

ПОТЕНЦИАЛ ПРИМЕНЕНИЯ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ЭКСПЕРТИЗЫ НАУЧНЫХ ПРОЕКТОВ

С.П. Юркевичюс, нач. отд. ФГБНУ НИИ РИНКЦЭ, канд. техн. наук, доц.,
jursp@extech.ru

Д.А. Вразнов, науч. сотр. Института оптики атмосферы им. В.Е. Зуева СО РАН,
vda@iao.ru

Рецензент: Е.В. Ляпунцова, ФГАОУ ВО «Московский государственный технический университет имени Н.Э. Баумана», д-р техн. наук, *lev7lev63@me.com*

Данный обзор посвящен перспективам внедрения методов искусственного интеллекта (ИИ) в области автоматизации процедур оценки качества научно-исследовательских проектов (НИР). Приводится декомпозиция стандартной системы оценки НИР, оценивается потенциал внедрения методов ИИ в каждом сегменте системы. Также рассматривается автоматизация оценки проектов согласно шкале уровня готовности технологии. К наиболее перспективным областям был отнесен этап экспертизы заявки за счет внедрения современных больших языковых моделей, способных облегчить работу экспертов по анализу текста на новизну, актуальность, плагиат.

Ключевые слова: искусственный интеллект, машинное обучение, системы поддержки принятия решения, экспертиза научных проектов, системы оценки проектов.

THE POTENTIAL OF USING ARTIFICIAL INTELLIGENCE METHODS FOR THE EXAMINATION OF SCIENTIFIC PROJECTS

S.P. Yurkevichyus, Head of Department, SRI FRCEC, Doctor of Engineering, Assistant Professor, *jursp@extech.ru*

D.A. Vrazhnov, Researcher, V.E. Zuev Institute of Atmospheric Optics RAS, Siberian Branch, *vda@iao.ru*

This review is devoted to the prospects for the introduction of artificial intelligence (AI) methods, in the field of automation of procedures for quality assessment of research projects (R&D). A decomposition of the standard system of R&D evaluation is given and the potential of AI methods implementation in each segment of the system is evaluated. The automation of project evaluation according to the technology readiness level scale is also considered. The most promising areas include the stage of the application examination, due to the introduction of modern large language models that can facilitate the work of experts in analyzing the text for novelty, relevance, and plagiarism.

Keywords: artificial intelligence, machine learning, decision support systems, expertise of scientific projects, project evaluation systems.

Введение

В последнее время отмечается взрывная динамика развития как теоретических основ искусственного интеллекта (ИИ), так и областей его применения. Наиболее полно термин ИИ описывает совокупность методов и технологий в области компьютерных наук, позволяющих вычислительным машинам имитировать такие возможности человека, как рассуждать, принимать решения или решать проблемы [1, 2]. К современным областям применения ИИ относят компьютерное зрение и робототехнику [3], системы поддержки принятия решения

(СППР) [4], в том числе ИИ – агенты-консультанты (ChatGPT [5], DeepSeek [6]), позволяющие агрегировать, систематизировать и интерпретировать огромные массивы информации.

Важной особенностью создания ИИ-систем является наличие размеченных экспертом данных либо эвристических правил, позволяющих реализовать процесс обучения модели на базе ИИ. Поэтому согласно одной из классификаций ИИ выделяет модели, основанные на правилах или знаниях (экспертные системы), и модели на основе данных (машинное обучение) (рис. 1).

