
МАШИННОЕ ОБУЧЕНИЕ В МЕДИЦИНЕ

Обзор научной литературы

Автор

Красавцева Дарья

БЭАД223

Содержание

1	Forecasting individual risk for long-term Posttraumatic Stress Disorder in emergency medical settings using biomedical data: A machine learning multicenter cohort study. [1]	4
2	Usefulness of machine learning and deep learning approaches in screening and early detection of breast cancer. [2]	4
3	Machine learning research towards combating COVID-19: Virus detection, spread prevention, and medical assistance. [3]	5
4	Theranostic roles of machine learning in clinical management of kidney stone disease. [4]	5
5	Application of machine learning in dementia diagnosis: A systematic literature review. [5]	6
6	Machine learning and physical based modeling for cardiac hypertrophy. [6]	6
7	A novel approach to forecasting the mental well-being using machine learning. [7]	7
8	An overview of deep learning in medical imaging focusing on MRI. [8]	7
9	Utility of four machine learning approaches for identifying ulcerative colitis and Crohn's disease. [9]	8
10	Prediction of neonatal death in pregnant women in an intensive care unit: Application of machine learning models. [10]	8
11	Early identification of bloodstream infection in hemodialysis patients by machine learning. [11]	9
12	Chronic kidney disease prediction based on machine learning algorithms. [12]	9

**13 Predicting the onset of diabetes-related complications after
a diabetes diagnosis with machine learning algorithms. [13] 10**

1 Forecasting individual risk for long-term Posttraumatic Stress Disorder in emergency medical settings using biomedical data: A machine learning multicenter cohort study.[1]

В данной статье рассматривается возможность применения машинного обучения для прогнозирования развития Посттравматического Стрессового Расстройства (ПТСР) на основе данных биомедицинских исследований, собранных в течение 48 часов после травмы. В исследовании был использован алгоритм LGMM - Latent Growth Mixture modeling. Это статистический метод, используемый в рамках моделирования структурных уравнений для оценки траекторий роста. Результаты показали, что применение многочисленной классификации методов машинного обучения позволяет точно прогнозировать риски развития ПТСР на ранних этапах после травмы на основе биомедицинских данных. Это исследование также показывает, что вычислительное многомерное моделирование с использованием ML может способствовать лучшему пониманию биологических путей, лежащих в основе индивидуальных различий в уязвимости к долгосрочному ПТСР.

2 Usefulness of machine learning and deep learning approaches in screening and early detection of breast cancer. [2]

В этом исследовании представлен систематический и подробный анализ различных подходов и механизмов машинного обучения, используемых в процессе диагностики рака молочной железы. Для диагностики использовались результаты МРТ, маммографии и другие гибридные методы. Во второй части исследования были применены такие методы машинного обучения, как neural networks (CNN, RNN, GAN), deep learning и гибриды на наборе данных в первой части. Основываясь на результатах этого исследования, подход CNN является наиболее точным и широко используемым методом диагностики среди подходов, связанных с NNS и DL. В этом исследовании одной из основных проблем является невозможность применить предложенные подходы в ML к большим и реальным наборам данных, а также отсутствие эффективных клинических пара-

метров в процессе диагностики рака молочной железы. В будущем использование технологии машинного обучения в сочетании с другими технологиями может сделать диагностику этого заболевания намного более эффективной и рентабельной.

3 Machine learning research towards combating COVID-19: Virus detection, spread prevention, and medical assistance. [3]

В данной статье рассматривается роль машинного обучения в выявлении или скрининге, прогнозировании и оказании медицинской помощи в связи с COVID-19. В исследовании так же было рассмотрено применение ML для поиска лекарств/повторного использования и разработки вакцин. В исследовании были рассмотрены другие основанные на ML инструменты, такие как чат-боты и искусственный интеллект вещей (AIoT). Инструменты ML применялись на трех основных этапах: "скрининге" (рентгеновские снимки, связанные с COVID-19, и компьютерная томография), "отслеживании и прогнозировании" (для прогнозирования числа случаев и отслеживания контактов) и "медицинской помощи" (понимания последовательностей и структуры белков о вирусе и о том, можно ли найти лекарство для борьбы с ним с помощью лекарства или вакцины). Одна из основных проблем, с которой столкнулись исследователи - отсутствие доступных данных. Нехватка данных означала, что исследователям приходилось использовать такие методы, как увеличение объема данных, обучение передаче данных и тонкая настройка моделей для повышения точности прогнозирования. Модели были бы надежнее, если бы были построены на большем количестве данных.

