

**Ю.Ю. МУХИН,**

к.э.н., исполнительный директор международного конгресса «Информационные технологии в медицине», Генеральный директор ООО «Консэф», Москва, mukhin@consef.ru

К.Ю. МУХИН,

магистр менеджмента, Финансовый Университет при Правительстве Российской Федерации, Департамент Менеджмента, Преподаватель-ассистент, KYMukhin@fa.ru

РЕИНЖИНИРИНГ ОБЩЕСТВЕННОГО ЗДРАВООХРАНЕНИЯ, ОСНОВАННЫЙ НА ПЕРСОНОЦЕНТРИРОВАННОЙ МОДЕЛИ, ГИБРИДНЫХ ПРОЕКТНЫХ ПОДХОДАХ И МЕТОДАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

УДК 338.24.01 338.242.2

Мухин Ю.Ю., Мухин К.Ю. Реинжиниринг общественного здравоохранения, основанный на персонцентрированной модели, гибридных проектных подходах и методах искусственного интеллекта (ООО «Консэф», Финансовый Университет, Москва)

Аннотация. В работе рассмотрена модель общественного здравоохранения нового типа, возникающая на платформе технологической революции информационных систем. Авторами приведено описание ключевых методов управленческого, технологического и математического взаимодействия с коллективной информационной инфраструктурой системы здравоохранения, которые окажут наибольшее влияние на происходящие изменения и составят основу цифрового общества, цифровой экономики и общественного здравоохранения нового типа. Методы рассмотрены с точки зрения их влияния на формирование новых подходов в процессе трансформации укладов экономики и общества. Представлена модель персонцентрированной системы здравоохранения, обобщены ее ключевые преимущества, изложены подходы к гибкой проектной реализации процесса перехода к новой модели.

Ключевые слова: Гибкие подходы к управлению, интеллектуальный анализ, искусственный интеллект, нейронные сети, проектный менеджмент, инноватика, гибкие методологии в управлении проектами, agile, цифровая экономика, здравоохранение.

UDC: 338.24.01 338.242.2

Mukhin Yu.Yu., Mukhin K.Yu. Reengineering of public health system, based on a person-centered model, hybrid project management approaches and methods of artificial intelligence (Consef, Financial University, Moscow)

Abstract. The paper considers new approaches of public health emerging on the platform of the technological revolution of information systems. The authors describe the key methods of managerial, technological and mathematical interaction with the collective information infrastructure of the healthcare system that will bring qualitative changes in the near future and will form the basis of a digital society, digital economy and public health of a new type. The methods are considered from the point of view of their influence on the formation of new approaches in the process of the economic and social transformation of waves of innovation. The model of the person-centered health care system and the forecast of its impact on the subject area are presented, the advantages of the new model are summarized and the agile project management approaches to its implementation in the current stage of development are presented.

Keywords: Agile, project management, artificial intelligence, healthcare, business-agility, medicine, innovations & innovation management, agile project management approaches, neural networks, digital economy.



1. Введение. Актуальность проблемы, новая парадигма здравоохранения и средства ее реализации

Бурное развитие технологий и новых областей научных знаний обуславливает качественные изменения в общественных институтах, которые еще совсем недавно казались незыблемыми с точки зрения базовых постулатов, системы ценностей, принципов построения научных знаний и их применения для нужд общества. Одним из важнейших практических аспектов использования этих достижений является модернизация системы общественного здравоохранения. Несмотря на географические, исторические и культурные различия, она всегда отличалась высоким уровнем иерархичности и конформизма, приверженностью фундаментальным основам медицинской науки, история которой уходит далеко за пределы нашей эры, формализацией целого свода процедур и процессов, ориентированных на поддержку общественного здоровья и принципы доказательной медицины. Научные эксперименты в существующей модели здравоохранения локализованы в определенных, как правило функциональных, исследовательских центрах вплоть до получения хорошо проверенного результата этих исследований на основе рандомизированных клинических исследований и широкого внедрения нового метода только после получения достоверного положительного результата.

Российская модель общественного здравоохранения доказала свою эффективность в части достижения таких ключевых показателей здоровья общества, как снижение смертности, повышение продолжительности жизни, борьбы с массовыми угрозами здоровью (инфекционные и неинфекционные заболевания), формирование культуры здорового образа жизни и ранней профилактики заболеваний. Вместе с тем, в силу описанных причин, она обладает высокой инертностью

и консерватизмом и нуждается на данный момент в дополнительных инструментах роста эффективности с учетом современных достижений научно-технического прогресса.

Технологическая революция в области медицинской техники и оборудования, а также в сфере биотехнологий, оптимизация структуры общественного здравоохранения позволили добиться значительного улучшения ключевых показателей общественного здоровья в России в первой половине 21 века. На данном этапе потенциал перечисленных инструментов роста продолжает играть свою важную роль. Распространение указанных научных и технических достижений по-прежнему обеспечивает значительный вклад в рост качества и доступности медицинской помощи, но уже не может удовлетворить потребности общества в уровне темпов роста ее эффективности в быстроизменяющемся мире. Возникает необходимость поиска новых драйверов, способных поддерживать уровень имеющихся темпов роста.

Одним из важнейших инструментов достижения указанной цели является изменение структуры взаимодействия системы общественного здравоохранения и гражданина, основанное на использовании новых подходов мотивации человека к контролю собственного здоровья, а также автоматизированной системе мониторинга его состояния. Внедрение новых подходов равноправного информационного взаимодействия человека, врача и медицинской науки, реализованного с помощью современных и перспективных методов обработки медицинской информации, может выступить эффективным средством достижения указанной цели и позволит реализовать главный принцип современной системы здравоохранения: принцип «4P» (от англ. Predictive, Preventive, Personalized, Participative), который в редакции П.П. Кузнецова – М.А. Шифрина [10, 11] звучит следующим образом:

✓ Возможность прогнозирования проблем (Предсказательность).



