

АНАЛИЗ ПУБЛИКАЦИОННОЙ АКТИВНОСТИ ПО ПРОБЛЕМЕ ВНЕДРЕНИЯ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ПРИ ПРОВЕДЕНИИ КЛИНИЧЕСКИХ ИССЛЕДОВАНИЙ

ANALYSIS OF PUBLICATION ACTIVITY ON THE PROBLEM OF IMPLEMENTING MACHINE LEARNING METHODS IN CLINICAL RESEARCH

A. Alekseev

Summary. The article presents an analysis of the research community's activity in using machine learning in clinical research. The main keywords were identified, and their clustering and analysis of the relationship between key terms were conducted. Using the capabilities of information systems such as Lens.org and VOSviewer, an extensive analysis of scientific publications was performed. The research methodology is based on analytical methods, including analysis and synthesis. A systems approach was applied for the structural analysis of researcher activity in the field of machine learning applications in clinical research. The problem of applying machine learning in clinical research remains generally understudied, which hinders its widespread adoption due to a limited understanding of the functional capabilities and potential of this technology in the field. The results of this study can be useful for scientists and practicing physicians interested in using machine learning to improve the quality and efficiency of clinical research.

Keywords: heatmaps, clinical research, information technology, deep learning.

Алексеев Антон Викторович

Аспирант, Московский государственный университет
технологий и управления имени К.Г. Разумовского,
г. Москва
alxvanton@gmail.com

Аннотация. В статье проведён анализ активности мирового сообщества исследователей по использованию машинного обучения в клинических исследованиях. Были определены основные ключевые слова, проведена их кластеризация и анализ взаимосвязи основных терминов. Используя инструменты информационных систем, такие как Lens.org и VOSviewer, был выполнен обширный анализ научных публикаций. Методология исследования основана на аналитических методах, включая методы анализа и синтеза. Системный подход был применен для структурного анализа активности исследователей в области использования машинного обучения в клинических исследованиях. Проблема применения машинного обучения в клинических исследованиях в целом остаётся недостаточно исследованной, что мешает её широкому применению из-за ограниченного понимания функциональных возможностей и потенциала этой технологии в сфере клинических исследований. Результаты данного исследования могут быть полезны для учёных и практикующих врачей, заинтересованных в использовании машинного обучения для улучшения качества и эффективности клинических исследований.

Ключевые слова: тепловые карты, клинические исследования, информационные технологии, глубокое обучение.

Введение

В целом, изучение научных публикаций выявило два основных тренда — рост числа работ по применению нейросетей в клинике и изменение способов их использования. Актуальность текущего исследования заключается в том, что оно основано на статьях, ключевые слова в которых, посвящены использованию автоматизированной диагностики и автоматизации прогнозов относительно эффективности лечения с использованием алгоритмов машинного обучения. Практическая значимость данных направлений очевидна, поскольку они становятся все более широко используемыми в клинической практике, а это подтверждает и рост числа публикаций на эти темы. Особые надежды исследователи связывают с машинным обучением, так как его алгоритмы могут существенно повысить точность медицинских прогнозов. Для проведения анализа индекса научной активности исследователей в этой области важно вы-

брать ключевые слова для получения системного представления о существующей литературе.

Одна из задач современной клинической практики — это построение помощи пациентам, основываясь на индивидуальном подборе средств и методов [1]. А именно методы машинного обучения способны предоставить эффективные решения в этой области [2].

Также считается, что экспертные системы предназначены для замены экспертов, когда решение проблем из-за их недостаточного количества, недостаточной эффективности в решении проблемы или в опасных условиях (вредных) условиях [3].

Методология

Для анализа публикационной активности были выбраны следующие ключевые термины: «машинное об-

учение в клинических исследованиях», «использование нейронных сетей в медицине» и «цифровые технологии в клинических науках». Проведенный анализ по указанным терминам позволяет оценить интерес научного сообщества к использованию машинного обучения в клинических исследованиях и уровень развития соответствующих научных направлений в мировом контексте.

Для сбора и анализа данных был использован функционал системы Lens.org [4], предоставляющей доступ к глобальным научным публикациям с индексацией цитирования. Этот веб-сервис дал возможность проанализировать статистику и динамику научных работ по выбранным ключевым словам и выражениям, а также определить периоды возрастания интереса к данной проблематике и текущий уровень активности исследований в этой области.