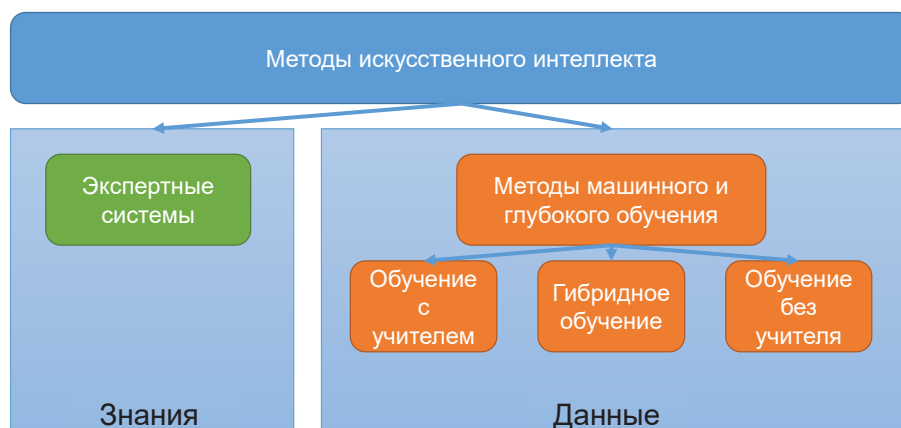


Рис. 1. Классификация методов ИИ

Экспертные системы широко применяются для решения формализованных задач с алгоритмизуемым решением. Примером таких систем являются СППР в медицине, основанные на протоколах и клинических рекомендациях [7]. Они не требуют значительных вычислительных мощностей, имеют небольшой набор оптимизируемых гиперпараметров.

Методы машинного обучения, включая глубокое обучение, позволяют строить модели данных со сложной структурой, обобщая известную информацию. Наиболее известными примерами являются большие языковые модели с миллиардами гиперпараметров и требующие использования дорогостоящих вычислительных кластеров для обучения моделей [8]. Следует отметить, что, несмотря на стремительный рост сфер применения ИИ, решение творческих задач с применением методов ИИ зачастую неудовлетворительно. Это связано с ограниченностью выборки и необходимостью извлечения семантики для оценки качества генерируемых идей [9, 10].

Оценка качества научных проектов – трудоемкая задача, которая сейчас решается с привлечением экспертов. Соответственно, внедрение методов ИИ в данную область является актуальным. В данной работе исследуются области вышеуказанной задачи, в которых целесообразно и возможно внедрение методов ИИ.

Возможные направления применения методов искусственного интеллекта для экспертизы научных проектов

В задаче научной экспертизы можно выделить два основных направления: экспертиза заявок научных проектов и экспертиза научных отчетов (рис. 2). Экспертиза заявок научных проектов в большей части требует проверки соответствия формальным требованиям, например количеству и качеству публикаций по теме проекта (часть научного задела), наличие опыта руководства проектами и достаточности материальной базы. В этой части оценки

научного проекта широко распространена практика экспертизы научного отчета на основе опросников/анкет с количественными оценками каждого критерия из списка. Итоговой оценкой отчета является взвешенная сумма баллов по критериям. Основной проблемой данного подхода является субъективность выбора соотношения баллов разделов опросника и, следовательно, невозможность существенно использовать экспертное мнение, основанное на опыте. Эту проблему гипотетически можно решить путем обучения алгоритма ИИ на выборке научных проектов, предварительно «размеченных» группой экспертов на набор классов, соответствующих «качеству» проекта, что позволит учесть более тонкие нюансы (признаки качества проекта) и повысить качество экспертизы новых проектов с применением методов ИИ.