✓ Профилактика (устранение проблем до их появления).

✓ Персонализация.

✓ Совместная работа врача и пациента (Партнерство).

Плюс «4П» по П.П. Кузнецову:

➤ безопасность персональных данных;

➤ прозрачность для всех участников процесса;

➤ защищенность от взаимных вредных воздействий (принцип «не навреди»);

➤ приверженность (поддержка) окружения пациента реализации стратегии лечения.

Техническая реализация представленного принципа возможна на базе одновременного интегрированного взаимодействия централизованных баз данных и знаний, получаемых из электронных медицинских карт пациентов, индивидуальных приборов контроля состояния здоровья и непосредственного участия гражданина в его мониторинге и контроле через информационную систему управления здоровьем. Интеграция описанной информационной системы в классическую систему здравоохранения позволит обеспечить реализацию персоноцентрированной модели охраны здоровья и качественное повышение показателей здоровья населения.

Важным аспектом обратной связи в этом случае является реализация персонального взаимодействия информационной системы управления здоровьем с каждым конкретным человеком, подразумевающая активное использование информационных систем, основанных на датацентрированной модели и методах искусственного интеллекта (Artificial Intelligence). Программные средства нового поколения, использующие методы интеллектуального анализа уже на текущем этапе развития позволяют автоматизировать значительное количество **формализуемых** медицинских действий (контроль одновременных воздействий, формализуемые консультации, непрерывный мониторинг состояния

и его экстренных отклонений и др.), а также часто обеспечивают допустимый уровень достоверности в анализе и прогнозе в сфере неформализованных знаний, позволяющих использовать их в качестве эффективного вспомогательного инструментария в медицине. Аналогичные подходы и средства применяются для построения обучающихся систем поддержки принятия врачебных решений, исследовательских целей, прогнозов, совершенствования клинических рекомендаций, страхового обеспечения, анализа статистических данных и выработки медицинских и технических средств нового поколения.

2. Отраслевая архитектура информационной среды. Точки роста. Смена отраслевой парадигмы под влиянием технологической революции

Мировая экономика сегодня стоит на пороге четвертой технологической революции, с которой связывают возможности кардинальной модернизации производства и сферы услуг, изменения качества жизни, а также появление таких явлений, как цифровое производство, реализация персоноцентрированной модели взаимодействия с гражданами, коллективное использование больших неструктурированных массивов данных для научных и социальных целей, децентрализация управления и т. д. Технологической основой для перехода к новой экономической парадигме являются интернет вещей (IoT), роботизация манипуляций и формализуемых операций, так называемая «уберизация» (изъятие излишних маргинально-транзакционных звеньев из цепочки поставки услуги) практически всех видов бизнеса, глобализация всех рынков и отраслей, внедрение систем искусственного интеллекта.

В широком смысле «Индустрия 4.0» характеризует текущий тренд развития от автоматизации и обмена данными, включая в себя



киберфизические системы, интернет вещей и облачные вычисления к созданию экосистем, основанных на моделях искусственного интеллекта и прямого кастомизированного взаимодействия с потребителем. Указанные изменения приносят новый уклад организации производства и социальных услуг, образования, науки и жизни человека.

Министр здравоохранения Российской Федерации В.И. Скворцова наглядно проиллюстрировала в своем выступлении 21 марта 2017 г. на заседании Совета по стратегическому развитию и приоритетным проектам под председательством Президента Российской Федерации Владимира Путина изменения, которые цифровые технологии принесут в сферу здравоохранения: «В 2020 году в целом завершится формирование единой государственной информационной системы сферы здравоохранения на основе единой электронной медицинской карты пациента. Это позволит внедрить перекрёстный электронный документооборот и личный кабинет пациента. С любого компьютера каждый человек сможет войти в свой личный кабинет и получить необходимую информацию, выписки из медицинских документов, записаться к врачу, связаться со страховым представителем. В системе будет создан архив цифровых изображений и будут внедрены интеллектуальные программы автоматизированного чтения изображений с резким увеличением точности и своевременности диагностики. Будет внедрена система автоматизированной экспертизы качества медицинской помощи: сопоставление алгоритма ведения конкретного больного со вшитыми в систему эталонными протоколами и критериями качества. К концу 2020 года система будет полностью интегрирована с информационно-аналитическими системами ОМС, обращения лекарств и медицинских изделий, что позволит проводить комплексный анализ ресурсоэффективности всей отрасли. Таким образом, к концу 2020 года мы сможем выйти на

качественно новый уровень управления здравоохранением, а также преемственности и качества оказания медицинской помощи. Планируется, что с 2020 года начнут внедряться автоматизированные программы обработки больших массивов данных BIG DATA, которые позволят уже автоматизированно выбирать оптимальные алгоритмы медицинского сопровождения для каждого человека с учётом телемедицинских консультаций».[12]

Появление новых технологий, их влияние на жизнедеятельность активно стимулирует законодательную и нормотворческую деятельность, призванную определить новые контуры жизни общества и порядок интеграции в него новых, ранее не существовавших общественных сущностей, а также информационной экосистемы, обеспечивающей их существование и развитие. В сфере здравоохранения на текущий момент законодательно оформлены следующие положения, регламентирующие информационное взаимодействие:

- Федеральным законом № 242-ФЗ от 29 июля 2017 г. «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам применения информационных технологий в сфере охраны здоровья»[7] закрепляются основные понятия, описывающие Единую государственную информационную систему в сфере здравоохранения, цели и задачи, операторов и пользователей, порядок, срок принятия и тип нормативных актов, регулирующих функционирование системы. Они призваны актуализировать, дополнить и скорректировать положения «Концепции создания единой государственной информационной системы в сфере здравоохранения» (далее – Концепция), утвержденной приказом Министерства здравоохранения и социального развития Российской Федерации от 28 апреля 2011 г. № 364 [8]. Отметим, что указанный документ был первым



и долгое время единственным нормативным документом, определяющим цели, задачи, архитектуру и структуру, порядок взаимодействия элементов Единой государственной информационной системы в сфере здравоохранения (ЕГИСЗ).

