

УДК 004.852, 001.891.5

*К. В. Воронцов¹, В. И. Игловиков², В. В. Стрижов¹, А. Е. Устюжанин^{1,3},
А. С. Хританков¹*

¹Московский физико-технический институт (национальный исследовательский университет)
²Lyft Inc

³Высшая школа экономики (национальный исследовательский университет)

Проблемы проведения экспериментов и воспроизводимости исследований в науках о данных

Приводится тезисное изложение результатов круглого стола по проблемам проведения экспериментов и воспроизводимости исследований в науках о данных. Указывается на различие форматов результатов научных и прикладных исследований, поясняется, в чем суть воспроизводимости результатов. Рассматриваются критерии качества результатов обоих видов исследований, в чем суть воспроизводимости результатов в каждом случае. Определяются направления дальнейшего развития методологии разработки предсказательных моделей, разработки алгоритмов и проведения экспериментов. Сформулированные рекомендации полезны для развития программ учебных курсов по машинному обучению.

Ключевые слова: воспроизводимость исследований, повторяемость экспериментов, науки о данных, машинное обучение, круглый стол.

K. V. Vorontsov¹, V. I. Iglovikov², V. V. Strijov¹, A. E. Ustuzhanin³, A. S. Khritanikov¹

¹Moscow Institute of Physics and Technology

^{1,2}Lyft Inc

^{1,3}Higher School of Economics

Roundtable: Challenges in repeatable experiments and reproducible research in data science

This article provides a summary of the results of a roundtable discussion on experimental design and research reproducibility in data science. A distinction is made between scientific and applied research, the issues of determining the quality for both types of research are considered, which is the essence of the reproducibility of the results in each case. In addition, an attempt is made to determine the directions for further development of infrastructure and methodology for the development of predictive models, algorithms and experiments. The recommendations formulated can be useful for the development of machine learning courses curricula.

Key words: research reproducibility, repeatability of experiments, data science, machine learning, roundtable.

1. Введение

Технологии машинного обучения являются результатом теоретических исследований и прикладных разработок в академической среде и в индустрии. Анализ представленных на

© Воронцов К. В., Игловиков В. И., Стрижов В. В., Устюжанин А. Е., Хританков А. С., 2021

Список авторов приведен в алфавитном порядке

© Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико-технический институт (национальный исследовательский университет)», 2021

конференциях AAAI, IJCAI, NeurIPS, ICLR работ показывает [1, 2], что доля экспериментальных исследований существенно выше чисто теоретических. Несмотря на имеющийся прогресс в области постановки и проведения экспериментов, воспроизводимость и повторяемость результатов пока еще не достигла ожидаемого уровня [3].

Более 70% из опрошенных исследователей не удалось хотя бы раз воспроизвести результаты, полученные ранее. В цитируемой 1140 раз (по данным Scopus на 16.04.2021) статье в журнале Nature [4] изучается отношение к воспроизводимости исследований и приводятся результаты опроса более полутора тысяч подписчиков этого журнала: ученых из области химии, физики, наук о земле, социологии, медицины и других. При этом 52% опрошенных указывают на наличие кризиса воспроизводимости в науке, а 31% считают, что сложности в воспроизведении исследований могут означать их недостоверность.

В более близком к машинному обучению мета-исследовании (исследовании, объектом изучения которого являются процесс или результаты научных исследований) [1] авторы указывают на факторы и переменные, влияющие на повторяемость экспериментов, и приводят результаты анализа 385 статей с эмпирическими исследованиями, представленных на конференциях AAAI (2014, 2016) и IJCAI (2013, 2016). Примерно в половине просмотренных авторами работ предствлен псевдокод алгоритма, описание проблемы, использованных для обучения данных, и более двух третей работ приводят постановку эксперимента. При этом вопросы исследования, цели исследования, проверяемые гипотезы, исходный код и анализ результатов чаще всего не включаются в опубликованные работы.