С целью наглядного представления доминирующих научных направлений и выявления паттернов сотрудничества в публикационной деятельности был задействован аналитический инструмент VOSviewer[5]. Результаты анализа с помощью этого инструмента позволили создать карты взаимосвязей между ключевыми терминами, авторами и публикациями, что дало возможность определить основные тенденции и направления исследований в данной области. Созданные с использованием инструмента VOSviewer библиометрические карты обеспечивают визуализацию количества научных статей в зависимости от различных параметров, таких как страны происхождения, авторские имена, хронологические рамки и наличие связей между исследователями, публикациями и ключевыми словами [6].

Использование указанных методов позволило получить обоснованную картину библиометрического анализа научной активности профессионального сообщества по проблеме использования машинного обучения в клинических исследованиях и определить актуальные направления исследований в данной области.

Результаты

Данный обзор научных статей позволяет выявить две основные тенденции. Во-первых, к количеству работ по нейросетям в клинике за последний период времени можно отметить устойчивый рост. Во-вторых, характеризуется какими-то общими чертами развитие методов применения нейросетей, что можно выделить, ориентируясь на анализ современных публикаций. Одно из самых примечательных направлений — автоматизированное машинное обучение для постановки диагнозов и определения эффективности терапии, это направление становится все более интересным в плане применения в реальной клинике, что отражено на увеличивающемся количестве публикаций

Кроме того, отмечается рост исследований, связанных с применением нейросетевых технологий для анализа медицинских изображений, включая снимки МРТ, компьютерную томографию и рентгенограммы, что подтверждает их значительный потенциал в области клинической практики [7,8,9].

Для анализа публикационной активности по данной проблеме были выбраны ключевые слова, такие как «машинное обучение», «клинические исследования», «использование искусственного интеллекта в медицине». Эти термины были использованы для поиска и анализа научных публикаций в базе данных. После сбора данных проведен анализ, направленный на выявление кластеров тематически связанных ключевых слов и терминов.

Каждый кластер был сопоставлен с определенным аспектом использования машинного обучения в клинических исследованиях. Например, один из кластеров был связан с применением машинного обучения для диагностики и лечения заболеваний, в то время как другой кластер отражал использование алгоритмов машинного обучения для анализа медицинских изображений.

Для анализа временной динамики публикаций был использован метод распределения ключевых слов по годам исследования. Это позволило выявить изменения в интересах научного сообщества со временем и выявить актуальные направления исследований в области применения машинного обучения в клинической практике.

Дополнительно, для визуализации результатов использовалась программа VOSviewer, позволяющая создавать тепловые карты. Это позволило более наглядно представить результаты анализа и сделать выводы о развитии и актуальности исследуемой проблемы в научном сообществе (Рисунок 1).

Анализ данного рисунка позволяет выделить несколько кластеров, которые можно условно обозначить как «Covid-19» (серый), «Применение новейших технологий» (темно-серый), «Устройства ввода» (светло-серый) и «Методы» (черные кластеры). Интерпретация кластеров основана на встречающихся в них ключевых словах, однако, следует отметить, что данное разделение довольно условно, поскольку и кластеры, и термины взаимосвязаны.

Особенностью кластера «Применение новейших технологий» является его «технический» уклон в привязке к принятию решений в клинических исследованиях. В рамках него, в частности, освещаются темы построения модели исследования, определения эффективности лекарственного препарата, обеспечения безопасности и удобства использования технологий, особенно в контексте исследуемого препарата.