- Федеральным законом № 242-ФЗ от 29 июля 2017 г. внесены следующие изменения в Федеральный закон от 21 ноября 2011 года № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации»: «Статья 91¹. Единая государственная информационная система в сфере здравоохранения 1. В целях обеспечения доступа граждан к услугам в сфере здравоохранения в электронной форме, а также взаимодействия информационных систем в сфере здравоохранения уполномоченным федеральным органом исполнительной власти создается, развивается и эксплуатируется единая государственная информационная система в сфере здравоохранения...» [6]

Таким образом, законодательно закреплено новое целевое назначение ЕГИСЗ. В предыдущей редакции, указанной в Концепции цель формулировалась следующим образом: «Основной целью создания Системы является обеспечение эффективной информационной поддержки процесса управления системой медицинской помощи, а также процесса оказания медицинской помощи» [8].

С точки зрения информационной архитектуры системы государственного информационного обеспечения в сфере здравоохранения также учтены векторы, формирующие современные подходы к ее развитию и использованию: интеграция ЕГИСЗ с другими отраслевыми информационными системами (системами ФОМС, ТФОМС, мониторинга движения лекарственных средств, РМИС (изменения в пп. 1 статьи 91 ФЗ от 21 ноября 2011 года № 323-ФЗ), возможность доступа

граждан к своей медицинской информации через личный кабинет «Мое здоровье» ЕПГУ.

Медицинские данные на текущий момент экспортируются в ЕГИСЗ из МИС различного уровня в формате стандартизированных электронных медицинских документов (СЭМД), являющихся ограниченным подмножеством Архитектуры клинических документов CDA Release 2, которые образуют Интегрированную электронную медицинскую карту (ИЭМК)[9].

На этапе создания ЕГИСЗ ключевой задачей является обеспечение всем государственным медицинским учреждениям условий для создания и ведения медицинской документации в электронном виде. На этом этапе целесообразно ограничение медицинской информации, передаваемой в ЕГИСЗ минимально необходимым набором данных, передаваемых в формате ИЭМК, т.к. оно обеспечивает формирование ядра и инфраструктуры системы, отработку ее базовых эксплуатационных характеристик, таких как: сбор и передачу данных, их хранение, анализ и статистическую обработку. Расширение целей ЕГИСЗ от сбора, анализа и использования информации в административно-статистических целях до формирования экосистемы цифрового здравоохранения позволит автоматизировать значительную часть действий с медицинской частью данных, обеспечить адаптивный мониторинг и контроль значительной части функций системы оказания медицинской помощи, в том числе качества и доступности медицинских услуг. Также это позволит формировать и развивать систему знаний и научных исследований, оптимизировать затраты на оказание медицинской помощи и предоставить пациенту дополнительные формы управления своим здоровьем.

Для реализации решений, основанных на интеллектуальном анализе, необходима консолидация данных и перевод их в структурированную форму, позволяющую трансформировать большие базы данных в базы знаний. Формирование деперсонифицированных баз



медицинских знаний с использованием русского языка достаточного объема необходимо для реализации процесса создания онтологий. Очевидно, что чем крупнее база знаний, тем более точный анализ смогут обеспечивать функциональные аппаратные модули для интеллектуальной обработки этих знаний, тем выше потенциал их эффективности в практическом здравоохранении, а соответственно и уровень принятия управленческих и медицинских решений. В этой связи путь на централизацию системы сбора и хранения медицинских данных и баз знаний на сегодняшний день является доминирующим трендом. В то же время программные модули, использующие эти знания, скорее всего будут развиваться по функциональному принципу: модульная архитектура интеллектуальных приложений позволит более полно реализовать их задачи и опережающие запросы отрасли, обеспечит конкурентную и экономически эффективную модель разработки и внедрения.

В схематичном виде модель информационной экосистемы общественного здравоохранения нового типа представлена на диаграмме №1.



Диаграмма 1. Информационная экосистема здравоохранения

В предложенной архитектуре информационной экосистемы здравоохранения базы медицинских знаний и онтологий непрерывно формируются и пополняются с использованием гибридных подходов: наряду с иерархическим принципом формирования централизованных ведомственных медицинских информационных систем, таких как ЕГИСЗ, они используют открытую распределенную архитектуру, при которой знания в этих базах группируются вокруг конкретного физического лица с возможностью их пополнения другими участниками экосистемы, в том числе и самим физическим лицом, а также приборами к нему привязанными. Облачная архитектура и современные технологии сбора, обработки, и хранения указанных данных позволяют собрать их, структурировать и использовать для последующего анализа. Собранные данные затем деперсонифицируются, трансформируются в специальный формат и поступают в базы медицинских знаний.

Программные комплексы используют базы знаний и онтологии для обработки с помощью методов искусственного интеллекта и возвращают результаты посредством автоматизированных интеллектуальных сервисов. Они используются при оказании медицинской помощи, администрировании системы общественного здравоохранения, стимулировании положительной динамики общественного и индивидуального здоровья. Обратная связь в конкретных медицинских эпизодах и организационно-финансовых, административных воздействиях также фиксируется. Результаты указанных рекомендаций или воздействий и общий результат законченного эпизода (проекта) возвращается в базу знаний, а также используется для дальнейшего обучения интеллектуальных систем.

Представленная экосистема цифрового здравоохранения и уровень развития интеллектуальных программных средств уже в обозримой перспективе (5–10 лет) позволит



проектировать и внедрять такие сервисы, как интеллектуальная система поддержки принятия врачебных решений, формализация простых административных и медицинских процедур для прямого контакта с человеком: мониторинг экстремальных отклонений состояния здоровья, логистика взаимодействия с медицинским персоналом и контроль медицинских назначений, интерактивное взаимодействие в случаях первой медицинской помощи, образовательные, просветительские и обучающие программы, профилактика здорового образа жизни и других однозначных и хорошо формализуемых процессов.