Организаторы конференций по машинному обучению NeurIPS (2019) и ICLR (2018, 2019) [2, 5] изучили, какие средства издатели и научные сообщества могли бы использовать для повышения воспроизводимости исследований. Организаторы предложили своим авторам использовать контрольные списки для самопроверки подаваемых статей, возможность представить исходный код и инструкции по воспроизведению результатов вместе со статьей, пригласили к публикации статьи с повторением исследований [5]. При этом наибольшей поддержкой у авторов пользовалась именно публикация исходного кода на открытой платформе, например Github.

Проблема воспроизводимости в машинном обучении находится в поле зрения научного сообщества не первый год. Так, Google Scholar приводит статьи, в которых обсуждаются вопросы воспроизводимости вычислительных экспериментов для эвристических методов еще с 1992 года [6–9], когда авторы попробовали совместить программы расчета результатов и скрипты подготовки публикации на одном CD-ROM диске для распространения коллегам [7]. В 2015 году была создана рабочая группа ACM (Association of Computing Machinery) по воспроизводимости исследований. На данный момент группой разработана система маркировки публикаций по уровню воспроизводимости результатов [10]. Эту схему использует, например, сообщество ACM по поиску информации SIGIR [11] с 2019 года, в том числе, в рамках совместной инициативы CLEF/NTCIR/TREC (CENTRE) [12] по сравнительному анализу алгоритмов и программных систем.

В индустрии [13] проблема воспроизводимости актуальна с точки зрения своевременного получения и внедрения результатов прогнозируемого качества. Возможное решение — в объединении идей DevOps [14] (сочетание Software Development and Operations в рамках единого рабочего процесса) и научного метода постановки эксперимента. Концепция MLOps (Machine Learning and Operations, по аналогии с DevOps) предполагает выстраивание единого потока работ подготовки данных и выполнения эксперимента от постановки задачи до внедрения и измерения результата. Известные облачные платформы и стартапы [15–17] предлагают программные средства для организации процессов создания и внедрения моделей машинного обучения в рамках MLOps.

В данной статье представлены результаты и рекомендации круглого стола, проведенного в рамках 63-й Всероссийской научной конференции МФТИ. Участники обсудили проблемы воспроизводимости экспериментов и повышения качества научных публикаций, возможные пути решения проблемы. В следующем разделе обсудим, в чем состоит проблема

воспроизводимости исследований и ее значимость. В третьем разделе приведем возможные решения, а также инструменты проведения экспериментов. В четвертом разделе обсудим вопросы подготовки специалистов в науках о данных.

2. Воспроизводимость и качество исследований

В чем состоит проблема воспроизводимости, и в каких ситуациях она особенно актуальна? Проблемы воспроизводимости возникают при изучении научных статей, при повторном использовании результатов других команд, при переносе результатов на другие наборы данных, при повторном использовании разработанных алгоритмов и предсказательных моделей, особенно в случае невозможности обращения к их авторам. При этом суть воспроизводимости и понимания качества результатов зависит от цели исследования или эксперимента.

В литературе указывают на множество причин невозможности воспроизводимости результатов, например, неполное описание в научной статье, недоступность исходного кода эксперимента или данных [1]. Кроме того, наблюдается эффект влияния самого экспериментатора [18,19], он заключается в более тщательной проработке и настройке собственных алгоритмов по сравнению с референтными. Этот эффект можно преодолеть, если кто-то еще воспроизводит результаты эксперимента без участия автора, например, если проводить соревнования и независимое сравнение алгоритмов в обзорных работах [20], а также применять автоматизированные методы проведения экспериментов.

Существуют социологические причины не делиться промежуточными результатами даже в небольших командах, в том числе вследствие научной конкуренции. Вероятно поэтому, на практике требования и стремление к воспроизводимости результатов исходят не от самих авторов, а от спонсоров исследования и издателей. При этом методологические аспекты подготовки и проведения повторяемых экспериментов не всегда хорошо проработаны. Перечисленные факторы не способствуют воспроизводимости экспериментов.

При определении понятия воспроизводимости результатов, можно выделить воспроизводимость результатов в малом, или *слабую воспроизводимость* — возможность получения заявленных в результатов на тех же входных данных, и *сильную воспроизводимость* — повторение результатов и их повторное использование другими исследователями (в англоязычной литературе это называется *repeatability* и *reproducibility* соответственно [7]).