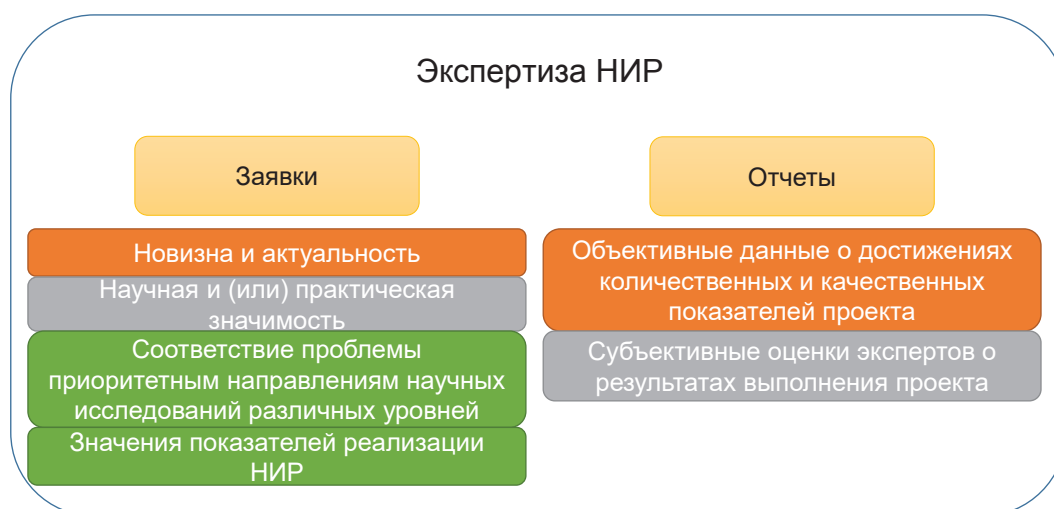


Рис. 2. Задачи научной экспертизы НИР

Оранжевым цветом выделены потенциальные области применения ИИ. Серым цветом — недоступные в настоящее время для решения методами ИИ. Зеленым цветом — имеющие готовое автоматизированное решение

К неформальным характеристикам проекта относятся новизна, актуальность, соответствие методов и подходов решаемым задачам и т.п. Одной из ключевых проблем в этой области является то, что потенциальный ИИ-агент на базе большой языковой модели должен обладать способностью осуществлять семантический анализ текста, то есть выделить смысл того, о чем авторы хотели сказать. ИИ-агенты обучаются на известной информации, и то, как они будут ее экстраполировать на неизвестную, остается открытым вопросом. Решение этой проблемы возможно с использованием интерпретируемых методов машинного обучения или проведения дополнительных исследований по надежности и адекватности заключений, полученных с использованием методов ИИ, работающих по принципу «черного ящика», к которым, в частности, относится большинство искусственных нейросетей. Для развития данной области в работе [11] предлагается общий фреймворк для агентов ИИ, основанный на принципах дивергентного и конвергентного мышления. В целом на текущий момент ни одна из современных больших языковых моделей не может на уровне человека понимать написанное. Тем не менее уже сейчас возможно привлечение ИИ-агентов к задачам выявления плагиата, проверки данных статистического анализа и других манипуляций с данными [12].

Другой сложностью экспертизы научного проекта или научного отчета является оценка его новизны. В узком смысле новизна трактуется как непохожесть на существующие тексты. Классические методы оценки схожести основаны на следующих метриках: косинусная (представляет собой косинус угла между двумя векторами в многомерном пространстве), tf-idf (Term Frequency-Inverse Document Frequency, учитывает частоту встречаемости слова в текстовом датасете), дивергенция Кульбака — Лейблера (статистическая мера измерения расстояния между вероятностными распределениями), ряд метрик для определения близости графов, а также метрики, используемые в методах кластеризации и классификации. Для обучения и валидации разрабатываемых моделей анализа текстов широко используется база TREC [13]. Этот датасет был разработан в рамках трехлетнего проекта по разработке методов анализа текстов, в частности оценки новизны и релевантности. Наилучшие достигнутые результаты — значение меры F1 (среднегармоническое между точностью и полнотой) в диапазоне 0,60–0,75, значения прецизионности и полноты 65–75 % [14]. В больших языковых моделях используются механизмы генерации семантических пространств на основе рекуррентных нейронных сетей [15], в частности с архитектурами автокодировщика (например, Universal Sentence Encoder, USE) [16], долговременной и кратковременной памятью (Long Short-Term Memory, LSTM) [17]. Для создания подобных моделей используются большие датасеты, такие как Wikipedia или Google News. Информация об аналогичных отечественных базах в открытом доступе отсутствует.

