



ДЕКАБРЬ 2025 | ВЫПУСК №2

КИБЕРНАУКА

МЕЖДУНАРОДНЫЙ НАУЧНЫЙ
ЭЛЕКТРОННЫЙ ЖУРНАЛ



<http://cybernauka.ru/>

**НАУЧНАЯ ОБЩЕСТВЕННАЯ ОРГАНИЗАЦИЯ
ПРОФЕССИОНАЛЬНАЯ НАУКА**

КИБЕРНАУКА

МЕЖДУНАРОДНЫЙ НАУЧНЫЙ ЭЛЕКТРОННЫЙ ЖУРНАЛ

№ 2/2025

www.cybernauka.ru

Нижний Новгород, 2025

УДК 004

ББК 30

Международный научный электронный журнал «Кибернаука», Нижний Новгород:
НОО «Профессиональная наука» - №2 - 2025. – 70 с.

ISSN 3033-8131

Статьи журнала содержат информацию, охватывающие исследования в области взаимодействия человека, технологий и социокультурной среды в условиях цифровой трансформации.

Журнал открыт для исследователей, преподавателей, аспирантов и практиков, заинтересованных в осмыслении роли технологий в современном мире.

Все включенные в журнал статьи прошли научное рецензирование и опубликованы в том виде, в котором они были представлены авторами. За содержание статей ответственность несут авторы.

Электронная версия журнала находится в свободном доступе на сайте www.cybernauka.ru

УДК 004

ББК 30

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Краснова Наталья Александровна — кандидат экономических наук, доцент, руководитель НОО «Профессиональная наука»

РЕДАКЦИОННЫЙ СОВЕТ

1. **Тойгамбаев Серик Кокибаевич** — кандидат технических наук, доцент. Член Союза журналистов России, Академик Международной общественной академии экологической безопасности природообустройства (МОАЭБП). профессор кафедры техническая эксплуатация технологических машин и оборудования природообустройства, РГАУ-МСХА им. К.А. Тимирязева

2. **Сагитов Рамиль Фаргатович** — кандидат технических наук, доцент, заместитель директора по научной работе в ООО «Научно-исследовательский и проектный институт экологических проблем», г. Оренбург

3. **Бюллер Елена Александровна** — кандидат экономических наук, доцент. ФГБОУ ВО «Адыгский государственный университет»

4. **Тимчук Оксана Григорьевна** — кандидат экономических наук, доцент кафедры «Экономики и цифровых бизнес-технологий». Иркутский национальный исследовательский технический университет

5. **Тихалева Елена Юрьевна** — кандидат юридических наук, доцент. Сотрудник Академии ФСО России

6. **Краснов Андрей Николаевич** — специалист в области математики, физики, программирования, старший программист АО «Intel»

**Материалы печатаются с оригиналов, поданных в оргкомитет,
ответственность за достоверность информации несут авторы статей**

© НОО Профессиональная наука, 2015-2025

Оглавление

БИОИНФОРМАТИКА И БИОКИБЕРНЕТИКА, НЕЙРОТЕХНОЛОГИИ, РОБОТОТЕХНИКА, VR/AR, КВАНТОВЫЕ ТЕХНОЛОГИИ, ЭНЕРГОЭФФЕКТИВНЫЕ ВЫЧИСЛЕНИЯ..... 6

Lipatov M.S., Nashuk D.S. ML models for predictive maintenance of hot-water boilers 6

Москаленко П.А., Нечитайлов В.В. Локальная адаптивная логика в системах управления упаковочными линиями при кратковременных сбоях электропитания 12

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ..... 22

Lipatov M.S., Isakov A.P. Development of a digital twin of a cooling tower for predictive condition management..... 22

Бугаенко А.П., Тарасова А.П., Интеграция внешних источников данных в процесс генерации текста в больших языковых моделях..... 26

Литвинова А.В., Чернова Д.В. Машинное обучение и разработка эффективных промышленных материалов..... 35

Сорокин Н.А., Леонова Н.Л. Машинное обучение в задачах анализа временных рядов для прогнозирования тенденций финансового рынка 49

КИБЕРЭКОНОМИКА И ЭКОНОМИКА ДАННЫХ 56

Коновалова В.К., Кириллова В.А. «Легкая» цифровизация как инструмент повышения эффективности логистики 56

БИОИНФОРМАТИКА И БИОКИБЕРНЕТИКА, НЕЙРОТЕХНОЛОГИИ, РОБОТОТЕХНИКА, VR/AR, КВАНТОВЫЕ ТЕХНОЛОГИИ, ЭНЕРГОЭФФЕКТИВНЫЕ ВЫЧИСЛЕНИЯ

УДК 004.021

Lipatov M.S., Nashuk D.S. ML models for predictive maintenance of hot-water boilers

Lipatov Maxim Sergeevich,

Senior Lecturer of the Department of Heat Power Installations and Heat Engines,
St. Petersburg State University of Industrial Technologies and Design.
Higher School of Technology and Energy

Nashuk Dmitriy Sergeevich,

Student of the Department of Heat Power Installations and Heat Engines,
St. Petersburg State University of Industrial Technologies and Design.
Higher School of Technology and Energy

Abstract. The article examines modern approaches to predictive maintenance of hot-water boilers based on machine learning methods. An overview of typical boiler failures, key diagnostic parameters and architectures of ML models used for predicting technical condition is presented. Special attention is paid to practical examples, advantages over traditional maintenance methods, and existing implementation barriers. It is shown that the integration of machine learning into monitoring systems can reduce the risk of accidents, minimize downtime and improve the energy efficiency of heating systems.

