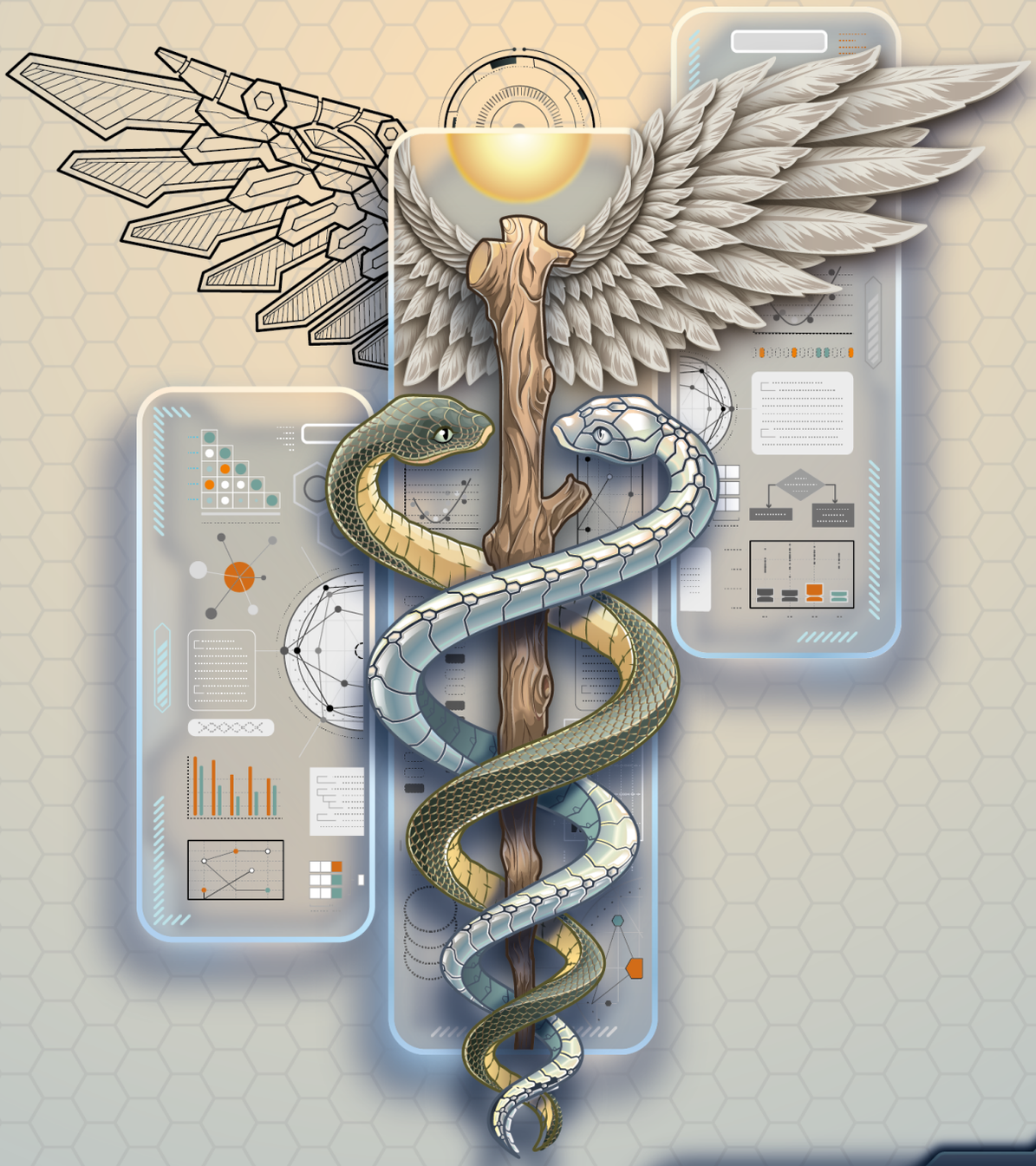
НАУКА О ДАННЫХ
и искусственный интеллект в медицине



А.Ю. КУЗЬМЕНКОВ
А.Г. ВИНОГРАДОВА
И.В. ТРУШИН
А.А. АВРАМЕНКО
М.Ю. КУЗЬМЕНКОВ

А.Ю. КУЗЬМЕНКОВ
А.Г. ВИНОГРАДОВА
И.В. ТРУШИН
А.А. АВРАМЕНКО
М.Ю. КУЗЬМЕНКОВ

НАУКА О ДАННЫХ
И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В МЕДИЦИНЕ

e-noto
Москва, 2026

УДК 004.8:61
ББК 5с51
Н 34

Н 34 **Наука о данных и искусственный интеллект в медицине.** / А.Ю. Кузьменков, А.Г. Виноградова, И.В. Трушин, А.А. Авраменко, М.Ю. Кузьменков — М.: Е-ното, 2026. — 744 с.