Для учебных и научных работ важна повторная используемость и применимость результатов и частей решения в новых условиях, а не фактический результат по показателям качества, таким как F1-мера или частота ошибок второго рода. Для промышленных работ важна воспроизводимость другими исследователями и повторяемость результатов при переносе предсказательных моделей с исследовательских на реальные данные. В отличие от исследовательских работ для удовлетворения ожидания и целей бизнеса важно продемонстрировать высокую точность и полезность результатов, скажем, для повышения дохода от продаж. Остальные требования реализуются через регуляторные ограничения и административно установленные регламенты работы, чтобы удовлетворить требованиям законодательства или поддержать исследовательский уровень организации.

3. Инструменты и процессы поддержки проведения экспериментов

Машинное обучение – это во многом экспериментальная дисциплина, предполагающая получение практически значимых результатов с использованием сложных программных средств группой исследователей [21–23]. При этом для обеспечения необходимого качества на практике рекомендуется организовать совместную командную работу с реальными данными и обмен результатами с другими командами, которые решают связанные задачи. Формирующаяся в настоящее время концепция MLOps объединяет программные средства, практики и непрерывные процессы проведения исследований, выпуска и наблюдения за

предсказательными моделями на базе единой вычислительной платформы и общих рабочих процессов.

Концепция MLOps применима и в научных исследованиях для многокритериальной оценки моделей и алгоритмов, и формирования переднего края исследований [24]. Понимание и соотнесение с этим передним краем помогает исследователям и прикладным специалистам воспользоваться подходящим инструментом для решения своей задачи. Такие программные средства, как MLflow, DVC, Kubeflow [15,17], MLdev [16], NextFlow [25], используются для проведения вычислительных экспериментов как в индустрии, так и при проведении научных исследований.

В то же время применение MLOps в индустрии помогает решить проблемы доступа к данным и внедрения предсказательных моделей за счет автоматизации и снижения накладных расходов. Облачные платформы и сервисы проведения экспериментов на реальных и обновляемых данных, измеряющие эффективность моделей и смещение данных, существенно снижают накладные расходы и повышают продуктивность работы. При использовании таких платформ работа аналитика данных становится в чем-то аналогична работе трейдера, добывающегося увеличения KPI бизнеса [26–28] в реальном режиме времени.

В отличие от естественных и фундаментальных наук результаты исследований в науках о данных во многом имеют непосредственное практическое применение. Вероятно, поэтому научных мегапроектов и коллабораций, как в экспериментальной физике, в машинном обучении пока нет. В то время как мегапроекты строятся вокруг больших установок, которые участникам сложно создать по отдельности, в машинном обучении есть масштабные проекты для ограниченного круга пользователей и небольшого числа проверяемых гипотез [29]. Проблема доступа к уникальным экспериментальным установкам не является первоочередной, при наличии финансирования даже небольшие группы могут проводить исследования на доступном оборудовании.

С другой стороны, решение прикладных задач с помощью технологий машинного обучения может требовать объединения ресурсов. Например, при решении проблем докинга, сворачивания белков [30], исследований мозга в коллективе необходима как экспертиза в предметной области, так и в области машинного обучения. Известны коллаборации и другого рода — соревнования, направленные на развитие методов и технологий в выбранной задаче, так называемого технологического барьера, например Up Great ПРО//Чтение [31] в машинном понимании текста или широко известный Netflix Prize [32] для рекомендательных систем.

4. Подготовка специалистов в науках о данных

Обучение специалистов и исследователей в области анализа данных — важная задача. В последнее время популярными стали краткосрочные курсы повышения квалификации, на которых будущие аналитики данных изучают программные средства и алгоритмы для повышения KPI бизнеса с использованием уже готовых методов. Обучение методикам выполнения экспериментов поможет в получении предсказуемых результатов в работе и создания ценности для бизнеса.

В то же время для высшего образования важным дополнением к текущим программам могут стать темы по проведению экспериментов и обеспечению качества прикладного анализа данных. Для поддержки процесса научной работы студентам нужно передавать ценности научного знания [33], понимания сути научного исследования. В чем значение воспроизводимости и повторяемости, совместных обсуждений и обмена идеями, донесения своих результатов до широкой аудитории, популяризации результатов для того, чтобы они были приняты научным сообществом? В этой сфере необходимы дополнительные усилия по систематизации, определению критериев качества исследований и развитию образовательных программ.