4 Theranostic roles of machine learning in clinical management of kidney stone disease. [4]

В этой статье рассматривается применение машинного обучения для выявления почечнокаменной болезни, прогнозирования типа камня, определения подходящего метода лечения и прогнозирования терапевтического результата. Использование машинного обу-

чения в медицинской визуализации, такой как ультразвуковая эластография (UE), компьютерная томография и магнитно-резонансная томография (MRI), повышает точность диагностики и снижает вероятность человеческих ошибок в широком спектре областей медицины. Различные алгоритмы машинного обучения, включая XGBoost, CNN, метод на основе ансамбля, k-ближайших соседей, ANN, SVM, RF и ряд других методов, улучшили производительность систем за счет повышения точности и чувствительности диагностики каменнопочечной болезни. Для повышения точности прогнозирования каменнопочечной болезни с ML следует сочетать с клинической информацией и медицинской визуализацией лабораторные анализы.

5 Application of machine learning in dementia diagnosis: A systematic literature review. [5]

В данной статье были систематизированы исследования по применению машинного обучения в прогнозировании деменции. Результаты этого обзора литературы показывают, что исследование ML при деменции является многообещающей темой. Однако литература по применению алгоритмов ML при нейродегенеративных заболеваниях ограничена по сравнению с литературой по ИИ в здравоохранении. Исследование статей показало, что наиболее часто используемыми инструментами машинного обучения для решения задач классификации для идентификации болезни Альцгеймера и умеренных когнитивных нарушений (MCI) были метод опорных векторов (SVM), ансамблевые методы и CNN.

6 Machine learning and physical based modeling for cardiac hypertrophy. [6]

В данной статье рассматривается применение машинного обучения для прогнозирования роста и регенерации сердца у отдельных пациентов. В исследовании были использованы модели, основанные на случайных лесах, градиентном ускорении и нейронных сетях, используемых для отслеживания гипертрофии сердца. Модель

была обучена с использованием истории болезни пациента и текущего уровня здоровья сердца, и сравнивалась с физической моделью, использовавшей процедуру конечных элементов для моделирования развития гипертрофии сердца. В результате исследования модели машинного обучения работали гораздо быстрее, однако были менее точной по сравнению с физическими моделями. Для повышения точности ML-моделей необходимо больше данных.

7 A novel approach to forecasting the mental well-being using machine learning. [7]

Результаты данного исследования представляют собой метод решения задачи прогнозирования психического здоровья с использованием алгоритмов машинного обучения. Исследователи использовали контролируемое обучение, при котором подмножество было помечено для построения машины, а помеченные данные использовались для прогнозирования новых классов информации. Этот метод занимает больше времени и требует больших усилий, но он более эффективен, чем неконтролируемая классификация, при которой система обучается на больших наборах данных. Исследователи использовали данные с российского психологического форума, извлеченные признаки и маркировка были выполнены психиатрами вручную. В результате работы модели были выведены классы для послеродовой депрессии, одиночеству и проблемам с уверенностью. Исследователи должны уделять приоритетное внимание устранению искажений данных, повышению качества данных и рассмотрению дополнительных внешних источников данных для повышения точности.

8 An overview of deep learning in medical imaging focusing on MRI. [8]

В данной статье рассматриваются применение машинного обучения и глубокого обучения в медицинской визуализации и МРТ, и как глубокое обучение применялось ко всей цепочке обработки данных МРТ от получения до извлечения изображений, от сегмен-

тации до прогнозирования заболеваний. По мере того как исследователи машинного обучения приобретут больше опыта, станет легче классифицировать проблемы в соответствии с тем, какой подход к решению является наиболее разумным. Как только в повседневный рабочий процесс клиник войдет достаточное количество высокоэффективных программных систем, основанных на математике, информатике, физике и инженерии, признание других подобных систем, вероятно, возрастет.

