных и поддержки принятия решений для автоматизации процессов заполнения ЭМК. Отличительной особенностью предложенной модели является формализация этапов структурирования текстовых данных и построение интеллектуальных моделей формирования рекомендаций к лечению диагностированных заболеваний;

– в построении иерархической модели данных амбулаторных карт пациентов. Основным отличием предложенной модели данных от реализованных в существующих МИС является возможность обработки разно-шаблонных XML-документов на основе рекурсивного подхода для обеспечения семантической интероперабельности;

– в разработке метода и алгоритма прогнозирования группы заболеваний пациентов. Предложенный метод использует уникальный узкоспециализированный корпус текстов, построенный на основе слабоструктурированных текстовых данных ЭМК;

– в разработке метода и алгоритма автоматической генерации индивидуальных листов назначений и рекомендаций к лечению. Предложенный метод позволяет автоматизировать процесс заполнения документов и отличается применением современных предобученных нейросетевых моделей трансформеров;

– в построении прототипа автоматизированного программного комплекса интеллектуальной поддержки принятия решений в медицинской практике. Отличительной особенностью программного комплекса является применение модулей искусственного интеллекта для диагностирования заболеваний и формирования рекомендаций к лечению на основе методов обработки естественных языков.

3. Значимость результатов, полученных автором, для науки и практики

Значимость полученных научных результатов состоит в том, что они вносят весомый вклад в развитие методологической базы исследований и разработок в области совершенствования методов извлечения и структурирования знаний из текстовой информации электронных-медицинских карт пациентов для поддержки принятия решений при диагностике и лечении заболеваний.

Практическая ценность результатов исследования, проведенного автором, состоит в разработке программного комплекса, позволяющего производить автоматизированный анализ состояния пациента и генерацию

индивидуального листа назначений и рекомендаций к лечению на основе методов глубокого обучения. Автоматизированный программный комплекс интеллектуальной поддержки принятия решений в медицинской практике позволяет автоматизировать процесс заполнения документов и допускает коррекцию в соответствии с экспертным мнением врача.

Личный вклад автора, судя по диссертации, заключается в разработке концептуальной модели анализа клинических данных и поддержки принятия врачебных решений; разработке иерархической модели данных амбулаторных карт пациентов; в планировании и проведении аналитических исследований и вычислительных экспериментов, для построения обучающего набора данных, необходимого для практической реализации разработанной системы поддержки принятия решений; разработке метода и алгоритма прогнозирования укрупненных групп заболеваний; разработке метода и алгоритма автоматической генерации индивидуальных листов назначений и рекомендаций к лечению и построении на основе полученных результатов прототипа автоматизированного программного комплекса интеллектуальной поддержки принятия решений в медицинской практике.

4. Публикации основных результатов исследований, реализация результатов работы

По теме диссертационного исследования опубликовано 8 научных работ, 3 из них в изданиях, рекомендованных ВАК РФ и отечественных изданиях, которые входят в международные базы данных и системы цитирования, а также 2 работы в изданиях, индексируемых Scopus и Web of Science, получено свидетельство о государственной регистрации одной программы для ЭВМ (Роспатент).

Основные результаты диссертационной работы обсуждены на научных конференциях: всероссийской научно-методической конференции «Университетский комплекс как региональный центр образования, науки и культуры» (Оренбург, 2023); международной научно-технической конференции «Перспективные информационные технологии» (Самара, 2022); международном семинаре «Вычислительные технологии и прикладная математика» (Владивосток, 2022); 2-ой международной научной конференции «Современные тенденции управления и цифровая экономика: от регионального развития до глобального экономического роста» (MTDE, Екатеринбург, 2020).

Результаты работы используются в практике медицинских исследований ГАУЗ «Оренбургской областной клинической больницы имени В.И. Войнова» и ГАУЗ «Бузулукской больницы скорой медицинской помощи им. академика Н.А. Семашко», а также в учебном процессе ФГБОУ ВО «Оренбургского государственного медицинского университета».

5. Рекомендации по использованию результатов и выводов диссертации

Предложенную в диссертации концептуальную модель анализа клинических данных и поддержки принятия решений для автоматизации процессов заполнения ЭМК, разработанную ее основе иерархическую модель структурирования данных амбулаторных карт пациентов и обработки разношаблонных XML-документов, метод и алгоритм прогнозирования группы заболеваний, а также метод и алгоритм генерации индивидуальных листов назначений и рекомендаций к лечению, рекомендуется использовать:

- *в научных организациях* – для проведения клинических исследований, выявления закономерностей, прогнозирования вспышек заболеваний или изменения в эпидемиологической ситуации; для оценки эффективности различных подходов к лечению на основе исследования индивидуальных листов назначений и рекомендаций к лечению;