5. Заключение

В заключение представим основные выводы и рекомендации.

Первое, экспериментальная проверка и сравнение предсказательных моделей и алгоритмов – это важная часть наук о данных. В зависимости от целей исследования по-разному понимается и качество эксперимента. При этом сложности связаны не только с инструментами и методами, но и с влиянием самих исследователей и их отношением к получаемым результатам.

Второе, при оценке научных исследований и учебных работ следует рассматривать совокупность критериев, достижение которых обеспечивает общеприменимость и воспроизводимость результатов. Здесь воспроизводимость научного эксперимента состоит в повторяемости на других наборах данных другими исследователями, возможность повторного использования результатов. Более сильный вариант воспроизводимости, к которому нужно стремиться – это получение другими исследователями тех же выводов и следствий, что и в исходном эксперименте.

Третье, качество прикладного исследования в индустрии во многом определяется достижением заданных КПЭ (ключевые показатели эффективности), воспроизводимость в этом случае предполагает подтверждение и повторение результатов на меняющихся реальных данных на промышленной системе. При этом вопросам безопасности, надежности, производительности и другим нефункциональным характеристикам качества следует уделить особое внимание как ключевым для бизнеса. Методология и инструменты MLOps, являясь развитием подходов из программной инженерии в применении к анализу данных, будет способствовать решению этих вопросов. Несмотря на быстрое развитие в этой области [34], на данный момент все еще не сформулировано, какие именно продукты и решения востребованы в этой сфере, и в чем состоят ключевые функции платформ автоматизации проведения экспериментов.

Четвертое, образовательные программы для аналитиков данных и для исследователей в области машинного обучения имеют разные цели и должны отличаться. Для будущих исследователей необходимо предусмотреть изучение тем по методологии исследований, проведению экспериментов и оценке полученных результатов.

6. Информация о проведении круглого стола

Круглый стол был проведен 23 ноября 2020 г. вместе с 63-й Всероссийской конференцией МФТИ. В работе круглого стола приняли участие:

Воронцов К. В. — доктор физ.-мат. наук, заведующий лабораторией Машинного интеллекта МФТИ, заведующий кафедрой Интеллектуальных систем ФПМИ.

Игловиков В. И. — кандидат физ.-мат. наук, старший исследователь, Kaggle Grandmaster.

Стрижов В. В. — доктор физ.-мат. наук, проф. кафедры Интеллектуальных систем МФТИ. Автор курса «Моя первая научная статья».

Устюжанин А. Е. — кандидат физ.-мат. наук, заведующий лабораторией методов анализа больших данных НИУ ВШЭ.

Хританков А. С. — модератор круглого стола, кандидат физ.-мат. наук, доцент кафедры АТП МФТИ, архитектор систем машинного обучения.

Литература

1. *Gundersen O.E., Gil Y., Aha D.W.* On reproducible AI: Towards reproducible research, open science, and digital scholarship in AI publications // AI magazine. 2018. V. 39.3. P. 56–68.
2. *Pineau J., Vincent-Lamarre P., Sinha K., Lariviere V., Beygelzimer A., d'Alche-Buc F., ... & Larochelle H.* Improving reproducibility in machine learning research (a report from