В работе [18] предлагается семантическая модель оценки новизны и актуальности на основе анализа групповых чатов, сохраненных во время «мозговых штурмов». Предлагаемая модель состоит из блока оценки релевантности, проработанности, новизны. Блок релевантности проводит удаление нерелевантных идей методами опорных векторов, вероятностными языковыми моделями [19]. Блок проработанности применяется для исключения полностью сформулированных идей, используя информацию о длине и связности текста. Блок новизны оценивает релевантные и непроработанные идеи по количеству обсуждений на текущий момент и по частоте встречаемости идеи в контекстном домене. Полученные результаты свидетельствуют о возможности применения больших языковых моделей для оценки новизны и актуальности проектов.

При оценке новизны идеи важно соблюдать баланс между ней и полезностью. На практике исследователи столкнулись с проблемой генерации галлюцинаций и идей в жестких рамках (меморизация, запоминание). Проблема галлюцинации моделей заключается в создании контента, далекого от реальности. При меморизации модель не может выйти за рамки существующих ограничений и генерирует идеи с минимальными модификациями, не отвечающими критериям новизны.

Одним из текущих трендов развития тематик НИР в РФ стал акцент на прикладные технологические разработки. Для оценки зрелости технологических проектов применяется шкала уровней готовности технологии (УГТ) (англ. Technology Readiness Level, рис. 3) [20, 21].

Основными недостатками шкалы УГТ также являются ее субъективность и неоднозначность выбора заявленным уровнем. Этим обусловлена необходимость разработки инструментов независимой оценки уровня УГТ. Пилотные работы в данном направлении показали перспективность использования библиометрических индикаторов в качестве предикторов уровня готовности технологии [22, 23]. К последним авторы отнесли наличие венчурного капитала, рыночной оценки, лицензированных технологий, исследовательских проектов, клинических испытаний, научных публикаций, патентов, новостей и упоминаний в СМИ. Более того, учитывая комплексный характер показателей и нелинейный характер зависимостей между предикторами и целевой переменной, было предложено использовать модели машинного обучения для предсказания уровня УГТ (рис. 4).



Рис. 3. Пример шкалы УГТ для авиакосмической отрасли [20]

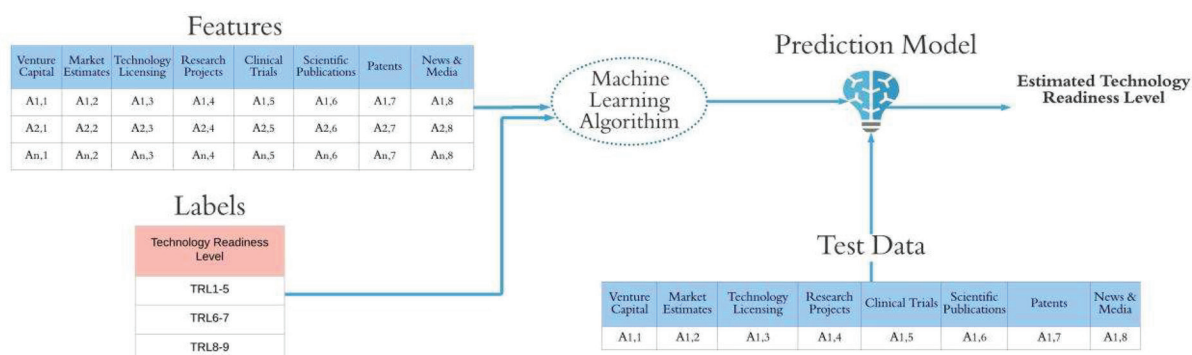


Рис. 4. Пример схемы применения модели машинного обучения для прогнозирования уровня технологической готовности [22]

Заметим, что вместо предсказания УГТ можно использовать методы машинного обучения для предсказания положения разрабатываемой технологии на кривой хайпа Гартнера.

Критерии соответствия проблемы приоритетным направлениям научных исследований, квалификация и опыт научного коллектива, использование ожидаемых результатов НИР на практике выражаются количественно, и их оценка не требует экспертизы человека-эксперта.