В дальнейшем, по мере накопления и развития медицинских знаний и онтологий, практического совершенствования программно-аппаратных средств, использующих методы искусственного интеллекта, становится возможным более глубокая их интеграция в систему оказания медицинской помощи. В частности, это позволит задействовать и использовать интеллектуальные симуляторы таких органов чувств, как зрение (распознавание и диагностика), слух, обоняние для непосредственного участия в процессе оказания медицинской помощи.

В настоящее время в сфере здравоохранения мы наблюдаем практически параллельное развитие двух основных групп информационных систем по типу используемого математического аппарата:

1. Создание, поддержка и развитие единой информационной инфраструктуры. С этой задачей непосредственно связаны управленческие задачи государственного статистического наблюдения и мониторинга, планирование и прогнозирование комплексных организационных и финансово-экономических инструментов и их результатов, в том числе страхового медицинского обеспечения, административно-хозяйственной деятельности, выполнения целого ряда показателей, в том числе

объема, качества и эффективности медицинской помощи.

Математические инструменты и модели, используемые для реализации указанных задач, довольно хорошо описаны, а специализированное программное обеспечение, разработанное и внедренное в различных странах мира (в том числе и в России) постоянно совершенствуется в рамках уже созданных национальных информационных сегментов в сфере здравоохранения, их законодательного и нормативно-правового обеспечения, международной интеграции.

2. Создание принципиально новых информационных инструментов, позволяющих автоматизировать целый ряд административных функций системы здравоохранения, в первую очередь направленных на взаимодействие с пациентом и обеспечение проведения различных медицинских манипуляций, а также позволяющих проводить сложный анализ. Указанные системы создаются, как правило, в виде функциональных модулей или приложений и используют данные и знания, получаемые с помощью экосистемы первой группы. При разработке указанных систем активно используются различные методы интеллектуального анализа. В перспективе программные средства этой группы будут интегрированы в группу 1.

В условиях технологического отставания от лидеров (национальные информационные системы США, Великобритании, Германии, Южной Кореи) целесообразно применять «опережающую» модель развития. Она заключается в параллельном развитии с последующим объединением трех основных технологических трендов:

- развитие инфраструктурной компоненты экосистемы информационного обеспечения здравоохранения;

- структурирование данных методами data mining с применением тиражируемых BI





систем, в том числе и отечественного производства для получения статистики и аналитики и использования ее в оперативном управлении и планировании;

- развитие инфраструктуры, методологии и прикладного применения систем с использованием искусственного интеллекта с функционалом, ориентированным на внедрение в наиболее чувствительных участках системы общественного здравоохранения. Наиболее актуальными представляются задачи по созданию систем персонифицированного взаимодействия (системы, направленные на автоматизацию логистики медицинских процедур и ЗОЖ, информационное интерактивное взаимодействие в хорошо формализованных областях), а также систем, направленных на мониторинг, анализ и прогноз в наиболее опасных медицинских нозологиях: болезни системы кровообращения, новообразования, заболевания органов пищеварения и дыхания.

Комплексные и скоординированные усилия, направленные одновременно в трех указанных направлениях позволят сократить, а при удачном стечении обстоятельств и ликвидировать к 2030 году технологическое отставание, обеспечить систему общественного здравоохранения мощным и современным драйвером роста и стимулировать отечественных

разработчиков к созданию и внедрению конкурентоспособных решений. Пример архитектуры комплексной информационной системы, использующей сочетание различных моделей анализа приведен на *рис. 2* [3]

Построение программы и соответствующих дорожных карт на такой длительный период должно учитывать возможность появления новых открытий и технологий, о которых мы пока не знаем, использование потенциала лучших современных решений (возможность интероперабельности и соответствие международным стандартам), возможности развития по различным сценариям, в зависимости от уровня развития технического прогресса, уровня доступности передовых технологий и ресурсных возможностей для их реализации, включая комплексный синергетический эффект различных программ развития.

3. Наиболее распространенные элементы математического аппарата, используемые при описании моделей интеллектуального анализа в медицине.

Успешные проекты в области создания сложных медицинских информационных систем – это интегральный результат работы комплексных команд специалистов, обладающих высокими компетенциями, часть из которых способна консолидировать коммуникации



Рис. 2. Обобщенная архитектура системы поддержки принятия решений



в смежных областях знаний. Как правило, в составе указанных команд присутствуют постановщики задач, профильные эксперты, математики, специалисты, обладающие компетенциями в области различных программных средств. Успешное проектирование, разработка и внедрение информационной системы – результат довольно длительного процесса взаимодействия ядра команды со всей экосистемой.

В этой связи мы посчитали целесообразным в рамках настоящей работы раскрыть смысл ряда ключевых понятий, используемых не только в процессе создания и внедрения современных информационных систем, но и все более активно употребляемых в медицинском сообществе. Многие из представленных понятий и методов развиваются и не имеют на данный момент четко определенных границ, а также по-разному трактуются представителями различных научных школ. Задача данного раздела – популярно и по возможности математически строго представить специалистам в области здравоохранения суть понятий и методов, используемых при описании отраслевых информационных систем.

Использование математического аппарата для описания, классификации и прогнозирования медицинских процессов применялось с момента образования системных знаний в медицине. Врачи всегда вели статистический учет, использовали его данные для прогнозирования и планирования своей деятельности. Диагностика, наблюдение, а также весь процесс лечения и реабилитации также описывался определенным набором данных, который использовался в медицинской науке для выявления закономерностей и структурирования знаний.

Начиная с пятидесятых годов 20 века, с появлением ЭВМ в здравоохранении начинают применяться более сложные математические методы. Прогнозы аналитического агентства Accenture[13] говорят о том, что к 2024 г.