Keywords: predictive maintenance, hot-water boiler, machine learning, IoT, thermal power engineering, equipment reliability.

Рецензент: Торопцев Василий Владимирович - кандидат технических наук, доцент.
ФГБОУ ВО «РГАУ-МСХА им. К.А. Тимирязева»

The modern centralized heat supply system of the Russian Federation is faced with the need to ensure reliable and uninterrupted operation of hot water boilers, which are the basis of heat generating capacities. The operation of boilers under conditions of high load, variable climatic conditions and often outdated equipment leads to various types of failures, resulting in economic losses and a decrease in the quality of heat supply to consumers. Traditional maintenance methods based on time-based routine maintenance or actual equipment failures do not effectively prevent sudden breakdowns and optimize equipment life. In this regard, the transition to a predictive maintenance strategy based on predicting the condition of equipment using modern digital technologies is becoming particularly relevant. Machine learning is a powerful tool for implementing such a strategy, allowing you to analyze large amounts of sensor data and identify hidden patterns that precede equipment failures [1].

Hot water boilers are complex technical systems in which various types of failures can occur. The most common problems are corrosion of heat exchange surfaces, scale formation

and deposits on the inner walls of pipes, overheating of individual components, clogging of burner devices, as well as failure of automation and control systems. These malfunctions develop gradually, manifesting themselves in changes in the operating parameters of the equipment long before the critical condition [2]. To detect such anomalies in a timely manner, constant monitoring of key parameters is necessary: water temperature at the boiler inlet/outlet, pressure in the circuit, coolant flow rate, flue gas composition, vibration characteristics of the housing and operating parameters of the burner devices [3]. Modern IoT monitoring systems allow collecting and transmitting data in real time, creating an information basis for the application of machine learning. (Fig. 1).

The use of machine learning algorithms for predictive maintenance of hot water boilers opens up new opportunities to improve the reliability of heat supply. Unlike the traditional approach, in which repairs are carried out according to a pre-set schedule regardless of the actual condition of the equipment, ML models allow you to determine the optimal moments for maintenance based on the analysis of current parameters and forecasting their dynamics. This approach not only reduces the risk of accidents, but also optimizes maintenance costs, avoiding both premature repairs and critical equipment wear.

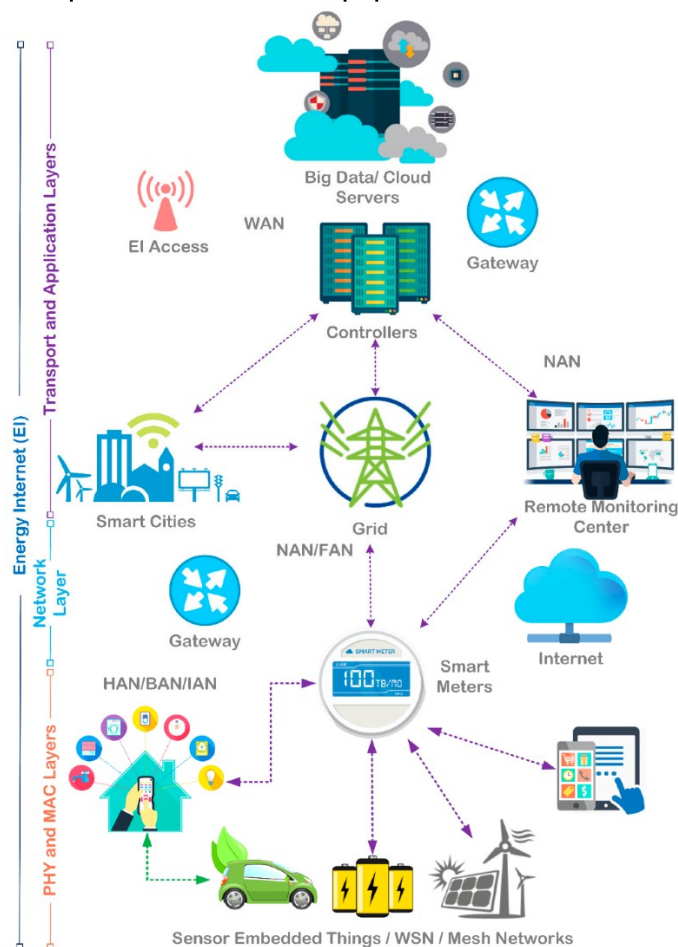


Figure 1. IoT-based predictive maintenance architecture

Modern ML algorithms for predictive maintenance can be classified according to several criteria. According to the type of tasks being solved, classification methods are distinguished (determining the condition of equipment as normal or abnormal), regression (predicting the remaining resource before failure) and anomaly detection (detecting deviations in operation without using labeled failure data). By architecture, algorithms are divided into traditional machine learning methods, deep learning, and hybrid approaches [4].