Заключение

В данной работе были рассмотрены основные этапы экспертизы научных проектов и научных отчетов. Проведенная декомпозиция на подзадачи и анализ литературы показали, что наиболее эффективной является автоматизация рутинных процедур на этапе экспертизы заявки. Это обусловлено большой долей субъективности оценок отчетного этапа и внедрения опросников с количественными показателями выполнения НИР, что не требует приме-

нения методов на базе ИИ. На этапе экспертизы заявок внедрение современных больших языковых моделей может существенно упростить работу экспертов в области анализа текста на новизну, актуальность, плагиат.

Статья подготовлена к изданию при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках Государственного задания на 2025 г. № 075-00682-25-02.

Список литературы

1. Helm J.M., et al. Machine learning and artificial intelligence: definitions, applications, and future directions // Current reviews in musculoskeletal medicine. 2020. Т. 13. С. 69–76.
2. Sheikh H., Prins C., Schrijvers E. Artificial intelligence: definition and background // Mission AI: The new system technology. Cham: Springer International Publishing, 2023. С. 15–41.
3. Murphy R.R. Introduction to AI robotics. MIT press, 2019.
4. Gupta S., et al. Artificial intelligence for decision support systems in the field of operations research: review and future scope of research // Annals of Operations Research. 2022. Т. 308. № 1. С. 215–274.
5. Kalla D., et al. Study and analysis of chat GPT and its impact on different fields of study // International journal of innovative science and research technology. 2023. Т. 8. №. 3.
6. Wu Z., et al. Deepseek-vl2: Mixture-of-experts vision-language models for advanced multimodal understanding // arXiv preprint arXiv:2412.10302. 2024.
7. Толмачев И.В. и др. Возможности и ограничения использования программных информационных устройств с искусственным интеллектом для диагностики и лечения заболеваний.
8. Minaee S., et al. Large language models: A survey // arXiv preprint arXiv:2402.06196. 2024.
9. Sarathy V., Scheutz M. Macgyver problems: Ai challenges for testing resourcefulness and creativity // Advances in Cognitive Systems. 2018. Т. 6. С. 31–44.
10. Runco M.A. AI can only produce artificial creativity // Journal of Creativity. 2023. Т. 33. № 3. С. 100063.
11. Mukherjee A., Chang H. The creative frontier of generative ai: Managing the novelty-usefulness tradeoff // arXiv preprint arXiv:2306.03601. 2023.
12. Molnár G., József C. AI Based Plagiarism Checking: Ease of use and applicable system for teachers to find similarities in students' assessments // 2022 IEEE 5th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE). IEEE, 2022. С. 000187–000192.
13. Soboroff I., Harman D. Novelty detection: the trec experience // Proceedings of human language technology conference and conference on empirical methods in natural language processing. 2005. С. 105–112.
14. Doboli S., et al. A cognitive inspired method for assessing novelty of short-text ideas // 2020 International Joint Conference on Neural Networks (IJCNN). IEEE, 2020. С. 1–8.
15. K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation, in Proc. of 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, Oct. 2014. P. 1724–1734.
16. Cer D., et al. Universal sentence encoder for English // Proceedings of the 2018 conference on empirical methods in natural language processing: system demonstrations. 2018. С. 169–174.
17. Tai K.S., Socher R., Manning C.D. Improved semantic representations from tree-structured long short-term memory networks // arXiv preprint arXiv:1503.00075. 2015.
18. Fisher D., et al. A real-time semantic model for relevance and novelty detection from group messages // 2022 International Joint Conference on Neural Networks (IJCNN). IEEE, 2022. С. 01–08.
19. Zhai C.X., Lafferty J. Two-stage language models for information retrieval // Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. 2002. С. 49–56.

20. Анохов И.В. Шкала уровня технологической готовности УГТ и перспективы ее модификации // Стратегии бизнеса. 2022. Т. 10. № 11. С. 289–294.
21. Hirshorn S. NASA System Engineering Handbook SP-2016-6105 Rev2 // Washington, DC. 2016.
22. Faide S. Assessing Bibliometrics for the Automation of Technology Readiness Level Assessments. University of Toronto (Canada), 2021.
23. Dastoor J., et al. A Bibliometric Approach to Characterizing Technology Readiness Levels Using Machine Learning // AIAA SCITECH 2023 Forum. 2023. С. 2686.