здравоохранение будет лидировать по количеству внедрений программных продуктов, основанных на моделях искусственного интеллекта, находящихся на переднем крае математической мысли и использующими значительные вычислительные ресурсы. На сегодняшний день при моделировании медицинских процессов используется около ста математических моделей и закономерностей. В данном разделе мы опишем только наиболее часто используемые в здравоохранении методы и понятия, обеспечивающие сквозные коммуникации разработчиков, организаторов и конечных пользователей.

*Под **Знаниями** понимается совокупность сведений, которая образует целостное описание, соответствующее некоторому уровню осведомленности об описываемом вопросе, предмете, проблеме и т.д.*

*Принятие решений требует **Информации**, которая основана на данных.*

***Данные** обеспечивают получение информации, которая поддерживает решения.*



***Онтология** – система понятий некоторой предметной области, которая представляется как набор сущностей, соединенных различными отношениями*

В настоящее время в сфере здравоохранения для построения программных средств интеллектуального анализа данных чаще других используются следующие математические методы и инструменты (по частоте их использования):

✓ **Интеллектуальный анализ данных** (Data Mining) – извлечение, обработка и анализ данных, в том числе с использованием



моделей искусственных нейронных сетей. В настоящей работе мы приведем общие представления о следующих архитектурах нейронных сетей:

- **Сверточные нейронные сети** (Convolutional neural network, далее CNN) и
- **Рекуррентные нейронные сети** (Recurrent neural networks, далее RNN): распознавание, эмоции и мимика на видео (RNN) сложный интеллектуальный анализ.

✓ **Байесовский подход** (Bayesian Analysis) используется в машинном обучении и статистике, а также в совокупности с нейронными сетями. Классический пример из медицины: оценка двух равновероятных диагнозов при проявлении схожих симптомов и наличия данных, связывающих симптоматику и диагноз.

✓ **Регрессионный анализ** (Regression analysis) – распространенный и хорошо изученный метод исследования статистических данных. Отдельно остановимся на **Временных рядах (Time series)**, используемых для описания и анализа циклических и нециклических колебаний (временных и т. д.).

Интеллектуальный анализ данных (Data Mining) – извлечение, обработка и анализ данных.

Несмотря на консервативность медицины во многих ее аспектах, технология Data Mining в последние годы активно применяется для различных исследований и в этой сфере человеческой деятельности. Традиционно для постановки медицинских диагнозов используются экспертные системы, которые построены на основе символьных правил, сочетающих, например, симптомы пациента и его заболевание. С использованием Data Mining при помощи шаблонов можно разработать базу знаний для экспертной системы.

Методы Data Mining используются при решении следующих задач:

- классификация: в результате решения задачи классификации обнаруживаются признаки, которые характеризуют группы объектов исследуемого набора данных – классы; по этим признакам новый объект можно отнести к тому или иному классу. Для решения задачи классификации могут использоваться такие инструменты, как: байесовские сети (Bayesian Networks); деревья решений; нейронные сети (neural networks) и др.;

- кластеризация: является логическим продолжением идеи классификации. Особенность кластеризации заключается в том, что классы объектов изначально не predetermined. Кластеризация предназначена для разбиения совокупности объектов на однородные группы (кластеры или классы) и дальнейшего их анализа;

- прогнозирование: в результате решения задачи прогнозирования на основе особенностей исторических данных оцениваются пропущенные или же будущие значения целевых численных показателей. Для решения таких задач широко применяются методы математической статистики, нейронные сети и др.;

- ассоциация: в ходе решения задачи поиска ассоциативных правил отыскиваются закономерности между связанными событиями в наборе данных. Отличие ассоциации от предыдущих задач Data Mining в том, что поиск закономерностей осуществляется не на основе свойств анализируемого объекта, а между несколькими событиями, которые происходят одновременно.

В общем случае процесс интеллектуального анализа и обработки данных состоит из следующих этапов: постановка задачи, отбор данных, очистка, обогащение, кодирование, извлечение знаний и сообщение. Процесс носит итеративный характер: на каждом шаге возможен возврат на один или более этапов. Например, на этапе извлечения знаний можно вернуться к этапу очистки. [2]



Байесовский подход заключается в переходе от априорных знаний к апостериорным с учетом наблюдаемых явлений. Основан на формуле Байеса. Позволяет оперировать на непрерывном интервале между «истиной» и «ложью», присваивая событиям промежуточные значения вероятности. При большом объеме знаний показывает математически корректные результаты. Благодаря этим свойствам активно используется в машинном обучении и статистике, а также в совокупности с нейронными сетями. Классический пример из медицины: оценка двух равновероятных диагнозов при проявлении схожих симптомов и наличия данных, связывающих симптоматику и диагноз.

Формулу Байеса в простейшем виде можно представить следующим образом:

Пусть $x = (x_1 \dots x_i)$ – экспериментальные данные (симптомы) – апостериорные значения, а H – множество независимых гипотез (диагнозов), образующих полную систему, суммарная вероятность которых равна 1 – априорные значения. $P(x|H)$ – вероятность того, что x произойдет, при условии, что H – произошло. Это значение называют значением правдоподобия. В случае непрерывных величин вместо значения вероятности используется функция плотности вероятности $p(x)$.

В этом случае формулу Байеса можно записать следующим образом:

$$p(H|x) = \frac{p(x|H)p(H)}{p(x)} \quad (1)$$

Нейронная сеть – это система программных нейронов, соединенных и взаимодействующих между собой.

Области применения нейронных сетей:

- Классификация.
- Прогнозирование.
- Распознавание.