Ensemble algorithms, especially gradient boosting, occupy a special place among traditional methods. The XGBoost and LightGBM models have demonstrated high efficiency in predicting boiler equipment failures due to their ability to process heterogeneous data, noise tolerance, and the ability to assess the importance of individual features. These algorithms build a sequence of weak models (usually decision trees), each of which corrects the errors of the previous one, which makes it possible to achieve high prediction accuracy even if there are missing values in the data. Recurrent neural networks, especially LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Units), are effective for analyzing time series characteristic of boiler operation parameters. These architectures are able to take into account long-term time dependencies in the data, which makes it possible to detect a gradual deterioration in the characteristics of the equipment long before the failure occurs. For example, LSTM networks can analyze trends in pressure drop changes in the boiler pipe system, which serves as an indicator of scale formation, or a change in exhaust gas temperature, indicating a decrease in heat exchange efficiency. Anomaly detection methods such as Isolation Forest and Autoencoders are particularly useful when historical failure data is scarce. Isolation Forest works on the principle of isolating anomalies in random trees, which makes it possible to efficiently identify rare events without the need for a large amount of labeled data. Autoencoders, in turn, are trained to reproduce the normal operating modes of the equipment, and a significant deviation between the input data and their recovery indicates an abnormal condition. This approach is especially valuable for detecting new, previously unknown types of failures.

Bayesian networks and hidden Markov models are used to predict the remaining service life, taking into account probabilistic transitions between equipment states. These methods make it possible not only to determine the current condition of the boiler, but also to assess the likelihood of a critical condition in the near future, which makes it possible to plan maintenance based on real needs [5]. For the practical implementation of ML solutions in the monitoring systems of hot water boilers, it is necessary to take into account the architectural features of the implementation. A typical predictive maintenance system includes the following components: a sensor level with temperature, pressure, flow sensors, and gas analyzers; a

data acquisition and transmission level with edge computing for preprocessing; a server level with data storage and ML models; and a visualization and management level for boiler room personnel. The effectiveness of various ML approaches to predictive maintenance of hot water boilers can be assessed by several key parameters presented in Table 1.

Table 1

Comparative characteristics of ML algorithms for predictive maintenance of hot water boilers

Algorithm	Prediction accuracy, %	Data requirements	Learning rate	Noise tolerance
Random Forest	85–92	Medium	Medium	High
XGBoost	88–95	Medium	High	High
LSTM	90–96	High	Low	Medium
Isolation Forest	80–88	Low	High	High
Autoencoders	85–93	High	Low	Medium



Figure 2. Automated boiler room

Some experience has already been gained in the practical application of ML solutions for predictive maintenance of hot water boilers. In one of the pilot projects implemented at a boiler house in Germany, an XGBoost-based system analyzed data from 50 sensors, including water temperature, pressure, flue gas composition, and pump energy consumption. The model

successfully predicted the formation of scale in pipe systems 14-21 days before the critical condition, which allowed cleaning to be carried out as planned and to avoid an emergency stop. The economic effect amounted to 180 thousand euros per year due to lower repair costs, reduced downtime and increased energy efficiency [6].

The integration of machine learning into boiler control systems is a promising area of digital transformation of the thermal power industry. ML models make it possible to implement the principles of predictive maintenance, significantly increasing the reliability of equipment and reducing operating costs. Ensemble methods, especially XGBoost and LightGBM, as well as recurrent neural networks for time series analysis demonstrate the greatest efficiency. However, the widespread adoption of these technologies is hindered by insufficient digitalization of infrastructure, a shortage of personnel, and problems with interpretability of models. Further development of this area should be focused on creating hybrid models that combine the physical laws of boiler operation with adaptive machine learning algorithms. This will increase the confidence of the operating personnel in the forecasts and ensure the practical applicability of the solutions. In addition, it is necessary to modernize the measuring base in boiler houses with the installation of modern IoT sensors and the creation of unified standards for data collection and storage. The implementation of these measures will not only improve the reliability of heat supply, but also achieve significant economic benefits by optimizing maintenance costs and increasing the energy efficiency of equipment. In the context of global challenges in the field of energy and ecology, the transition to intelligent control systems for hot water boilers is becoming not just a technological trend, but a necessary condition for the sustainable development of thermal energy.

Further development of this area should be focused on creating hybrid models that combine the physical laws of boiler operation with adaptive machine learning algorithms. This will increase the confidence of the operating personnel in the forecasts and ensure the practical applicability of the solutions. In addition, it is necessary to modernize the measuring base in boiler houses with the installation of modern IoT sensors and the creation of unified standards for data collection and storage (Fig 2).

The implementation of these measures will not only improve the reliability of heat supply, but also achieve significant economic benefits by optimizing maintenance costs and increasing the energy efficiency of equipment. In the context of global challenges in the field of energy and ecology, the transition to intelligent control systems for hot water boilers is becoming not just a technological trend, but a necessary condition for the sustainable development of thermal energy.

Библиографический список

1. Цифровизация и новые функциональные требования, предъявляемые к тепловым электростанциям / С. Л. Горобченко, Д. А. Ковалев, Ю. И. Беленький [и др.] // Известия Тульского государственного университета. Технические науки. – 2023. – № 7. – С. 714-717.
2. Гиматдинов, И. Р. Алгоритмы машинного обучения для прогнозирования отказов оборудования на основе данных телеметрии / И. Р. Гиматдинов // Энергетика и автоматизация в современном обществе : Материалы VIII Всероссийской научно-практической конференции обучающихся и преподавателей, Санкт-Петербург, 15 мая 2025 года. – Санкт-Петербург: Санкт-Петербургский государственный университет промышленных технологий и дизайна, 2025. – С. 57-62.
3. Лисовин, Г. Б. Цифровая трансформация ТЭК / Г. Б. Лисовин // Потенциал и вызовы развития возобновляемой энергетики : Сборник научных статей Всероссийской научно-практической конференции с международным участием, Невинномысск, 19–23 декабря 2022 года. – Ставрополь: Северо-Кавказский федеральный университет, 2023. – С. 76-82.
4. Жигалов, К. Ю. Аспекты применения машинного обучения для предиктивного анализа состояния сложного технологического оборудования на основе исторических данных и текущих показателей / К. Ю. Жигалов // Актуальные аспекты развития науки и общества в эпоху цифровой трансформации : Сборник материалов XVI Международной научно-практической конференции, Москва, 25 июля 2024 года. – Москва: Центр развития образования и науки, 2024. – С. 93-101.
5. Максименков, А. С. Применение метода нечеткой логики при управлении котельной / А. С. Максименков // Актуальные вопросы науки, общества и образования : сборник статей Международной научно-практической конференции. В 2 ч., Пенза, 05 декабря 2024 года. – Пенза: Наука и Просвещение, 2024. – С. 78-80.
6. Хисматуллин, А. С. Система управления водогрейным котлом с использованием нечеткой логики / А. С. Хисматуллин, Р. Ш. Ишмуратов, А. М. Хафизов // Международный технико-экономический журнал. – 2022. – № 2. – С. 77-82.