- the neurips 2019 reproducibility program) // arXiv preprint. 2018. [Электронный ресурс] URL: <https://arxiv.org/abs/2003.12206> (дата обращения: 20.03.2021).
3. *Hutson M.* Artificial intelligence faces reproducibility crisis // *Science*. 2018. V. 359(6377). P. 725–726.
 4. *Baker, M., Penny, D.* Is there a reproducibility crisis? // *Nature*. 2016. V. 533(7604). P. 452–454.
 5. *Pineau J., Sinha K., Fried G., Ke R.N., Larochelle H.* ICLR Reproducibility Challenge 2019 // *ReScience C*. 2020. V. 5(2). P. 5.
 6. *Stodden V.* The legal framework for reproducible scientific research: Licensing and copyright // *Computing in Science & Engineering*. 2008. V. 11(1). P. 35–40.
 7. *Claerbout J.F., Karrenbach M.* Electronic documents give reproducible research a new meaning / *Society of Exploration Geophysicists // SEG Technical Program Expanded Abstracts*. 1992. P. 601–604.
 8. *LeVeque R.J., Mitchell I.M., Stodden V.* Reproducible research for scientific computing: Tools and strategies for changing the culture // *Computing in Science & Engineering*. 2012. V. 14(4). P. 13–17.
 9. *Barr R.S., Golden B.L., Kelly J.P., Resende M.G., Stewart W.R.* Designing and reporting on computational experiments with heuristic methods // *Journal of heuristics*. 1995. V. 1(1). P. 9–32.
 10. ACM Artifact Review and Badging [Электронный ресурс], URL: <https://www.acm.org/publications/policies/artifact-review-and-badging-current> (дата обращения: 20.03.2021).
 11. *Ferro N., Kelly D.* SIGIR initiative to implement ACM artifact review and badging // *ACM SIGIR Forum*. New York, NY, USA : ACM. V. 52, N 1. P. 4–10.
 12. CLEF/NTCIR/TREC REproducibility (CENTRE) [Электронный ресурс], URL: <https://www.centre-eval.org/> (дата обращения: 20.03.2021).
 13. Tech Trends 2021 // *Deloitte Insights*. 2020 [Электронный ресурс], URL: <https://www2.deloitte.com/us/en/insights/focus/tech-trends.html> (дата обращения: 20.03.2021).
 14. *Ким Г., Бер К., Снафффорд К.* Проект «Феникс»: Роман о том, как DevOps меняет бизнес к лучшему. Москва : Эксмо, 2014. 384 с.
 15. A curated list of awesome MLOps tools [Электронный ресурс], URL: <https://github.com/kelvins/awesome-mlops> (дата обращения: 20.03.2021).
 16. MLdev experiment automation software [Электронный ресурс], URL: <https://gitlab.com/mlrep/mldev> (дата обращения: 20.03.2021).
 17. Machine Learning Operations [Электронный ресурс], URL: <https://github.com/pyseany/awesome-machine-learning-operations> (дата обращения: 20.03.2021).
 18. *Kerr N.L.* HARKing: Hypothesizing after the results are known // *Personality and social psychology review*. 1998. V. 2(3). P. 196–217.
 19. *Hollenbeck J.R., Wright P.M.* Harking, sharking, and tharking: Making the case for post hoc analysis of scientific data // *Journal of Management*. 2016. V. 4, I. 1. P. 5–18.
 20. How to win a kaggle competition [Электронный ресурс], URL: <https://www.import.io/post/how-to-win-a-kaggle-competition/> (дата обращения: 20.03.2021).