- *в образовательных учреждениях* – при изучении дисциплин: «Медицинская информатика и статистика», «Системы искусственного интеллекта» и «Системный анализ и принятие решений» направлений подготовки студентов 31.05.01 Лечебное дело и 01.03.02 Прикладная математика и информатика в целях ознакомления с современными методами поддержки принятия решений;

- *в медицинских учреждениях* – для прогнозирования и выявления потенциальных рисков и заболеваний на ранних стадиях; автоматизированной генерации рекомендаций к лечению для снижения вероятности опечаток в назначениях и повышении согласованности между врачами; снижения временных затрат на оформление документации при заполнении электронных медицинских карт; обучения и поддержки медицинского персонала для принятия более обоснованных решений на основе анализа данных;

- *в органах государственной власти и муниципального самоуправления* – для статистического анализа и мониторинга здоровья населения; при разработке стратегических и краткосрочных программ развития интеллектуальных систем поддержки принятия решений в медицинской практике.

6. Содержание диссертации и ее завершенность

Диссертация состоит из введения, пяти разделов, общих выводов, рекомендаций, списка использованных источников (107 наименований) и приложений, изложенных на 114 страницах машинописного текста, включая 38 рисунков и 17 таблиц.

Область исследования соответствует паспорту специальности 2.3.1. Системный анализ, управление и обработка информации, статистика:

- пункт 2 «Формализация и постановка задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта»;

- пункт 5 «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта»;

- пункт 12 «Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации».

Во введении сформулирована решаемая проблема, обоснована актуальность темы исследования и определены основные его направления.

В первом разделе выполнен обзор научных работ в области поддержки принятия решений на основе анализа медицинских документов и построения моделей машинного обучения с помощью методов обработки естественных языков. Определено, что эффективность моделей машинного обучения, построенных на структурированных данных, снижается из-за необходимости предварительного отбора признаков, что может исключить потенциально полезную информацию из рассмотрения и резко сокращает выборки данных. Выдвинута гипотеза о том, что использование слабоструктурированной текстовой информации электронных-медицинских карт пациентов методами обработки естественных языков для диагностирования заболеваний и формирования индивидуальных листов назначений и рекомендаций к лечению может повысить эффективность принятия врачебных решений. Сформулированы цель и задачи диссертационного исследования.

Во втором разделе представлена концептуальная модель анализа клинических данных и поддержки принятия решений для автоматизации процессов заполнения ЭМК, которая проводит формализацию этапов структурирования текстовых данных и построения интеллектуальных моделей формирования рекомендаций к лечению диагностированных заболеваний.

Построена иерархическая модель данных амбулаторных карт пациентов, а также разработаны модуль автоматизированной выгрузки данных XML-протоколов и модуль выделения информации из разношаблонных файлов протоколов различных случаев лечения пациентов. Предложен подход к рекурсивному перебору узлов XML, с последовательным анализом наличия содержимого и созданием дерева записи оказанной услуги в медицинской организации.

В третьем разделе диссертационной работы автором описаны подходы к прогнозированию укрупненных групп заболеваний на основе слабоструктурированных текстовых данных жалоб пациентов и данных объективного осмотра из ЭМК, приведены результаты реализации моделей прогнозирования на основе методов машинного обучения.

Разработан метод и алгоритм прогнозирования группы заболеваний пациентов с ССЗ, который продемонстрировал сбалансированную точность 85,20 %±1,07%. Для извлечения признаков из слабоструктурированной текстовой информации использованы методы NLP, а для обучения группы моделей машинного обучения использован уникальный узкоспециализированный размеченный корпус текстов и укрупненные группы заболеваний по МКБ-10.

В четвертом разделе представлены подходы к решению задачи языкового моделирования для автоматической генерации текста индивидуальных листов назначений и рекомендаций к лечению, алгоритмы токенизации и языковые модели на базе архитектуры трансформер.

Разработаны метод и алгоритм генерации индивидуальных листов назначений и рекомендаций к лечению в рамках диагностированных заболеваний на основе модели глубокого обучения GPT 3 Large, предварительно обученной на русскоязычном тексте общего назначения. Значение метрики BLEU на основе униграмм и биграмм в среднем для всего корпуса языковых моделей составило 0,667 и 0.357 соответственно.

В пятом разделе представлены результаты разработки автоматизированного программного комплекса интеллектуальной поддержки принятия врачебных решений при диагностике и лечении заболеваний. Для оценки эффективности разработанного программного комплекса для пациентов с сердечно-сосудистыми заболеваниями проведена апробация результатов исследования в ГАУЗ «Бузулукской больнице скорой медицинской помощи» и ГАУЗ «Оренбургской областной клинической больнице имени В.И. Войнова». Результаты исследований показали, что при работе с документацией затраты рабочего времени снизились в среднем на 6,9%.