УДК 621.316

Москаленко П.А., Нечитайлов В.В. Локальная адаптивная логика в системах управления упаковочными линиями при кратковременных сбоях электропитания

Local adaptive logic in packaging line control systems for short-term power failures

Москаленко Павел Анатольевич,

магистрант кафедры информационно-измерительных технологий и систем управления,
Санкт-Петербургский государственный университет промышленных технологий и дизайна.
Высшая школа технологии и энергетики

Нечитайлов Василий Васильевич,

Канд. техн. наук, доцент кафедры теплосиловых установок и тепловых двигателей,
Санкт-Петербургский государственный университет промышленных
технологий и дизайна. Высшая школа технологии и энергетики

Moskalenko Pavel Anatolievich,

Master's student of the Department of Information and Measuring Technologies and Control
Systems, St. Petersburg State University of Industrial Technologies and Design.

Higher School of Technology and Energy

Nechitaylov Vasily Vasilyevich,

Candidate of Technical Sciences, Associate Professor of the Department of heat power
installations and heat engines, St. Petersburg State University of Industrial Technologies and Design.
Higher School of Technology and Energy

Аннотация. В статье рассматривается применение локальной адаптивной логики для обеспечения автономного восстановления упаковочных линий после кратковременных отключений питания. Описана децентрализованная архитектура на базе отечественных модулей ОВЕН (МК110, МВ110), шлюзов МОХА, датчиков Danfoss и пускателей ПМЛ. Приведён алгоритм автономного перезапуска, реализованный во встраиваемом исполнителе SP-RTM-50, и результаты лабораторных испытаний: время восстановления сокращено с 320 до 45 с, брак – с 18–24 до 0–2 единиц. Показано, что комплект стоимостью порядка 55 000₽ окупается менее чем за 3 месяца при 8 сбоях в смену.

Ключевые слова: локальная адаптивная логика, упаковочная линия, ОВЕН, SCADA+, автономный перезапуск, ПМЛ.

Abstract. The article discusses the use of local adaptive logic to provide autonomous recovery of packaging lines after short-term power outages. A decentralized architecture is described on the basis of domestic OWEN modules (MK110, MV110), MOXA gateways, Danfoss sensors and PML. The algorithm of autonomous restart, implemented in the embedded performer SP-RTM-50, and the results of laboratory tests are given: the recovery time is reduced from 320 to 45 s, the scrap - from 18-24 to 0-2 units. It is shown that a kit costing about 55,000₽ pays off in less than 3 months with 8 failures per shift.

Keywords: local adaptive logic, packaging line, OWEN, SCADA+, autonomous restart, PML.

Рецензент: Торопцев Василий Владимирович - кандидат технических наук, доцент.
ФГБОУ ВО «РГАУ-МСХА им. К.А. Тимирязева»

Современные упаковочные линии, несмотря на внешнюю стандартизацию – роторные питатели, конвейеры, пневмопресс-формовщики, системы термосварки – в действительности представляют собой тонко настроенные механо-электронные комплексы, где даже доли секунды имеют значение. При скорости работы 6 000–12 000 упаковок в час каждая позиция сервопривода должна быть выверена с точностью до 0.1 мм, а давление в пневмосистеме – стабильно поддерживаться в пределах ± 0.03 МПа. В таких условиях кратковременный провал напряжения продолжительностью менее двух секунд, который в быту может пройти незамеченным, на производстве запускает каскадную цепочку событий: сначала теряется питание ПЛК, затем – синхронизация между осями, падает давление в ресивере, и уже к моменту, когда оператор подходит к панели, оборудование оказывается в аварийной позиции – с зажатым продуктом, деформированным картоном или разорванным полимерным полотном. Именно этот разрыв между физическим инерционным торможением механизмов (1–2 с) и логическим восстановлением управления (часто 3–5 мин) и составляет основную причину брака, который в 80 % случаев технически предотвратим [1].

На предприятиях Северо-Западного федерального округа, по данным отраслевого отчёта Ассоциации предприятий упаковки за 2024 год, более половины всех кратковременных простоев (до 5 минут) вызваны именно электромагнитными помехами и просадками напряжения в сетях 0.4 кВ – особенно в периоды пиковых нагрузок или при коммутациях на подстанциях. При этом большинство действующих линий построены по схеме централизованного управления: датчики → шлюз → SCADA-сервер → ПЛК → исполнительные механизмы. Такая архитектура удобна для мониторинга, но критически уязвима к задержкам передачи команды. Цикл «обнаружение сбоя → передача на сервер → принятие решения → отправка команды → исполнение» занимает в среднем от 300 до 800 мс – времени, которого недостаточно для перехвата управляемого замедления приводов. В результате сервоприводы, потеряв связь, переходят в режим «свободного хода», и их оси, вращаясь по инерции, уходят из расчётных позиций. При попытке последующего пуска ПЛК обнаруживает расхождение между фактическим и ожидаемым положением и блокирует старт до ручной перекалибровки – операция, требующая от 5 до 15 минут и почти неизбежно сопровождающаяся формированием 20–30 единиц брака.