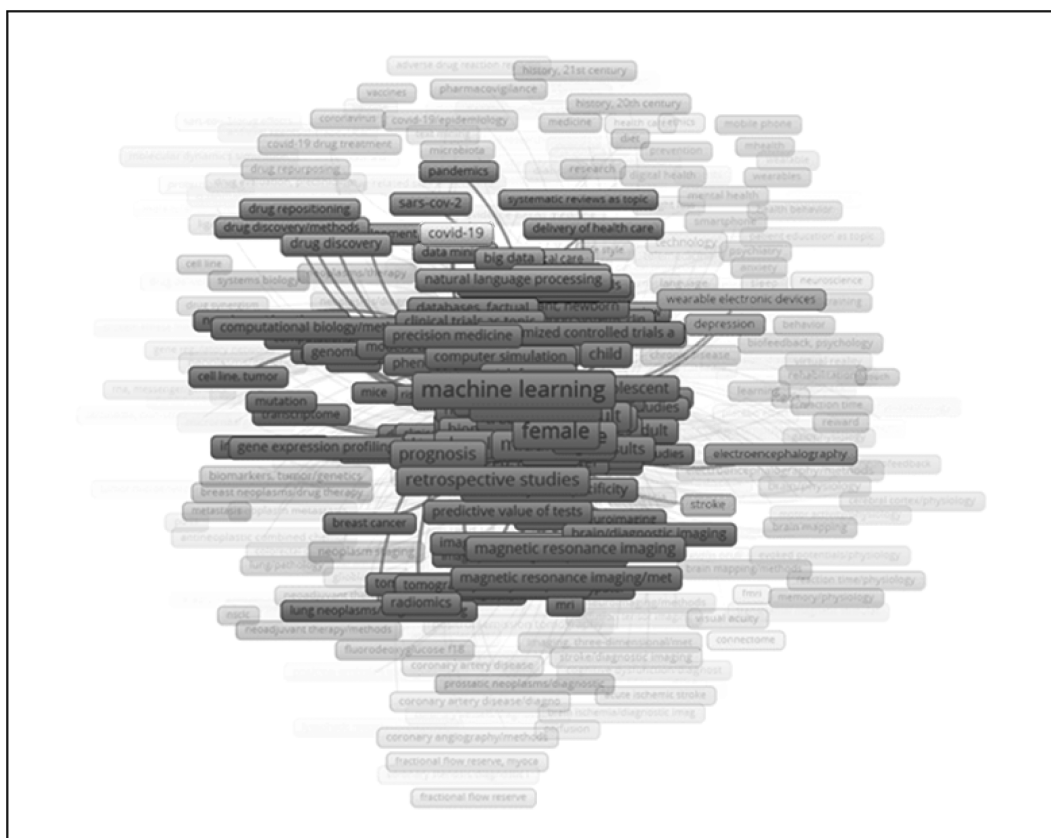


Рис. 1. Тематическая карта исследований, построенная по ключевым терминам, «Machine learning in Clinical Trials»

Причем, если добавить временную шкалу на данные кластеры, то можно увидеть, что такие термины, как «machine learning», «retrospective studies», относящиеся к терминам машинного обучения, используются в более актуальных статьях, нежели термины, относящиеся к области клинических исследований (Рисунок 2). Что говорит о том, что только недавно исследователи всего мира заинтересовались применении систем машинного обучения в разрезе клинических исследований.

Далее проведем анализ, по ключевым словам, «Artificial intelligence in Clinical trials with teacher». Чтобы убедиться, что данная тема действительно заинтересовала мировое сообщество. На представленном рисунке отображены ключевые слова с различной окраской, отражающей годы их публикации. Анализ данных позволяет сделать вывод о том, что в начальные периоды времени научных изысканий ключевые слова «drug discovery» и «computer simulation» преобладали в научной литературе. Однако, в последующие годы наблюдается значительное увеличение упоминаний таких терминов как «Artificial Intelligent» и «Machine Learning», что свидетельствует о повышенной актуальности темы исследования.

Это позволяет провести анализ динамики и изменений в интересах исследователей с течением времени. Например, ранние публикации могут содержать ключевые слова, связанные с основами исследования, в то

время как более поздние работы акцентируют внимание на современных методах и технологиях, таких как искусственный интеллект и машинное обучение. Это свидетельствует о постоянной эволюции в области клинических исследований и адаптации новейших технологий для решения актуальных задач.

Для определения актуальности исследования по проблеме применения машинного обучения в клинических исследованиях был проведен анализ научной литературы с использованием ключевых слов, фокусирующихся на применении систем машинного обучения при проведении клинических исследований. Одним из таких ключевых слов было сочетание терминов «Machine learning in Clinical Trials».

