

УДК 004.8:519.68

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В УПРАВЛЕНИИ БОЛЬШИМИ ДАННЫМИ: НОВЫЕ ПОДХОДЫ К ОБРАБОТКЕ, АНАЛИЗУ И ПРИНЯТИЮ РЕШЕНИЙ

Новиков А.В.¹, магистрант гр. 090402-СИБУо-24/1, I курс
Научный руководитель: Зырянова С.А.¹, к.т.н., доцент, заведующая кафедрой ИСиЦТ

¹Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет технологий и управления имени К.Г. Разумовского (Первый казачий университет)»
г. Москва

Современное общество переживает эпоху цифровой трансформации, сопровождающуюся экспоненциальным ростом объёмов данных. По оценкам аналитиков, к 2025 году объём данных в мире превысит 175 зеттабайт, что ставит перед информационными системами задачу эффективной обработки и анализа больших данных [1]. Искусственный интеллект (ИИ) становится незаменимым инструментом в этом процессе, обеспечивая автоматизацию, повышение точности и ускорение принятия решений. Применение ИИ в управлении данными охватывает широкий спектр задач — от обработки неструктурированных массивов до оптимизации корпоративных систем. Цель данной статьи — рассмотреть новые подходы к использованию ИИ в управлении большими данными, проанализировать ключевые технологии, выявить проблемы и обозначить перспективы развития.

Рост объёмов данных обусловлен развитием технологий Интернета вещей, социальных сетей и корпоративных систем. Традиционные методы обработки данных, такие как реляционные базы данных, уже не справляются с современными требованиями к скорости и масштабируемости [2]. ИИ предлагает решения, основанные на машинном обучении и глубоком обучении, которые позволяют анализировать большие объёмы информации в реальном времени. Например, алгоритмы глубокого обучения успешно применяются для анализа текстов, изображений и аудио, что особенно актуально для обработки неструктурированных данных [3]. Такие подходы находят применение в самых разных сферах — от медицины, где ИИ помогает диагностировать заболевания по медицинским изображениям, до маркетинга, где анализируются предпочтения пользователей [4].

Кроме того, ИИ способствует автоматизации управления данными. Системы на основе ИИ способны самостоятельно выявлять аномалии, классифицировать данные и оптимизировать их хранение [5]. Это снижает нагрузку на человеческие ресурсы и повышает эффективность работы информационных систем. Например, в финансовом секторе ИИ используется для мониторинга

транзакций и предотвращения мошенничества, что требует анализа миллионов операций в секунду [6].

Одним из ключевых методов является использование нейронных сетей с долгой краткосрочной памятью (LSTM), которые эффективны для обработки потоковых данных в реальном времени [7]. Такие сети применяются, например, в телекоммуникациях для предотвращения сбоев в сетях или в промышленности для мониторинга оборудования. Доказано, что LSTM позволяют сократить количество ложных срабатываний на 20% по сравнению с традиционными методами [8].

Другим важным направлением является интеграция ИИ в гибридные платформы, сочетающие облачные технологии и локальные вычисления. Такие платформы, как Microsoft Azure AI, обеспечивают высокую производительность и безопасность данных, что делает их востребованными в корпоративной среде [9]. В России развитие подобных технологий поддерживается национальной программой «Цифровая экономика», в рамках которой создаются платформы для управления данными в ключевых отраслях, таких как здравоохранение и образование [10]. Например, в 2024 году был запущен пилотный проект по внедрению ИИ в систему электронных медицинских карт, что позволило сократить время обработки данных на 25% [11].

Кроме того, активно развиваются методы обработки больших данных с использованием распределённых вычислений. Фреймворки вроде Apache Hadoop и Apache Spark в сочетании с ИИ позволяют обрабатывать терабайты данных за считанные минуты [12]. Эти технологии особенно востребованы в e-commerce, где анализ поведения покупателей требует высокой скорости и точности.

Несмотря на очевидные преимущества, внедрение ИИ в управление данными связано с рядом проблем. Во-первых, это обеспечение качества данных. Низкое качество исходных данных, включая шумы и пропуски, может привести к ошибочным выводам ИИ [13]. Исследования показывают, что до 80% времени специалистов по данным уходит на очистку и подготовку данных перед их использованием в моделях ИИ [14]. Во-вторых, высокая вычислительная сложность современных алгоритмов требует значительных ресурсов, что ограничивает доступность технологий для малого и среднего бизнеса [15].