Решение этой проблемы заключается не в усилении централизованного звена, а в его частичной децентрализации – в вынесении критически важных решений на самый нижний уровень иерархии, непосредственно в шкаф управления линией. Локальная адаптивная логика – это не отдельное устройство, а принцип организации: алгоритм,

исполняемый автономно, без зависимости от сервера, на базе компактных, но надёжных промышленных модулей, таких как ОБЕН МК110-8DN.4R и MV110-224.8AC. Эти устройства отличаются прежде всего гальванической изоляцией между цепями питания, измерения и интерфейса (до 3000 В), что позволяет устанавливать их в непосредственной близости от частотных преобразователей без риска сбоев из-за высокочастотных помех [2]. Модуль МК110-8DN.4R обеспечивает восемь дискретных входов для приёма сигналов от датчиков положения, концевых выключателей, аварийных реле, а также четыре релейных выхода, способных напрямую управлять пускателями – без промежуточных реле. Это принципиально важно: цепь «модуль → пускатель → привод» замыкается за 15–20 мс, тогда как при участии центрального ПЛК – за 300–500 мс.

Для измерения аналоговых параметров – в первую очередь давления в пневмосистеме – используется датчик Danfoss MBS 3000, компактный (диаметр 42 мм), выполненный из нержавеющей стали AISI 316L, с диапазоном 0–1.6 МПа и выходным сигналом 4–20 мА. Его ключевое преимущество – встроенная защита от гидроудара, демпфирующий сапун позволяет выдерживать до миллиона циклов резких перепадов давления без дрейфа нуля. Датчик подключается к аналоговому входу модуля MV110-224.8AC, который, в свою очередь, имеет отдельную гальваническую развязку для каждой пары каналов – это исключает взаимное влияние сигналов от разных зон линии, например, от системы закупорки и наполнения.

Физическое исполнение команд обеспечивает пускатель КЭАЗ ПМЛ-2200-024-УХЛ4 – нереверсивное устройство с кнопками «Пуск» и «Стоп», катушкой 220 В и степенью защиты IP54. Его особенность – пониженный порог срабатывания: включение гарантировано при напряжении 185 В (84 % от номинала), что позволяет использовать его в схемах с источниками бесперебойного питания, выдающими несинусоидальное напряжение. Это свойство критично при кратковременных сбоях, когда АВР или ИБП включаются с задержкой, и напряжение в сети оказывается нестабильным [3]. Пускатель допускает установку в производственных помещениях с относительной влажностью до 95 %, что соответствует требованиям ВСН 60-89 к оборудованию, размещаемому в зонах с повышенной влажностью.

Подключение всех компонентов осуществляется с соблюдением требований к электромагнитной совместимости. Для цепей управления используется контрольный кабель КИПЭВнг(А)-FRLS 5×0.75 – с медными жилами, экранирующей оплёткой (эквивалентное сечение 50 мм²) и оболочкой пониженной горючести. Экран заземляется только с одной стороны – со стороны шкафа управления – во избежание контурных

токов, которые при длине линии более 15 м могут вызывать ошибку измерения давления до 8 %. Питание модулей, шлюза и реле обеспечивает промышленный ИБП ОВЕН АВВ-20.4 (24 В, 4 А·ч), способный поддерживать работу локального узла в течение 2 ч 45 мин при нагрузке 30 Вт. В отличие от компьютерных ИБП, он переходит в автономный режим менее чем за 10 мс, что гарантирует непрерывность питания даже при полном исчезновении 220 В.

Связь с централизованной системой осуществляется через шлюз MOXA MGate MB3180, поддерживающий до 16 TCP-мастеров и 31 RTU-устройство по протоколу Modbus. Его функция автокалибруемого таймаута позволяет адаптироваться к изменяющимся условиям канала – особенно полезно при модернизации старых линий с нестандартной кабельной инфраструктурой. Однако при потере связи шлюз не блокирует локальные циклы: архитектура построена по принципу «централизованное наблюдение – децентрализованное действие», что соответствует рекомендациям Ростехнадзора (РД 11-349-01): базовая безопасность процесса должна обеспечиваться автономно, независимо от состояния верхних уровней [4].

Алгоритм локальной логики реализован в среде SCADA+ SP-DEV-1 и размещён во встраиваемом исполнителе SP-RTM-50, установленном непосредственно в шкафу управления. Цикл исполнения – не более 50 мс, что в 6–10 раз быстрее централизованного подхода. Логика строится вокруг последовательности «обнаружение → диагностика → действие → проверка → фиксация». При падении напряжения 24 В модуль МК110 фиксирует сбой после 200-мс фильтра дребезга (это исключает ложные срабатывания при коммутационных перенапряжениях). Затем анализируется давление в пневмосистеме: если оно не ниже 0.4 МПа, формируется импульс на дискретный выход для включения пускателя и одновременно подаётся сигнал «Разрешение запуска» на вход частотного преобразователя Danfoss FC 51. Если давление недостаточно – запускается компрессор на 15 с, после чего проверка повторяется. После успешного пуска включается режим «мягкого старта»: частота нарастает линейно в течение 2 с, что предотвращает гидроудар и перегрузку привода. В течение последующих 10 с проверяется стабильность параметров – давление, сигнал «Готов» от сервопривода, завершение цикла формовщика. При неудаче допускается до двух повторных попыток с паузой 120 с для охлаждения. После третьей неудачи система блокируется и формирует аварийный сигнал, который при восстановлении связи передаётся в SCADA.