С использованием программного обеспечения VOSviewer и базы данных Lens.Org был произведен поиск научных публикаций, содержащих указанные ключевые слова. Результаты поиска охватили 10,368 публикаций, в которых встречаются данные термины. Для визуализации совпадений и взаимосвязей между терминами была построена визуальная карта, исключая из анализа наиболее распространенные слова.

Полученные результаты показали плотную структуру взаимосвязей, где крупный кластер, связанный с обучением, был тесно связан с кластером компьютерных

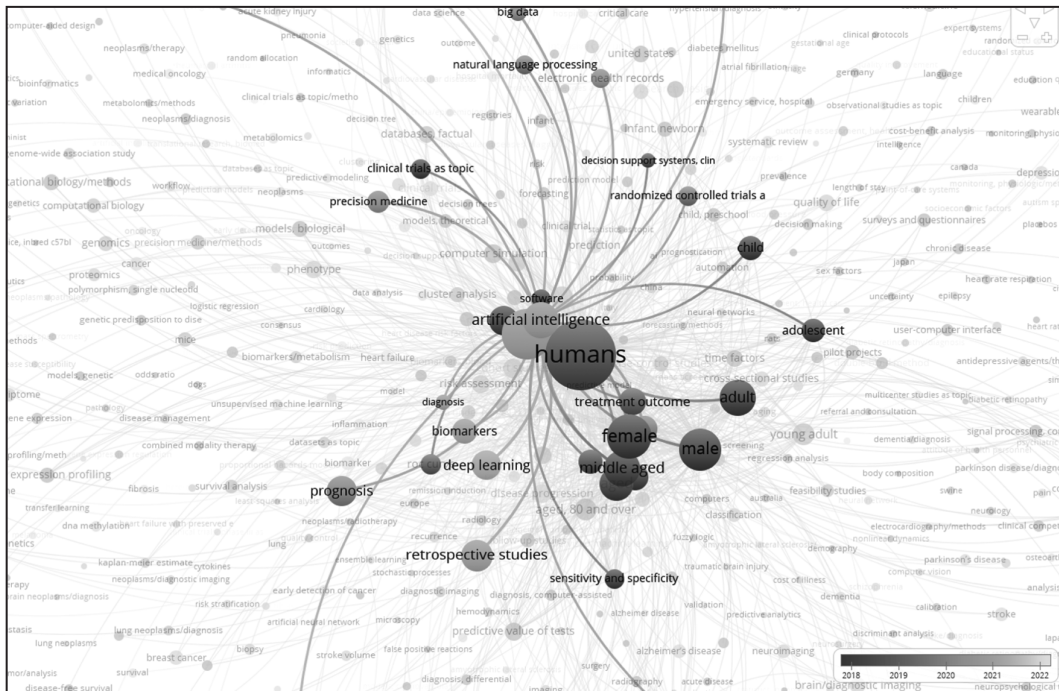


Рис. 4. Тематическая карта исследований, построенная по ключевым терминам, «Artificial intelligence in Clinical trials with teacher»

нейросетевых технологий. Также были выявлены связи с кластерами искусственного интеллекта и клиническими исследованиями, связанными с людьми (Рисунок 4).

Анализ указывает на активное развитие в области применения искусственного интеллекта и систем машинного обучения в клинических исследованиях, а также на интенсивный научно-исследовательский процесс в этом направлении.

Заключение

Проведенный анализ исследовательской деятельности по проблеме применения нейросетевых технологий при проведении клинических исследований позволяет сделать вывод о том, что на данный момент данная область является развивающейся и только набирает интерес ученых по всему миру. Несмотря на широкое развитие сферы машинного обучения применение их в сфере клинических исследований остается фрагментарным

и несистемным. Вследствие чего компьютерное моделирование является ключевым инструментом анализа и оптимизации работы систем здравоохранения. [10].

В современный период отмечается определенный рост количества исследований, связанных с использованием алгоритмов машинного обучения в сфере клинических исследований. Однако, количество научных работ по данной проблеме остается недостаточным для формирования системного подхода к изучению проблемы.

Наличие ограниченного количества публикаций по данной теме может свидетельствовать о неопределенности представлений научного сообщества относительно потенциала области применения искусственного интеллекта и систем машинного обучения в клинических исследованиях на широком уровне. Однако интерес явно прослеживается и заметен потенциал применения технологий машинного обучения при проведении клинических исследований.