21. *Wirth R., Hipp J.* CRISP-DM: Towards a standard process model for data mining // Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. London, UK : Springer-Verlag. 2000. Т. 1.
22. *Huber S. [et al.]*. DMME: Data mining methodology for engineering applications—a holistic extension to the CRISP-DM model // *Procedia Cirp*. 2019. V. 79. P. 403–408.
23. *Shafique U., Qaiser H.* A comparative study of data mining process models (KDD, CRISP-DM and SEMMA) // *International Journal of Innovation and Scientific Research*. 2014. V. 12., N 1. P. 217–222.
24. Papers with Code [Электронный ресурс], URL: <https://paperswithcode.com> (дата обращения: 20.03.2021).
25. *Di Tommaso P., Chatzou M., Floden E.W., Barja P.P., Palumbo E., Notredame C.* Nextflow enables reproducible computational workflow // *Nature biotechnology*. 2017. V. 35(4). P. 316–319.
26. Система YТ [Электронный ресурс], URL: <https://habr.com/ru/company/yandex/blog/311104/> (дата обращения: 20.03.2021).
27. Sberbank Huge Data Platform [Электронный ресурс], URL: [https://www.tadviser.ru/index.php/Проект:Сбербанк_\(новая_технологическая_платформа\)](https://www.tadviser.ru/index.php/Проект:Сбербанк_(новая_технологическая_платформа)) [https://www.tadviser.ru/index.php/Проект:Сбербанк_\(новая_технологическая_платформа\)](https://www.tadviser.ru/index.php/Проект:Сбербанк_(новая_технологическая_платформа)) (дата обращения: 20.03.2021).
28. Sber.DS [Электронный ресурс], URL: <https://habr.com/ru/company/sberbank/blog/489158/> (дата обращения: 20.03.2021).
29. *Brown T.B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., ... & Amodei D.* Language models are few-shot learners // *arXiv preprint*. 2020. [Электронный ресурс] URL: <https://arxiv.org/abs/2005.14165> (дата обращения: 20.03.2021).
30. *Senior A.W., Evans R., Jumper J. [et al.]*. Improved protein structure prediction using potentials from deep learnin // *Nature*. 2020. V. 577. P. 706–710. URL: <https://doi.org/10.1038/s41586-019-1923-7>
31. Технологический конкурс ПРО//ЧТЕНИЕ [Электронный ресурс], URL: <https://ai.upgreat.one> (дата обращения: 20.03.2021).
32. Netflix Prize [Электронный ресурс], URL: https://en.wikipedia.org/wiki/Netflix_Prize (дата обращения: 20.03.2021).
33. Учебный курс «Моя первая научная статья» [Электронный ресурс], URL: <https://m1p.org> (дата обращения: 20.03.2021).
34. *Makinen S., Skogstrom H., Laaksonen E., Mikkonen T.* Who Needs MLOps: What Data Scientists Seek to Accomplish and How Can MLOps Help? // *IEEE/ACM 1st Workshop on AI Engineering – Software Engineering for AI (WAIN) of 43rd International Conference on Software Engineering (ICSE)*. 2021.

References

1. *Gundersen O.E., Gil Y., Aha D.W.* On reproducible AI: Towards reproducible research, open science, and digital scholarship in AI publications. *AI magazine*. 2018. V. 39.3. P. 56–68.
2. *Pineau J., Vincent-Lamarre P., Sinha K., Lariviere V., Beygelzimer A., d’Alche-Buc F., ... & Larochelle H.* Improving reproducibility in machine learning research (a report from the neurips 2019 reproducibility program). *arXiv preprint*. 2018. [Электронный ресурс] URL: <https://arxiv.org/abs/2003.12206> (дата обращения: 20.03.2021).

3. *Hutson M.* Artificial intelligence faces reproducibility crisis. *Science*. 2018. V. 359(6377). P. 725–726.
4. *Baker, M., Penny, D.* Is there a reproducibility crisis? *Nature*. 2016. V. 533(7604). P. 452–454.
5. *Pineau J., Sinha K., Fried G., Ke R.N., Larochelle H.* ICLR Reproducibility Challenge 2019. *ReScience C*. 2020. V. 5(2). P. 5.
6. *Stodden V.* The legal framework for reproducible scientific research: Licensing and copyright. *Computing in Science & Engineering*. 2008. V. 11(1). P. 35–40.
7. *Claerbout J.F., Karrenbach M.* Electronic documents give reproducible research a new meaning. *Society of Exploration Geophysicists. SEG Technical Program Expanded Abstracts*. 1992. P. 601–604.
8. *LeVeque R.J., Mitchell I.M., Stodden V.* Reproducible research for scientific computing: Tools and strategies for changing the culture. *Computing in Science & Engineering*. 2012. V. 14(4). P. 13–17.
9. *Barr R.S., Golden B.L., Kelly J.P., Resende M.G., Stewart W.R.* Designing and reporting on computational experiments with heuristic methods. *Journal of heuristics*. 1995. V. 1(1). P. 9–32.
10. ACM Artifact Review and Badging [Electronic resource], URL: <https://www.acm.org/publications/policies/artifact-review-and-badging-current> (date accessed: 20.03.2021).
11. *Ferro N., Kelly D.* SIGIR initiative to implement ACM artifact review and badging. *ACM SIGIR Forum*. New York, NY, USA : ACM. V. 52, N 1. P. 4–10.
12. CLEF/NTCIR/TREC REproducibility (CENTRE) [Electronic resource], URL: <https://www.centre-eval.org/> (date accessed: 20.03.2021).
13. Tech Trends 2021. Deloitte Insights. 2020 [Electronic resource], URL: <https://www2.deloitte.com/us/en/insights/focus/tech-trends.html> (date accessed: 20.03.2021).
14. *Kim G., Behr K., Spafford K.* The phoenix project: A novel about IT, DevOps, and helping your business win. *IT Revolution*. 2014.
15. A curated list of awesome MLOps tools [Electronic resource], URL: <https://github.com/kelvins/awesome-mlops> (date accessed: 20.03.2021)
16. MLdev experiment automation software [Electronic resource], URL: <https://gitlab.com/mlrep/mldev> (date accessed: 20.03.2021).
17. Machine Learning Operations [Electronic resource], URL: <https://github.com/pyseany/awesome-machine-learning-operations> (date accessed: 20.03.2021).
18. *Kerr N.L.* HARKing: Hypothesizing after the results are known. *Personality and social psychology review*. 1998. V. 2(3). P. 196–217.
19. *Hollenbeck J.R., Wright P.M.* Harking, sharking, and tharking: Making the case for post hoc analysis of scientific data. *Journal of Management*. 2016. V. 4, I. 1. P. 5–18.
20. How to win a kaggle competition [Electronic resource], URL: <https://www.import.io/post/how-to-win-a-kaggle-competition/> (date accessed: 20.03.2021).
21. *Wirth R., Hipp J.* CRISP-DM: Towards a standard process model for data mining. *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*. London, UK : Springer-Verlag. 2000. T. 1.