Важно, что все временные параметры – длительность фильтра, время запуска компрессора, период мягкого старта, количество попыток – являются настраиваемыми

и хранятся в энергонезависимой памяти SP-RTM-50 [5]. Это позволяет адаптировать логику под разные типы линий: для асептической упаковки жидких продуктов, где критична стерильность, требуется более агрессивный перезапуск с принудительной продувкой азотом; для линий гофротары – наоборот, более осторожный подход из-за высокой инерции механизмов. Для наглядности в таблице 1 приведены рекомендуемые параметры для трёх основных классов упаковочных линий.

Таблица 1

Рекомендуемые временные параметры локальной логики для различных типов упаковочных линий

Параметр	Упаковка жидких продуктов (асептическая)	Упаковка сухих сыпучих (чай, крупы)	Упаковка картонной тары (гофрокороба)
Макс. длительность сбоя для восстановления, с	1.5	2.2	3.0
Мин. давление для перезапуска, МПа	0.35	0.40	0.45
Длительность «мягкого старта», с	1.0	1.5	2.5
Макс. кол-во попыток	2	3	1
Требуется ли компрессор при сбое?	Да (азотная продувка)	Нет	Да (пневмопресс)

Такой подход позволяет не просто реагировать на сбой, а адаптироваться к его характеру – отсюда и название «локальная адаптивная логика». Она не требует замены основного оборудования, не нарушает существующие схемы, но добавляет системе «второй слой» устойчивости – тот самый, который превращает аварию в инцидент, а инцидент – в событие, не требующее человеческого вмешательства.

Для проверки работоспособности предложенной архитектуры был собран лабораторный стенд, имитирующий участок упаковочной линии – формовку, наполнение и закупорку – с типичными для реального производства параметрами инерции, давления и тактовой частоты. В основу стенда положена роторная упаковочная машина УПАК-3М, модернизированная под управление через интерфейсные модули ОБЕН. В качестве исполнительных устройств использовались сервопривод Delta ASDA-B3 (0.75 кВт) для роторного питателя, частотный преобразователь Danfoss FC 51 для конвейера, пневмопресс-формовщик с цилиндром Ø50 мм и ресивером объёмом 10 л, а также термосварочная головка с электронагревателями 300 Вт. Давление в пневмосистеме контролировалось датчиком Danfoss MBS 3000 (0–1.6 МПа, 4–20 мА), положение осей – энкодерами Omron E6B2-CWZ6C (1000 имп./об), а состояние конечных операций – бесконтактными датчиками ОБЕН ПДУ-2103 (NPN, 3–30 В). Все сигналы поступали на модули МК110-8DN.4R и

MV110-224.8AC, установленные на DIN-рейке в шкафу управления, рядом с пускателем КЭАЗ ПМЛ-2200-024-УХЛ4 и ИБП ОВЕН АBB-20.4. Связь с ПК, на котором работала SCADA+ версии 6.2.1, осуществлялась через шлюз MOXA MB3180 по протоколу Modbus TCP. Стенд позволял имитировать кратковременные отключения питания (0.3–5 с) с помощью реле времени Finder 80.22, а также моделировать падение давления путём принудительного сброса воздуха через электромагнитный клапан Festo MNJ9-MS.

Тестирование проводилось по трём сценариям: лёгкий сбой (0.5 с), средний (2 с) и тяжёлый (4.5 с). При лёгком сбое инерционное вращение приводов сохранялось в полной мере – оси смещались не более чем на 15–18 мм, давление в ресивере падало до 0.42 МПа, и локальная логика восстанавливала стабильную работу за 38–42 с без единого бракованного изделия. При среднем сбое наблюдались умеренные дисбалансы: питатель и формовщик выходили из фазы на 40–60 мм, давление снижалось до 0.31 МПа, компрессор включался автоматически, и после одной повторной попытки система стабилизировалась за 52–58 с при 1–2 единицах брака (упаковки с неполной закупоркой, пригодные для ручной доупаковки). При тяжёлом сбое – когда инерция полностью терялась, а давление падало до 0.15 МПа – алгоритм выполнял две попытки перезапуска с паузой 120 с, после чего блокировался и формировал аварийный сигнал, требующий ручного вмешательства; в этом случае брак составлял 8–12 изделий, но время простоя ограничивалось 210 с вместо 410–470 с при ручном перезапуске. В сравнении с централизованной системой на базе Siemens S7-1200, где при тех же условиях восстановление занимало 320–380 с и сопровождалось 18–24 единицами брака, локальная архитектура показала не только ускорение реакции, но и значительное снижение риска повреждения оборудования – за счёт отсутствия рывков при «холодном старте» и плавного нарастания частоты в режиме мягкого пуска [6].