ЛИТЕРАТУРА

- Белозерова Е.В., Данилов А.В., Исаенкова Е.А., Калинина Л.Б., Манерова О.А., Усов Ю.И. Прогнозирование развития гипертонической болезни с использованием моделей машинного обучения в подсистеме дистанционного кардиомониторинга // Менеджер здравоохранения. 2022. №2. URL: <https://cyberleninka.ru/article/n/prognozirovanie-razvitiya-gipertonicheskoy-bolezni-s-ispolzovaniem-modeley-mashinnogo-obucheniya-v-podsisleme-distantsionnogo> (дата обращения: 01.03.2024).
- Кротова Ольга Сергеевна, Москалев Игорь Викторович, Хворова Любовь Анатольевна, Назаркина Оксана Михайловна Реализация эффективных моделей классификации медицинских данных методами интеллектуального анализа текстовой информации // Известия АлтГУ. 2020. №1 (111). URL: <https://cyberleninka.ru/article/n/realizatsiya-effektivnyh-modeley-klassifikatsii-meditsinskih-dannyh-metodami-intellektualnogo-analiza-tekstovoy-informatsii> (дата обращения: 15.03.2024).

3. Drygin D.S., Pronkin N.N. Application of artificial intelligence in medicine // International journal of professional science. 2020. №1. URL: <https://cyberleninka.ru/article/n/application-of-artificial-intelligence-in-medicine> (дата обращения: 01.03.2024).
4. Lens [Электронный ресурс], URL: <https://www.lens.org/lens/> (дата обращения: 01.03.2024)
5. VOSviewer [Электронный ресурс], URL: <https://www.vosviewer.com/> (дата обращения: 01.03.2024)
6. Chvanova, M.S., Mitrofanova, I.P., Molchanov, A.A., Shlenov, Y.V., Podlesny, D.V. Terminological Maps of Publications on the Transformation of Students' Professional Orientations in the Context of their Internet Socialization Proceedings of the 2021 IEEE International Conference «Quality Management, Transport and Information Security, Information Technologies», IT and QM and IS 2021, 2021, pp. 605–611
7. Переверзев Владимир Сергеевич, Казьмин Аркадий Иванович, Сажнев Максим Леонидович, Пантелеев Андрей Андреевич, Колесов Сергей Васильевич. Искусственный интеллект для прогнозирования различных состояний в вертебрологии: систематический обзор // Гений ортопедии. 2021. №6. URL: <https://cyberleninka.ru/article/n/iskusstvennyy-intellekt-dlya-prognozirovaniya-razlichnyh-sostoyaniy-v-vertebrologii-sistematicheskiy-obzor> (дата обращения: 24.03.2024).
8. Абдуганиева Ш.Х., Никонорова М.Л. Цифровые решения в медицине // Крымский журнал экспериментальной и клинической медицины. 2022. №2. URL: <https://cyberleninka.ru/article/n/tsifrovye-resheniya-v-medicine> (дата обращения: 24.03.2024).
9. Абдуалимов Тимур Пулатович, Обрезан Андрей Григорьевич. Возможности искусственного интеллекта в прогнозировании поражения коронарных артерий // Кардиология: Новости. Мнения. Обучение. 2022. №1 (28). URL: <https://cyberleninka.ru/article/n/vozmozhnosti-iskusstvennogo-intellekta-v-prognozirovanii-porazheniya-koronarnyh-arteriy> (дата обращения: 24.03.2024).
10. Прокофьева Елизавета Сергеевна, Зайцев Роман Дмитриевич. Анализ клинических путей пациентов в медицинских учреждениях на основе методов жесткой и нечеткой кластеризации // Бизнес-информатика. 2020. №1. URL: <https://cyberleninka.ru/article/n/analiz-klinicheskikh-putey-patsientov-v-meditsinskih-uchrezhdeniyah-na-osnove-metodov-zhestkoy-i-nechetkoy-klasterizatsii> (дата обращения: 19.03.2024).

© Алексеев Антон Викторович (alxvanton@gmail.com)

Журнал «Современная наука: актуальные проблемы теории и практики»