22. *Huber S., et al.*, DMME: Data mining methodology for engineering applications—a holistic extension to the CRISP-DM model. *Procedia Cirp*. 2019. V. 79. P. 403–408.
23. *Shafique U., Qaiser H.* A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*. 2014. V. 12, N 1. P. 217–222.
24. Papers with Code [Electronic resource], URL: <https://paperswithcode.com> (date accessed: 20.03.2021).
25. *Di Tommaso P., Chatzou M., Floden E.W., Barja P.P., Palumbo E., Notredame C.* Nextflow enables reproducible computational workflow. *Nature biotechnology*. 2017. V. 35(4). P. 316–319.
26. Sistema YT [Electronic resource], URL: <https://habr.com/ru/company/yandex/blog/311104/> (date accessed: 20.03.2021).
27. Sberbank Huge Data Platform [Electronic resource], URL: [https://www.tadviser.ru/index.php/Проект:Сбербанк_\(новая_технологическая_платформа\)](https://www.tadviser.ru/index.php/Проект:Сбербанк_(новая_технологическая_платформа))[https://www.tadviser.ru/index.php/Проект:Сбербанк_\(новая_технологическая_платформа\)](https://www.tadviser.ru/index.php/Проект:Сбербанк_(новая_технологическая_платформа)) (date accessed: 20.03.2021).
28. Sber.DS [Electronic resource], URL: <https://habr.com/ru/company/sberbank/blog/489158/> (date accessed: 20.03.2021).
29. *Brown T.B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., ... & Amodei D.* Language models are few-shot learners // arXiv preprint. 2020. URL: <https://arxiv.org/abs/2005.14165>
30. *Senior A.W., Evans R., Jumper J., et al.*, Improved protein structure prediction using potentials from deep learnin. *Nature*. 2020. V. 577. P. 706–710. URL:<https://doi.org/10.1038/s41586-019-1923-7>
31. Tekhnologicheskij konkurs PRO//CHTENIE [Electronic resource], URL: <https://ai.upgreat.one> (date accessed: 20.03.2021).
32. Netflix Prize [Electronic resource], URL: https://en.wikipedia.org/wiki/Netflix_Prize (date accessed: 20.03.2021).
33. Course «My first scientific paper» [Electronic resource], URL: <https://m1p.org> (date accessed: 20.03.2021).
34. *Makinen S., Skogstrom H., Laaksonen E., Mikkonen T.* Who Needs MLOps: What Data Scientists Seek to Accomplish and How Can MLOps Help? IEEE/ACM 1st Workshop on AI Engineering - Software Engineering for AI (WAIN) of 43rd International Conference on Software Engineering (ICSE). 2021.

Поступила в редакцию 02.05.2021