В заключении приведены основные научные и практические результаты, полученные в процессе проводимых исследований, и сформулированы соответствующие выводы.

Тема диссертационного исследования соответствует его содержанию. Поставленные задачи последовательно раскрываются в ходе диссертации. Основные научные положения разработаны, исследованы, обоснованы и освещены автором в научных публикациях. Автореферат объективно и полно отражает основные положения диссертации, написан грамотно и оформлен в соответствии с требованиями ВАК РФ. Содержание и стиль автореферата упрощает осмысление сути диссертационной работы. Она отличается целостностью и лаконичностью изложения материала. В целом диссертационная работа является завершенным научным исследованием.

7. Вопросы, замечания и пожелания по работе

По представленной работе имеются следующие вопросы, замечания и пожелания:

1. Для оценки точности моделей машинного обучения при прогнозировании укрупненной группы заболеваний использована метрика сбалансированной точности (*balanced accuracy*), которая позволяет равнозначно оценивать редкие и наиболее распространенные классы заболеваний. Однако, эффективность прогнозных моделей в медицине также необходимо рассматривать с точки зрения клинической значимости ошибок, чтобы оценить ложно-положительные и ложно-отрицательные результаты прогноза отдельных заболеваний (метрики *Precision*, *Recall* и *F1-Score*).

2. При построении интеллектуальной модели генерации индивидуальных листов назначений и рекомендаций к лечению автором сделано допущение, что протоколы МИС и уровень квалификации врачей соответствует стандартам. Но данное условие целесообразно подтвердить на практике, например, по результатам экспертной оценки на соответствие протоколов клиническим рекомендациям Министерства здравоохранения РФ.

3. В работе производится оценка эффективности принятия решений в медицинской практике на основе анализа данных средней длительности и структуры приёма больных врачом-терапевтом с учетом основных трудовых операций, необходимых к проведению на первичном осмотре. При этом не рассматривается вопрос об адекватности результатов прогнозирования диагноза и генерации рекомендаций на повторном приеме или при обращении за коррекцией лечения незаконченного клинического случая.

4. В качестве рекомендации по расширению области применения разработанного метода, на наш взгляд, целесообразно рассмотреть возможность применения предложенного подхода в отношении более широкого класса заболеваний (болезней нервной системы, эндокринной системы, расстройства питания, нарушения обмена веществ и др.), а также расширения обучающего набора данных за счет включения медико-географических особенностей территорий отдельных регионов, не ограничиваясь пациентами Оренбургской области.

8. Заключение

Диссертационная работа Гришиной Любови Сергеевны является законченной научно-квалификационной работой, в которой содержится решение научной задачи, имеющей значение для развития теории и практики разработки программных комплексов интеллектуальной поддержки принятия решений в медицинской практике на основе методов обработки естественных языков. Результаты исследования направлены на совершенствование методов обработки слабоструктурированной текстовой информации электронных-

медицинских карт пациентов с целью повышения эффективности принятия решений в процессе диагностирования заболеваний и формирования индивидуальных листов назначений и рекомендаций к лечению.

Считаем, что диссертационная работа Гришиной Любви Сергеевны «Методы и алгоритмы интеллектуальной поддержки принятия решений в медицинской практике на основе обработки естественных языков» выполнена на актуальную тему, имеет научную новизну и практическую значимость, соответствующую пп. 9-14 Положения о присуждении ученых степеней ВАК РФ, утвержденного постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842 (с изменениями и дополнениями от 25.01.2024), предъявляемым к диссертациям на соискание ученой степени кандидата технических наук, а ее автор Гришина Любовь Сергеевна заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика.

Отзыв был заслушан и одобрен на расширенном заседании кафедры вычислительной математики и кибернетики института информатики, математики и робототехники федерального государственного бюджетного образовательного учреждения высшего образования «Уфимский университет науки и технологий» (протокол № 9 от 27.08.2024 г.).

Заведующий кафедрой вычислительной математики и кибернетики УУНиТ,
доктор технических наук,
доцент

Г. Р. Шахмаметова

27 августа 2024 года



ФИО: Шахмаметова Гюзель Радиковна

Научная специальность докторской диссертации 05.13.01 - Системный анализ, управление и обработка информации (в промышленности)

Организация: Федеральное государственное бюджетное образовательное учреждение высшего образования «Уфимский университет науки и технологий»

Адрес: 450076, Приволжский федеральный округ, Республика Башкортостан, г. Уфа, ул. Заки Валиди, дом 32.

Тел.: + 7 (347) 229-96-16, E-mail: rector@uust.ru.

Дата составления отзыва: 27.08.2024г.