9 Utility of four machine learning approaches for identifying ulcerative colitis and Crohn's disease. [9]

В данном ретроспективном исследовании различные методы машинного обучения применялись для идентификации язвенного колита и болезни Крона. Исследователи собрали данные о 414 пациентах с воспалительными заболеваниями кишечника (ВЗК), 423 здоровых людях контрольной группы и 344 пациентах с заболеваниями кишечника без ВЗК. 70% данных было использовано для обучения, 30% для тестирования и еще одну группу использовали для внешней верификации. Для оценки эффективности диагностики и прогнозирования моделей ML использовалась площадь под кривой рабочих характеристик. Данное исследование было проведено со многими ограничениями, необходимы более новые ML-модели и больший набор данных.

10 Prediction of neonatal death in pregnant women in an intensive care unit: Application of machine learning models. [10]

Целью этого исследования была оценка применимости моделей машинного обучения для прогнозирования неонатальной смертности на основе материнских и клинических характеристик беременных женщин в отделении интенсивной терапии. В обсервационном перекрестно-групповом исследовании были собраны подробные данные, охватывающие диагноз матери, характеристики матери, акушерские характеристики и исходы новорожденного. Оценки бета-коэффициентов в наборе обучающих данных послужили основой для применения

модели к набору валидационных данных. Для построения модели был использован метод градиентного бустинга, модель продемонстрировала чувствительность и специфичность в 98%. Недостаток исследования заключается в ограниченности разнообразия примененных алгоритмов машинного обучения.

11 Early identification of bloodstream infection in hemodialysis patients by machine learning. [11]

Целью данного исследования было создание диагностической и прогностической модели для раннего выявления инфекции кровотока у пациентов, находящихся на гемодиализе. Для прогнозирования риска развития инфекции на основе клинических данных пациента использовались такие алгоритмы машинного обучения, как пошаговая логистическая регрессия (SLR), логистическая регрессия Лассо (LLR), метод опорных векторов (SVM), дерево решений, случайный лес (RF) и метод повышения градиента (XGBoost). Для оценки производительности таких моделей использовались точность (ACC) и площадь под рабочей кривой объекта. Значения аддитивного объяснения Шепли (SHAP) использовались для объяснения прогностического значения каждого признака на выходных данных моделей. В результате исследования было проведено сравнение эффективности 6-ти моделей машинного обучения в прогнозировании инфекции кровотока у пациентов, наибольшую точность показал метод повышения градиента. Для подтверждения эффективности этой модели требуются более масштабные многоцентровые исследования.

12 Chronic kidney disease prediction based on machine learning algorithms. [12]

В данном исследовании для прогнозирования хронической болезни почек и постановки диагноза на ранних стадиях болезни были использованы такие алгоритмы машинного обучения, как метод опорных векторов (SVM), метод k-ближайших соседей (KNN), деревья принятия решений с градиентным бустингом, гибридные ал-

горитмы и другие. При предварительной проверке данных для восполнения недостающих значений был использован метод среднего. Были проведены исследования связей между различными факторами, чтобы можно было сократить количество признаков и исключить избыточную информацию. Для проверки соответствия результатов использовалась 10-кратная перекрестная проверка.

13 Predicting the onset of diabetes-related complications after a diabetes diagnosis with machine learning algorithms. [13]

В данном ретроспективном исследовании для прогнозирования осложнений, связанных с диабетом, через один, два и три года после постановки диагноза были использованы данные о 610000 пациентах за 2013-2017 года. Для прогнозирования и оценки эффективности были использованы такие алгоритмы, как логистическая регрессия, дерево принятия решений, случайный лес и повышение экстремального градиента. Алгоритмы показали сравнительно высокую эффективность, значения площади под кривой рабочей характеристики (точности) варьировались от 60% до 69% в зависимости от рассматриваемого осложнения. Основные ограничения данного исследования связаны с доступностью данных и отсутствием данных о более длительных периодах наблюдения для проверки сохранения закономерности в долгосрочной перспективе.