Экономический эффект от внедрения локальной адаптивной логики требует дифференцированного подхода – брак в упаковочном производстве неоднороден по стоимости и последствиям. Упаковка с нарушением герметичности или маркировки подлежит полному изъятию из оборота, тогда как дефекты внешнего вида (морщины, смещение печати) могут быть допущены к реализации по сниженной цене. Это позволяет выделить три категории брака: лёгкий (внешний вид – стоимость потери 15–30 % от номинала), средний (частичная негерметичность – 60–80 % потери), тяжёлый (полная потеря герметичности, нарушение состава – 100 % потери). Для типовой линии, упаковывающей сок в тетрапак (стоимость одной упаковки с продуктом – 12.5₽), расчёт потерь при 21 единице брака (среднее значение для централизованной системы)

выглядит следующим образом: при 60 % тяжёлого, 30 % среднего и 10 % лёгкого брака экономический ущерб составляет:

$$(0.6 \cdot 21 \cdot 12.5) + (0.3 \cdot 21 \cdot 0.7 \cdot 12.5) + (0.1 \cdot 21 \cdot 0.2 \cdot 12.5) = 157.5 + 55.1 + 5.3 \\ = 217.9 \text{ руб. за инцидент.}$$

При 8 сбоях в смену и 25 рабочих днях в месяц это даёт 43 580₽/мес. В системе с локальной логикой, где брак сокращается до 1–2 изделий (в основном лёгкий), ущерб снижается до 3.8–7.5₽/инцидент, или 760–1500₽/мес. Таким образом, чистая экономия от снижения брака составляет от 42 000₽ до 43 000₽ ежемесячно. К этому следует добавить сокращение простоев: при средней скорости линии 10 000 уп./ч каждый час простоя – это упущенная выручка 125 000₽. Снижение времени восстановления с 320 до 45 с экономит $275 \text{ с} \times 8 = 36.7 \text{ мин/смену}$, или 76 450₽/мес дополнительно. Общий эффект – до 119 000₽/мес, при стоимости комплекта локальной логики 54 800₽.

Сравнение с зарубежными аналогами подтверждает экономическую и техническую целесообразность отечественного решения. Система на базе Siemens S7-1212C DC/DC/DC (82 000₽), ET200SP (35 000₽), датчиков SITRANS P320 (19 500₽) и пускателей Sirius 3RA2 (12 300₽) обойдётся в 148 800₽ – почти втрое дороже, при сопоставимом функционале. Более того, цикл исполнения логики в S7-1200 составляет 100–150 мс (по данным Siemens Catalog ST 70), что не позволяет реализовать «в-flight restart» – перезапуск приводов в режиме инерционного вращения. Аналогичная ситуация с Allen-Bradley CompactLogix 5370: стоимость комплекта – от 165 000₽, цикл сканирования – 80–200 мс (Allen-Bradley Publication 1769-UM018, 2024). При этом ни одна из этих систем не обеспечивает автономной работы при полной потере связи с сервером – все программы хранятся в процессоре, и при его отключении логика останавливается. В отличие от них, архитектура на базе ОВЕН и SP-RTM-50 построена по принципу «логика – в исполнителе», что соответствует современным трендам edge computing и отказоустойчивого проектирования [4].

Для визуализации эффективности локального восстановления использовались трендовые графики SCADA+, отражающие динамику давления в пневмосистеме и положения сервооси питателя в течение 120 с, включая момент сбоя ($t = 60 \text{ с}$). Для наглядной иллюстрации практической реализации локальной адаптивной логики на Рисунке 1 приведена фотография типового промышленного шкафа управления, собранного в соответствии с предложенной архитектурой. На изображении отчётливо видны DIN-рейки с установленными модулями (в центре – чёрный блок, соответствующий по габаритам и расположению модулю ОВЕН МК110-8DN.4R),

пускателями и автоматическими выключателями (жёлтые и красные элементы), а также кабельная разводка с использованием экранированных проводов. В правой части шкафа расположена панель с сигнальными лампами и кнопками аварийной остановки – стандартный элемент для линий, работающих в режиме непрерывного производства. Такая компоновка типична для модернизации действующих упаковочных линий и не требует замены шкафа целиком – достаточно выделить 300–400 мм ширины на DIN-рейке для установки новых модулей. Подобные изображения доступны в открытых источниках: например, на официальном сайте компании «ЭлектроШкаф» (раздел «Типовые решения») или в каталогах Schneider Electric – раздел «Промышленные шкафы управления».



Рисунок 1. Типовой промышленный шкаф управления упаковочной линией с компонентами локальной адаптивной логики

Такая визуализация подчёркивает простоту и доступность решения: ни одна из деталей не является специальной или заказной – все элементы имеются в свободной продаже, сертифицированы и совместимы друг с другом без дополнительных интерфейсных преобразователей. Это особенно важно для предприятий с ограниченными ресурсами, где критична не только функциональность, но и скорость внедрения – шкаф, подобный показанному на рисунке, может быть собран и запущен в течение одной смены силами одного электромонтёра и одного программиста [5].

Важно подчеркнуть, что локальная адаптивная логика не противопоставляется централизованной SCADA – она дополняет её, занимая строго определённый уровень

иерархии. Как указано в Методических рекомендациях Минстроя России по цифровизации промышленных предприятий (2024), оптимальная архитектура включает три уровня: низший (0–1) – для базовой устойчивости и безопасности, средний (2) – для мониторинга и оперативного управления, высший (3) – для аналитики и прогнозирования. Локальная логика однозначно относится к уровню 0–1: она не анализирует причины сбоя, не строит прогнозы, не формирует отчёты – она лишь обеспечивает «безопасный минимум»: сохранение давления, позиций и готовности к пуску. Это позволяет SCADA-оператору сосредоточиться на действительно сложных задачах – настройке рецептур, балансировке нагрузки, взаимодействии с MES, – не тратя время на рутинные перезапуски.