References

1. Helm J.M., et al. (2020) Machine learning and artificial intelligence: definitions, applications, and future directions. Current reviews in musculoskeletal medicine. T. 13. P. 69–76.
2. Sheikh H., Prins C., Schrijvers E. (2023) Artificial intelligence: definition and background. Mission AI: The new system technology. Cham: Springer International Publishing. P. 15–41.
3. Murphy R.R. (2019) Introduction to AI robotics. MIT press.
4. Gupta S., et al. (2022) Artificial intelligence for decision support systems in the field of operations research: review and future scope of research. Annals of Operations Research. T. 308. No. 1. P. 215–274.
5. Kalla D., et al. (2023) Study and analysis of chat GPT and its impact on different fields of study. International journal of innovative science and research technology. T. 8. №. 3.
6. Wu Z., et al. (2024) Deepseek-v12: Mixture-of-experts vision-language models for advanced multimodal understanding. arXiv preprint arXiv:2412.10302.
7. Tolmachev I.V. et al. *Vozmozhnosti i ogranicheniya ispol'zovaniya programmnykh informatsionnykh ustroystv s iskusstvennym intellektom dlya diagnostiki i lecheniya zabolevaniy* [Opportunities and limitations of using software-based information devices with artificial intelligence for diagnosing and treating diseases].
8. Minaee S., et al. (2024) Large language models: A survey. arXiv preprint arXiv:2402.06196.
9. Sarathy V., Scheutz M. (2018) Macgyver problems: Ai challenges for testing resourcefulness and creativity. Advances in Cognitive Systems. T. 6. P. 31–44.
10. Runco M.A. (2023) AI can only produce artificial creativity. Journal of Creativity. T. 33. No. 3. P. 100063.
11. Mukherjee A., Chang H. (2023) The creative frontier of generative ai: Managing the novelty-usefulness tradeoff. arXiv preprint arXiv:2306.03601.
12. Molnár G., József C. (2022) AI Based Plagiarism Checking: Ease of use and applicable system for teachers to find similarities in students' assessments. IEEE 5th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE). IEEE. P. 000187–000192.
13. Soboroff I., Harman D. (2005) Novelty detection: the trec experience. Proceedings of human language technology conference and conference on empirical methods in natural language processing. P. 105–112.
14. Doboli S., et al. (2020) A cognitive inspired method for assessing novelty of short-text ideas. International Joint Conference on Neural Networks (IJCNN). IEEE. P. 1–8.
15. K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation, in Proc. of 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, Oct. 2014, P. 1724–1734.
16. Cer D., et al. (2018) Universal sentence encoder for English. Proceedings of the 2018 conference on empirical methods in natural language processing: system demonstrations. P. 169–174.
17. Tai K.S., Socher R., Manning C.D. (2015) Improved semantic representations from tree-structured long short-term memory networks. arXiv preprint arXiv:1503.00075.
18. Fisher D., et al. (2022) A real-time semantic model for relevance and novelty detection from group messages. International Joint Conference on Neural Networks (IJCNN). IEEE. P. 01–08.
19. Zhai C.X., Lafferty J. (2002) Two-stage language models for information retrieval. Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. P. 49–56.

20. Anokhov I.V. *Shkala urovnya tekhnologicheskoy gotovnosti UGT i perspektivy ee modifikatsii* [Scale of the level of technological readiness of Technology Readiness Level and prospects of its modification] *Strategii biznesa* [Business strategies]. 2022. T. 10. № 11. P. 289–294.
21. Hirshorn S. (2016) NASA System Engineering Handbook SP-2016-6105 Rev2. Washington, DC.
22. Faidi S. (2021) Assessing Bibliometrics for the Automation of Technology Readiness Level Assessments. University of Toronto (Canada).
23. Dastoor J., et al. (2023) A Bibliometric Approach to Characterizing Technology Readiness Levels Using Machine Learning. AIAA SCITECH 2023 Forum. P. 2686.