Программный Нейрон – это вычислительная единица, которая получает информацию, присваивает ей весовые коэффициенты (W),

производит над ней простые вычисления и передает ее дальше (рис. 3). Они делятся на три основных типа: входной, скрытый и выходной. В том случае, когда нейросеть состоит из большого количества нейронов, вводят термин слоя. Соответственно, есть входной слой, который получает информацию, скрытые слои, которые ее обрабатывают и выходной слой, который выводит результат. В процессе обучения нейроны скрытых слоев самостоятельно изменяют весовые коэффициенты связей.

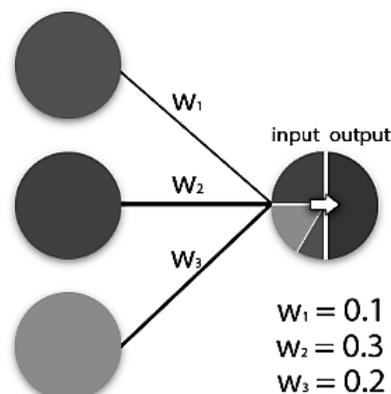


Рис. 3. Пример структуры программного нейрона [15]

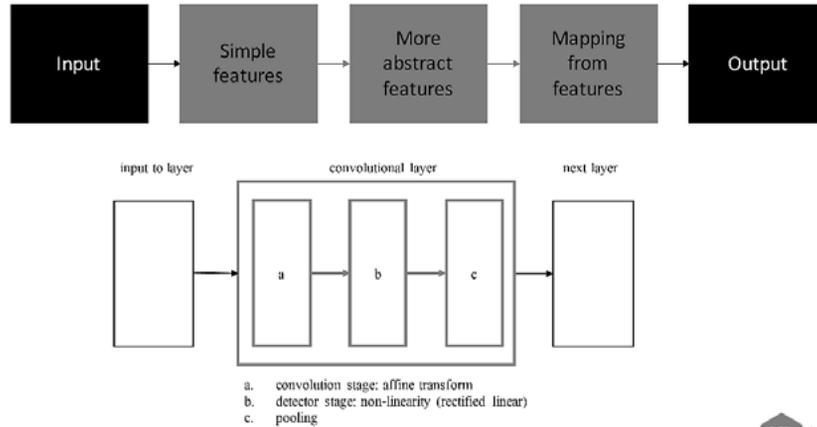
Сверточная нейронная сеть

На рис. 4 черным обозначены те части, которые не учатся, все остальные слои способны обучаться (самостоятельно изменять весовые коэффициенты связей). Существует множество определений того, что находится внутри каждого сверточного слоя. Одно из принятых обозначений – один слой с тремя компонентами разделяют на этап свертки (convolution stage), стадия распознавания (detector stage) и объединение (pooling stage).

Название архитектура сети получила из-за наличия операции свёртки, суть которой в том, что каждый фрагмент изображения умножается на матрицу (ядро) свёртки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения.



Архитектура



(Goodfellow, 2016)



Рис. 4. Архитектура сверточной нейронной сети [14]

Рекуррентные нейронные сети (Recurrent neural networks) используют тогда, когда важно распознавать динамические задачи. В частности, они применяются для распознавания видеоизображений во временной динамике. Эта задача актуальна для распознавания эмоций, визуальных симптомов

и других задач, для которых важен анализ изменений, пусть и краткосрочных. Основное отличие от обычных нейронных сетей – наличие обратной связи с предыдущими состояниями системы или объекта, информация о которых передается на один из слоев сети. Структура рекуррентной сети представлена на рис. 5.

Рекуррентные сети для распознавания эмоций на видео

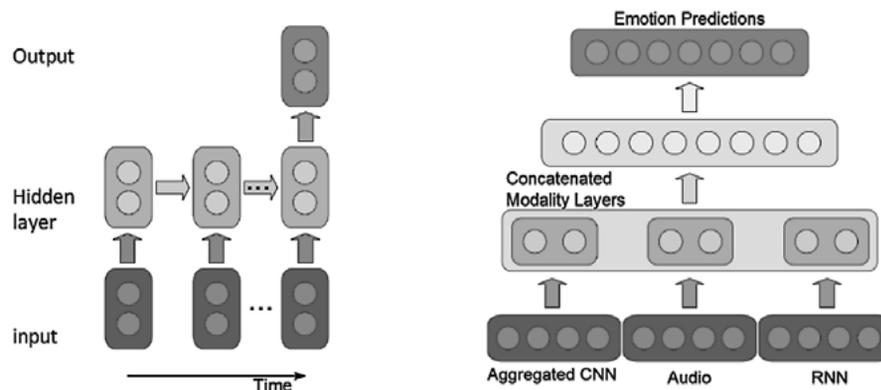


Рис. 5. Структура рекуррентной нейронной сети [14]



Регрессионный анализ

Регрессией называется зависимость среднего значения одной случайной величины от некоторой другой (или от нескольких случайных величин), а регрессионным анализом – раздел математической статистики, объединяющий прикладные методы исследования регрессионных зависимостей.

«Этот термин в статистике впервые был использован Френсисом Гальтоном (1886) в связи с исследованием вопросов наследования физических характеристик человека. В качестве одной из характеристик был взят рост человека; при этом было обнаружено, что в целом сыновья высоких отцов, что не удивительно, оказались более высокими, чем сыновья отцов с низким ростом. Более интересным было то, что разброс в росте сыновей был меньшим, чем разброс в росте отцов. Так, проявлялась тенденция возвращения роста сыновей к среднему (regression to mediocrity), то есть «регресс». Этот факт был продемонстрирован вычислением среднего роста сыновей отцов, рост которых равен 56 дюймам, вычислением среднего роста сыновей отцов, рост которых равен 58 дюймам, и т. д. После этого результаты были изображены на плоскости, по оси ординат которой откладывались значения среднего роста сыновей, а по оси абсцисс – значения среднего роста отцов. Точки (приблизённо) легли на прямую с положительным углом наклона меньше 45°; важно, что регрессия была линейной.»