Итог

В медицинской диагностике на данный момент используется множество методов машинного обучения. Наиболее активно ML используется для диагностики заболеваний и в персонализированной медицине, для решения этих задач применяют обучение с учителем или с частичным привлечением учителя. Модели машинного обучения могут использоваться для анализа клинических данных пациентов и прогнозирования риска возникновения определенных заболеваний. Например, модели могут предсказывать вероятность развития диабета, сердечно-сосудистых заболеваний, онкологических заболеваний и других медицинских состояний на основе данных о пациентах. Однако некоторые исследования, описанные в

данных статьях, основывались на недостаточном объеме данных, из-за чего полученные результаты не всегда можно было считать точными. Машинное обучение имеет потенциал существенно повысить точность диагностики, персонализировать подход к лечению и улучшить результаты здравоохранения, что делает его важным инструментом в медицинской практике.

Список литературы

- [1] K. Schultebrasucks, M. Sijbrandij, I. Galatzer-Levy, J. Mouthaan, M. Olff, and M. van Zuiden, “Forecasting individual risk for long-term posttraumatic stress disorder in emergency medical settings using biomedical data: A machine learning multicenter cohort study,” <https://www.sciencedirect.com/science/article/pii/S2352289521000059>, 2021.
- [2] M. Ghorbian and S. Ghorbian, “Usefulness of machine learning and deep learning approaches in screening and early detection of breast cancer,” <https://www.sciencedirect.com/science/article/pii/S2405844023096354>, 2023.
- [3] O. Shahid, S. P. Mohammad Nasajpour and, R. M. Parizi, M. Han, M. Valero, F. Li, M. Aledhari, and Q. Z. Sheng, “Machine learning research towards combating covid-19: Virus detection, spread prevention, and medical assistance,” <https://www.sciencedirect.com/science/article/pii/S1532046421000800>, 2021.
- [4] S. Sassanarakkit, S. Hadpech, and V. Thongboonkerd, “Theranostic roles of machine learning in clinical management of kidney stone disease,” <https://www.sciencedirect.com/science/article/pii/S200103702200561X>, 2022.
- [5] G. Kantayeva, J. Lima, and A. I. Pereira, “Application of machine learning in dementia diagnosis: A systematic literature review,” <https://www.sciencedirect.com/science/article/pii/S2405844023088345>, 2023.
- [6] B. Milićević, M. Milošević, V. Simić, A. Preveden, L. Velicki, Đorđe Jakovljević, Z. Bosnić, M. Pičulin, B. Žunković, M. Kojić, and N. Filipović, “Machine learning and physical based modeling

- for cardiac hypertrophy,” <https://www.sciencedirect.com/science/article/pii/S2405844023039312>, 2023.
- [7] A. Rayan and S. Alanazi, “A novel approach to forecasting the mental well-being using machine learning,” <https://www.sciencedirect.com/science/article/pii/S1110016823009717#s0015>, 2023.
- [8] A. S. Lundervold and A. Lundervold, “An overview of deep learning in medical imaging focusing on mri,” <https://www.sciencedirect.com/science/article/pii/S0939388918301181>, 2019.
- [9] J. Pei, G. Wang, Y. Li, L. Li, C. Li, Y. Wu, J. Liu, and G. Tian, “Utility of four machine learning approaches for identifying ulcerative colitis and crohn’s disease,” <https://www.sciencedirect.com/science/article/pii/S2405844023106475>, 2023.
- [10] M. Espinola-Sánchez, S. Sanca-Valeriano, A. Campaña-Acuña, and J. Caballero-Alvarado, “Prediction of neonatal death in pregnant women in an intensive care unit: Application of machine learning models,” <https://www.sciencedirect.com/science/article/pii/S240584402307901X>, 2023.
- [11] T. Zhou, Z. Ren, Y. Ma, L. He, J. Liu, J. Tang, and H. Zhang, “Early identification of bloodstream infection in hemodialysis patients by machine learning,” <https://www.sciencedirect.com/science/article/pii/S2405844023054713>, 2023.
- [12] M. A. Islam, M. Z. H. Majumder, and M. A. Hussein, “Chronic kidney disease prediction based on machine learning algorithms,” <https://www.sciencedirect.com/science/article/pii/S2153353923000032>, 2023.
- [13] T. Mora, D. Roche, and B. Rodríguez-Sánchez, “Predicting the onset of diabetes-related complications after a diabetes diagnosis with machine learning algorithms,” <https://www.sciencedirect.com/science/article/pii/S0168822723006733>, 2023.