ISBN 978-5-906023-48-3

«Наука о данных и искусственный интеллект в медицине» — уникальное практическое руководство, объединяющее современные подходы анализа больших данных и искусственного интеллекта с прикладными аспектами здравоохранения. Эта книга охватывает весь цикл работы с биомедицинскими данными: от планирования клинических исследований и сбора информации до углубленного статистического анализа и публикации результатов. Читатель познакомится с основами программирования на языке R, различными системами управления базами данных и алгоритмами машинного обучения, а также получит возможность отработать полученные навыки на реальных наборах биомедицинских данных. Издание станет незаменимым помощником как для опытных лечащих врачей и ученых-исследователей, так и для студентов вузов и специалистов из разных областей, желающих овладеть новыми компетенциями.

Рецензенты:

Колбин Алексей Сергеевич — доктор медицинских наук, профессор, заведующий кафедрой клинической фармакологии и доказательной медицины ФГБОУ ВО «Первый Санкт-Петербургский государственный медицинский университет имени академика И.П. Павлова» Министерства здравоохранения Российской Федерации

Середа Андрей Петрович — доктор медицинских наук, профессор кафедры восстановительной медицины, лечебной физкультуры и спортивной медицины, курортологии и физиотерапии Академии постдипломного образования ФГБУ «Федеральный научно-клинический центр специализированных видов медицинской помощи и медицинских технологий Федерального медико-биологического агентства» Российской Федерации

Балыкина Юлия Ефимовна — кандидат физико-математических наук, доцент кафедры математического моделирования энергетических систем ФГБОУ ВО «Санкт-Петербургский государственный университет»

УДК 004.8:61
ББК 5с51
Н 34

ISBN 978-5-906023-48-3

«Большое дерево вырастает из маленького,
девятиэтажная башня начинает строиться из горстки земли,
путешествие в тысячу ли начинается с одного шага...»
Лао-цзы

«...изучение науки о данных начинается со следующей страницы.»
Коллектив авторов

Оглавление

Авторы

Кузьменков Алексей Юрьевич — доктор медицинских наук, заместитель директора по стратегическим разработкам НИИ антимикробной химиотерапии, профессор кафедры микробиологии, основатель, руководитель и преподаватель проекта «Цифровая кафедра» с курсом «Технологии науки о данных в медицине» ФГБОУ ВО «Смоленский государственный медицинский университет» Министерства здравоохранения Российской Федерации.

Виноградова Алина Геннадьевна — кандидат медицинских наук, старший научный сотрудник отдела стратегических разработок и проектов НИИ антимикробной химиотерапии, доцент кафедры микробиологии, научный руководитель и преподаватель проекта «Цифровая кафедра» с курсом «Технологии науки о данных в медицине» ФГБОУ ВО «Смоленский государственный медицинский университет» Министерства здравоохранения Российской Федерации.

Трушин Иван Витальевич — младший научный сотрудник отдела стратегических разработок и проектов НИИ антимикробной химиотерапии, преподаватель проекта «Цифровая кафедра» с курсом «Технологии науки о данных в медицине» ФГБОУ ВО «Смоленский государственный медицинский университет» Министерства здравоохранения Российской Федерации.

Авраменко Андрей Алексеевич — младший научный сотрудник отдела стратегических разработок и проектов НИИ антимикробной химиотерапии, технический организатор проекта «Цифровая кафедра» с курсом «Технологии науки о данных в медицине» ФГБОУ ВО «Смоленский государственный медицинский университет» Министерства здравоохранения Российской Федерации.

Кузьменков Михаил Юрьевич — кандидат юридических наук, специалист по машинному обучению отдела стратегических разработок и проектов НИИ антимикробной химиотерапии ФГБОУ ВО «Смоленский государственный медицинский университет» Министерства здравоохранения Российской Федерации.

Предисловие

Наука о данных (Data Science) — это междисциплинарная область, которая объединяет методы, алгоритмы и программные средства, разработанные для извлечения полезной информации из большого объема данных, хранящихся в различных форматах. Соответственно, основной целью Data Science является обнаружение в данных скрытых закономерностей или получение знаний — т.е. такой информации, которая может быть использована для принятия обоснованных и оптимальных решений.