Задачей регрессионного анализа является нахождение «наилучшей» функции f , описывающей зависимость y от x . Оценка в основном производится по методу наименьших квадратов, либо по методу максимального правдоподобия (что возможно только при известном распределении величин y), методу моментов, квантильным регрессиям.

При использовании регрессионного анализа важно правильно выбрать вид и степень сложности регрессионной модели. Классический путь состоит в учете биологических,

физических и других предпосылок, а качество полученной модели оценивается по величине остаточных отклонений. Разработаны критерии оптимальной сложности регрессии в зависимости от дисперсии остаточных отклонений и объема выборки.

Если x_i и y_i – наблюдаемые величины, e_i – случайная ошибка с нулевым математическим ожиданием, то регрессия записывается в виде:

$$y_i = f(x_i) + e_i, \quad i = 1, 2, \dots, N,$$

где f – функция регрессии.

Если x_i – скалярная величина (число), то регрессия называется парной (связывающей пару случайных величин), если x_i – вектор, то множественной.

Временные ряды (Time series). Под временным рядом понимают упорядоченную во времени последовательность величин $Y(t) = Y[1], Y[2], \dots, Y[t]$, где t – момент времени [22, 23]. Принято выделять несколько основных элементов временного ряда: сезонная компонента – $S(t)$, циклическая компонента – $C(t)$, детерминированная компонента (или трендовая) – $G(t)$ и случайная (или стохастическая) компонента – $E(t)$. Циклическая компонента отличается от сезонной тем, что продолжительность цикла как правило больше, чем один сезонный период, и разные циклы могут иметь разную продолжительность.

Временные ряды используются для описания заболеваний, обладающих циклической симптоматикой: психические расстройства, хронические заболевания, сезонные обострения аллергического характера.

Формулу временного ряда можно представить, как:

$$Y(t) = C(t) + S(t) + G(t) + E(t) \quad (2)$$

Основной задачей прогнозирования временного ряда является поиск наилучшей оценки эмпирических данных с целью предсказания поведения наблюдаемой величины $Y(t+h)$, где h – это шаг прогноза.

Для определения модели временного ряда последовательно элиминируют сезонные



и циклические компоненты, далее детерминированную составляющую, и, в конечном итоге, обрабатывают остаточный ряд.

Одним из первых этапов анализа временного ряда является его визуальное изучение с целью выявления тренда либо сезонной составляющей.

4. Заключение

Основным отличием персонцентрированного здравоохранения является наличие информационной экосистемы, позволяющей комплексно выстраивать прямые адаптивные связи человека, как центрального объекта, со специалистами соответствующих общественных институтов. Именно поэтому вектор развития интеллектуальных систем в первую очередь ориентирован на создание среды взаимодействия для связей, где второй частью связи выступает человек, и где он еще долгое время не может быть заменен машиной: здравоохранение, наука, образование.

Распределенная и многовекторная модель внедрения в такой чувствительной области, как здравоохранение, может существовать только при наличии соответствующей законодательной и нормативной базы, направленной на обеспечение требований, предъявляемых к программным продуктам медицинского назначения, как с точки зрения технологической (надежность, безопасность, интероперабельность), так и с точки зрения медицинской (обеспечение исполнения отраслевых нормативов, стандартов и порядков оказания медицинской помощи, клинических рекомендаций, специальных нормативных требований, специфических этических требований). Новые технологические вызовы формируют однозначный запрос на совершенствование существующей законодательной и нормативно-правовой базы.

Внедрение представленной модели обладает целым рядом преимуществ, которые за отсутствием значимого практического опыта

пока не были количественно описаны, вместе с тем являются довольно очевидными:

1. Оптимизация взаимодействия физического лица и системы общественного здравоохранения, позволяющая в обозримый срок практически полностью автоматизировать всю административную часть процесса. Это повлечет за собой значительное снижение, а возможно и устранение кадрового дефицита в данной области и качественно ускорит все процессы.

2. Повышение квалификации уже существующих специалистов – врачей, что позволит значительно сократить время оказания скорой и неотложной помощи, а также обеспечит доступность современных медицинских методов и процедур, сократит транспортные и иные издержки, связанные с их оказанием.

3. Оперативный анализ и интеллектуальная обратная связь обеспечит значительное повышение точности и эффективности планирования ресурсов на всех уровнях системы, возможность прогнозирования отклонений и формирования превентивных защитных процедур.

4. Возможность мотивации человека к ведению здорового образа жизни путем создания системы мониторинга, позволяющей контролировать выполнение соответствующих действий пациентом и его окружением.

5. Развитие научных исследований в части здоровья человека с использованием инструментов, позволяющих моделировать процессы, ранее требовавшие значительных временных и других ресурсов.

6. Разработка и внедрение нового класса персонифицированных медицинских препаратов.

7. Внедрение гибкой системы страхового обеспечения здравоохранения, оптимизация ресурсов и формирование персональных мотивационных механизмов.

8. Формирование единого центра национальной компетенций и координации развития информатизации, статистики, аналитики и логистики здравоохранения.