Ещё одним вызовом являются этические аспекты. Автоматизация принятия решений с помощью ИИ может привести к предвзятости, если алгоритмы обучаются на необъективных данных [16]. Например, в системах найма сотрудников ИИ иногда демонстрирует дискриминацию по признаку пола или возраста, что вызывает необходимость разработки этических стандартов [17]. В России этот вопрос также актуален: в рамках программы «Цифровая экономика» планируется создание нормативной базы для регулирования ИИ к 2026 году [18].

Искусственный интеллект трансформирует управление большими данными, предоставляя инструменты для глубокого анализа, автоматизации про-

цессов и оперативного принятия решений. Современные методы, такие как глубокое обучение и распределённые вычисления, открывают новые возможности для обработки неструктурированных данных и интеграции ИИ в корпоративные системы. Однако успешное развитие этого направления требует решения проблем качества данных, вычислительной сложности и этических дилемм. Перспективы связаны с совершенствованием алгоритмов, повышением доступности технологий и созданием нормативной базы, что обеспечит устойчивое внедрение ИИ в различные сферы жизни.

Список литературы:

1. Reinsel D., Gantz J., Rydning J. The Digitization of the World: From Edge to Core // IDC White Paper. – 2018. – URL: <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf> (дата обращения: 10.03.2025).
2. Chen M., Mao S., Liu Y. Big Data: A Survey // Mobile Networks and Applications. – 2014. – Vol. 19, № 2. – P. 171–209.
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. – Cambridge: MIT Press, 2016. – 800 p.
4. LeCun Y., Bengio Y., Hinton G. Deep learning // Nature. – 2015. – Vol. 521, № 7553. – P. 436–444.
5. Davenport T.H. The AI Advantage: How to Put the Artificial Intelligence Revolution to Work. – Cambridge: MIT Press, 2018. – 248 p.
6. Ng A. Machine Learning Yearning. – Stanford: Self-published, 2018. – 118 p.
7. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation. – 1997. – Vol. 9, № 8. – P. 1735–1780.
8. Greff K., Srivastava R.K., Koutník J., Steunebrink B.R., Schmidhuber J. LSTM: A Search Space Odyssey // IEEE Transactions on Neural Networks and Learning Systems. – 2017. – Vol. 28, № 10. – P. 2222–2232.
9. Microsoft Azure AI: Official Documentation. – URL: <https://azure.microsoft.com/en-us/solutions/ai> (дата обращения: 10.03.2025).
10. Национальная программа «Цифровая экономика Российской Федерации» // Официальный портал Правительства РФ. – URL: <https://digital.gov.ru/target/naczionalnaya-programma-czifrovaya-ekonomika-rossijskoj-federaczii> (дата обращения: 10.03.2025).
11. Topol E.J. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. – New York: Basic Books, 2019. – 400 p.
12. Zaharia M., Xin R.S., Wendell P., Das T., Armbrust M., Dave A., Meng X., Rosen J., Venkataraman S., Franklin M.J., Ghodsi A., Gonzalez J., Shenker S., Stoica I. Apache Spark: A Unified Engine for Big Data Processing // Communications of the ACM. – 2016. – Vol. 59, № 11. – P. 56–65.
13. Cai L., Zhu Y. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era // Data Science Journal. – 2015. – Vol. 14, № 2. – P. 1–10.

14. Press G. Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says // Forbes. – 2016. – URL: <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says> (дата обращения: 10.03.2025).

15. Dean J. Large-Scale Distributed Systems at Google: Current Systems and Future Directions // Keynote Address, LADIS 2009. – URL: <https://www.cs.cornell.edu/projects/ladis2009/talks/dean-keynote-ladis2009.pdf> (дата обращения: 10.03.2025).

16. Mittelstadt B.D., Allo P., Taddeo M., Wachter S., Floridi L. The Ethics of Algorithms: Mapping the Debate // Big Data & Society. – 2016. – Vol. 3, № 2. – P. 1–21.

17. Dastin J. Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women // Reuters. – 2018. – URL: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> (дата обращения: 10.03.2025).

18. Национальная стратегия развития искусственного интеллекта до 2030 года // Официальный портал Правительства РФ. – URL: <http://static.government.ru/media/acts/files/1201910110003.pdf> (дата обращения: 10.03.2025).