Наука о данных включает в себя различные дисциплины: математическую статистику; системы искусственного интеллекта; управление базами данных; сбор, анализ и визуализацию данных и многие другие. Наибольшее развитие рассматриваемая наука получила с начала 2010-х годов благодаря взрывному росту объемов хранимой в мире информации и вычислительной мощности компьютерной техники. Также немаловажным стимулом к ее развитию стало совершенствование в эти годы алгоритмов машинного обучения, таких как нейронные сети, бустинг, бэггинг и др. На сегодняшний день популярность науки о данных вышла уже далеко за пределы чисто технических сфер, и ее инструменты стали широко применяться в маркетинге, логистике, государственном управлении и — наиболее интересующем нас — здравоохранении.

В медицинской науке, в частности в биомедицинских исследованиях, статистические методы анализа данных использовались еще с середины XX века, что естественным образом создало благоприятную почву для внедрения подходов науки о данных в здравоохранение. К тому же, на сегодняшний день здравоохранение является одной из крупнейших, быстрорастущих и наиболее приоритетных отраслей мировой экономики, которая генерирует огромное количество данных: диагностических, генетических, эпидемиологических, фармакологических, финансово-страховых, медико-социальных и др.

Сегодня технологии Data Science и искусственного интеллекта переосмысливают сферу медицинских услуг, предоставляя уникальные возможности для улучшения диагностики, лечения и профилактики заболеваний. Однако, зачастую, специалисты здравоохранения не обладают достаточными компетенциями ни в сфере информационных технологий, ни в сфере статистического анализа биомедицинских данных для того, чтобы создавать конкурентоспособные продукты. Привлечение в отрасль специалистов по анализу данных (т.н. «дата-сайентистов») с техническим образованием тоже не является панацеей, т.к. для создания эффективных, а главное безопасных медицинских продуктов необходимо глубокое понимание предметной области, связанной со здоровьем человека.

С учетом вышесказанного, наш авторский коллектив, обладающий многолетним опытом преподавания медицинских дисциплин, разработки программных решений для практического здравоохранения, а также создания и реализации курса «Технологии науки о данных в медицине» в рамках проекта «Цифровая кафедра» Смоленского государственного медицинского университета, подготовил для специалистов в области здравоохранения, медицинской информатики и аналитики, а также других заинтересованных специалистов практическое руководство «Наука о данных и искусственный интеллект в медицине». Настоящее издание, делая акцент на практической значимости излагаемого материала, призвано дать читателям комплексное понимание науки о данных в медицине, включающее все важные информационно-технологические и исследовательские аспекты.

Практическое руководство «Наука о данных и искусственный интеллект в медицине» состоит из девяти глав. В первой главе приводятся основные понятия доказательной медицины и принципы организации биомедицинских исследований. Во второй главе рассматриваются основы программирования на языке R, а также алгоритмы сбора и обработки биомедицинских данных. В третьей главе подробно описываются системы управления базами данных, обеспечивающие хранение информации в цифровом структурированном виде. В четвертой главе приводится обзор разнообразных способов компьютерной визуализации, позволяющих наглядно отображать полезную информацию, скрывающуюся в биомедицинских данных. В пятой главе рассматриваются общие, а в шестой — частные вопросы статистического анализа, раскрывающие особенности применения методов математической статистики для формальной оценки и интерпретации биомедицинских данных, а также для формирования обоснованных заключений. В седьмой главе приводятся базовые понятия «Больших данных» (Big Data) и «Машинного обучения» (Machine learning), а также разбираются конкретные примеры реализации алгоритмов машинного обучения для анализа биомедицинских данных. В восьмой главе описываются популярные инструменты и технологии, позволяющие облегчить процесс организации, разработки и представления результатов исследовательских проектов в области науки о данных. Девятая глава дополняет практическую значимость настоящего издания, предлагая подборку реальных биомедицинских данных для отработки навыков, полученных в ходе прочтения.

Мы надеемся, что наш труд будет способствовать притоку в отечественное здравоохранение высококлассных специалистов в области науки о данных и искусственного интеллекта, которые выведут отрасль на новую невиданную высоту.

«Наука о данных и искусственный интеллект в медицине» — уникальное практическое руководство, объединяющее современные подходы анализа больших данных и искусственного интеллекта с прикладными аспектами здравоохранения. Эта книга охватывает весь цикл работы с биомедицинскими данными: от планирования клинических исследований и сбора информации до углубленного статистического анализа и публикации результатов. Читатель познакомится с основами программирования на языке R, различными системами управления базами данных и алгоритмами машинного обучения, а также получит возможность отработать полученные навыки на реальных наборах биомедицинских данных. Издание станет незаменимым помощником как для опытных лечащих врачей и учёных-исследователей, так и для студентов вузов и специалистов из разных областей, желающих овладеть новыми компетенциями.

ISBN 978-5-906023-48-3

приоритет

AMRhub.ru