ЛИТЕРАТУРА



1. Осипов Г.С. «Методы искусственного интеллекта» Физматлит, 2011.
2. Барсемян А.А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP: учеб. Пособие / А.А. Барсемян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. – СПб.: БХВ-Петербург, 2007.
3. Чернышева Г.Ю. (составитель). Интеллектуальный анализ данных. Саратов: СГСЭУ, 2012.
4. Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning. The MIT Press, 2016 г.
5. Степанова М.Г., Самодумкин С.А., Лемешева Т.Л. Математические методы диагностики в медицинских интеллектуальных системах. Минск: БГУИР, 2001.
6. Федеральный закон от 21 ноября 2011 г. № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации».
7. Федеральный закон от 29 июля 2017 г. № 242-ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации по вопросам применения информационных технологий в сфере охраны здоровья».
8. Приказ Министерства здравоохранения и социального развития Российской Федерации от 28 апреля 2011 г. № 364.
9. Интегрированная электронная медицинская карта. Описание стандартизированных электронных медицинских документов (СЭМД). // URL: <http://portal.egisz.rosminzdrav.ru/materials/46>. Дата обращения: 04.09.2017.
10. Кузнецов П.П. Персональный менеджер здоровья: требования к информационному обеспечению. Доклад на конференции IT&Med, Москва, 2017. Адрес страницы: // URL: http://itmedforum.ru/sites/default/files/10.25-10.40_kuznecov.pdf. Дата обращения: 04.09.2017.
11. Шифрин М.А. 4П-медицина – синергия биомедицины и информатики. Доклад на конференции IT&Med, Москва, 2017. Адрес страницы: // URL: http://itmedforum.ru/sites/default/files/9.40-9.55_shifrin.pdf. Дата обращения: 04.09.2017.
12. Выступление Министра Вероники Скворцовой на заседании Совета по стратегическому развитию и приоритетным проектам под председательством Президента Российской Федерации Владимира Путина. Адрес страницы: // URL: <https://www.rosminzdrav.ru/news/2017/03/21/5225-vystuplenie-ministra-veroniki-skvortsovoy-na-zasedanii-soveta-po-strategicheskomu-razvitiyu-i-prioritetnym-proektam-pod-predsedatelstvom-prezidenta-rossiyskoy-federatsii-vladimira-putina>. Дата обращения: 04.09.2017.
13. Artificial intelligence: Healthcare's new nervous system. Accenture consulting. // URL: <https://www.accenture.com/us-en/insight-artificial-intelligence-healthcare>. Дата обращения: 04.09.2017.
14. Нейронные сети: практическое применение. Наталия Ефремова, research scientist, NtechLab // URL: <https://habrahabr.ru/post/322392/> Дата обращения: 2017.09.04.
15. Нейронные сети для начинающих. // URL: <https://habrahabr.ru/post/312450/>. Дата обращения: 04.09.2017.
16. Лебедев С.В., Жукова Н.А. Слияние медицинских данных на основе онтологий. Онтология проектирования, Т. 7, № 2 (24)/2017. – С. 145–159.
17. Нейросеть в 11 строчек на Python. // URL: <http://iamtrask.github.io/2015/07/12/basic-python-network/>. Дата обращения: 04.09.2017.





18. Мухин Ю.Ю., Коссова Е.В. Подходы к оценке полной (совокупной) стоимости владения (ТСО) для медицинских информационных систем. Экономические критерии и их влияние на оптимизацию информационной структуры медицинской организации. // Информационно-измерительные и управляющие системы. – 2010, Т. 8, № 12.
19. Мухин Ю.Ю., Лебедев Г.С. Подходы к параметрической оценке и сопоставлению функций медицинских информационных систем // Информационные и измерительные управляющие системы. М.: 2013, № 10, Т. 11. – С. 19–31.
20. Ветров Д.П., Кропотов Д.А. Байесовские методы машинного обучения. Учебное пособие. М.: 2007.
21. Гусев А.В., Зарубина Т.В. Поддержка Принятия врачебных решений в Медицинских информационных системах Медицинской организации. «Врач и информационные технологии», 2017. – № 2, – С. 60–72.
22. Чураков Е.П. Прогнозирование эконометрических временных рядов. – М.: Финансы и статистика, 2008.
23. Егоров Д.Б., Санников А.Г., Захаров С.Д. Шваб Д.В., Валеев Р.И. Анализ и прогнозирование общественно опасных действий психически больных современными математическими методами. «Врач и информационные технологии», 2017. – № 3.
24. Константинова Н.С., Митрофанова О.А. Онтологии как системы хранения знаний. // Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы», 2008. – 54 с.
25. Мухин К.Ю. Противоречивость долгосрочных и краткосрочных целей: пути согласования // Культура и туризм как инструменты повышения человеческого потенциала нации: Труды Всероссийской научно-практической конференции. Национальная академия туризма, Санкт-Петербург, 14–15 апреля 2016 г. – С.-Пб.: Д.А.Р.К., 2016. – 416 с., 275–280 с.
26. DAF Community Presentation 10-19-2016.
27. Conceptualizing a Data Infrastructure for the Capture, Use, and Sharing of Patient-Generated Health Data in Care Delivery and Research through 2024. Draft White Paper for a PGHD Policy Framework. Accenture, 2017
28. Beck K, Beedle M., Bennekum A., Cockburn A., Cunningham W., Fowler M., Grenning J, Highsmith J, Hunt A., Jeffries R., Kern J., Marick B., Martin R., Mellor S., Schwaber K., Sutherland J, Thomas D. Manifesto for Agile Software Development. 2001.
29. Schwaber K. and Sutherland J. The Definitive Guide to Scrum: The Rules of the Game, 2013.
30. // URL: <https://scrumtrek.ru/blog/agile-transformatsiya-i-korporativnaya-kultura/>. Дата обращения: 07.09.2017.
31. A Robust Health. Data Infrastructure. AHRQ Publication № 14–0041-EF. April 2014
32. // URL: <https://www.ohdsi.org/data-standardization/>. Дата обращения: 04.09.2017.
33. //URL: [https://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D1%8F_\(%D0%BC%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0\)](https://ru.wikipedia.org/wiki/%D0%A0%D0%B5%D0%B3%D1%80%D0%B5%D1%81%D1%81%D0%B8%D1%8F_(%D0%BC%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0)). Дата обращения: 04.09.2017.

Неопубликованные материалы (Готовятся к выходу):

34. Мухин К.Ю. Agile-эволюция проектных подходов как вынужденное требование времени. // «Инициативы XXI века», 2017.
35. Мухин К.Ю. Гибкая имплементация стандартов управления проектами P2M PMAJ & P5 GPM в контексте инновационного предпринимательства. // «Управленческие науки в современном мире», сборник докладов международной научной конференции, 2017.