Внедрение локальной адаптивной логики в упаковочные линии – это не революция, а эволюционное развитие, основанное на рациональном использовании доступных, сертифицированных, отечественных компонентов. Оно не требует замены действующего оборудования, не нарушает существующие схемы, но добавляет системе гибкость, автономность и отказоустойчивость – качества, всё более востребованные в условиях нестабильной внешней инфраструктуры. При стоимости комплекта в 55 000₽ и ежемесячной экономии до 119 000₽ срок окупаемости составляет менее 15 дней, а вклад в надёжность, безопасность и ресурсосбережение трудно переоценить. Перспективы развития – в интеграции с цифровыми двойниками (для имитации сценариев сбоев при проектировании), с системами технического зрения (для автоматической оценки качества восстановления) и с платформами edge-аналитики (для прогнозирования износа компрессоров и сервоприводов по трендам давления и тока). Но уже сегодня, без ожидания «завтрашних технологий», локальная адаптивная логика позволяет сделать упаковочное производство не просто автоматизированным, а по-настоящему устойчивым.

Библиографический список

1. Якшибаев, А. Р. Достижения в области адаптивных и оптимальных систем управления: обзор последних исследований / А. Р. Якшибаев, О. В. Борисова // Формирование профессиональной направленности личности специалистов - путь к инновационному развитию России: Сборник статей V Всероссийской научно-практической конференции, Пенза, 18–19 декабря 2023 года. – Пенза: Пензенский государственный аграрный университет, 2023. – С. 501-503.
2. Васюткин, Д. П. Способы описания бизнес-процессов с помощью ИПИ-технологий / Д. П. Васюткин, О. В. Борисова // Современные цифровые технологии:

Материалы II Всероссийской научно-практической конференции, Барнаул, 01 июня 2023 года / Под общей редакцией А.А. Беушев, А.С. Авдеев, Е.Г. Боровцов, А.Г. Зрюмова. – Барнаул: Алтайский государственный технический университет им. И.И. Ползунова, 2023. – С. 187-189.

3. Фаизов, Н. Н. Автоматизация противопожарных систем с использованием частотного регулирования / Н. Н. Фаизов, О. В. Борисова // Современные тенденции развития фундаментальных и прикладных наук: Материалы VII Всероссийской научно-практической конференции, Брянск, 25 января 2024 года. – Брянск: Брянский государственный инженерно-технологический университет, 2024. – С. 374-376.

4. Панфилова, Д. В. Сравнительный анализ датчиков давления для автоматизации газовых котельных / Д. В. Панфилова, О. В. Борисова // Актуальные аспекты развития науки и общества в эпоху цифровой трансформации: сборник материалов XII Международной научно-практической конференции, Москва, 05 декабря 2023 года. – Москва: Алеф, 2023. – С. 231-234.

5. Борисова, О. В. Автоматизация управления информационными системами / О. В. Борисова, А. Р. Зиннатуллин // Современные исследования: теория, практика, результаты: Сборник материалов Международной научно-практической конференции, Москва, 29 декабря 2023 года. – Москва: Центр развития образования и науки, ООО "Издательство АЛЕФ", 2023. – С. 427-430.

6. Габидуллин, И. Р. Особенности расчета надежности систем / И. Р. Габидуллин, О. В. Борисова // Грани науки 2023: Сборник трудов Международной научно-практической конференции, Ростов-на-Дону-Мадрид, 27 ноября 2023 года. – Ростов-на-Дону - Мадрид: ДиректСайнс? ООО "Сфера", 2023. – С. 141-147.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ

УДК 004.896

Lipatov M.S., Isakov A.P. Development of a digital twin of a cooling tower for predictive condition management

Lipatov Maxim Sergeevich,

Senior Lecturer of the Department of Heat Power Installations and Heat Engines,
St. Petersburg State University of Industrial Technologies and Design.
Higher School of Technology and Energy

Isakov Aleksandr Petrovich,

Student of the Department of Industrial Thermal Power Engineering,
St. Petersburg State University of Industrial Technologies and Design.
Higher School of Technology and Energy

Abstract. *The article discusses the problem of suboptimal maintenance of industrial cooling towers, leading to resource overruns and unplanned downtime. As a solution, the concept of a digital twin is proposed – a highly accurate dynamic mathematical model synchronized with a physical object in real time.*

Keywords: *digital twin, cooling tower, predictive analytics, predictive maintenance, mathematical modeling, residual resource.*

Рецензент: Торопцев Василий Владимирович - кандидат технических наук, доцент.
ФГБОУ ВО «РГАУ-МСХА им. К.А. Тимирязева»

The cooling tower functions as a critical element in the circulating water supply circuit of thermal power and chemical industry facilities. Its energy efficiency has a direct impact on the efficiency of the main technological cycle. Existing maintenance strategies based on fixed time intervals or reactive troubleshooting demonstrate systemic inefficiency. This is manifested in the occurrence of excessive costs associated with untimely maintenance operations, or significant economic losses due to unplanned equipment downtime. An urgent engineering task is to develop tools for the transition to a predictive management model based on continuous monitoring and predictive analysis of the actual technical condition. For complex thermal engineering equipment, such as a cooling tower, the implementation of this approach requires the construction of its highly accurate dynamic model – a digital twin [1]. The purpose of this study is to describe the architectural principles, key mathematical models, and practical aspects of implementing a digital twin of a cooling tower for predicting the condition and optimizing the repair cycle.

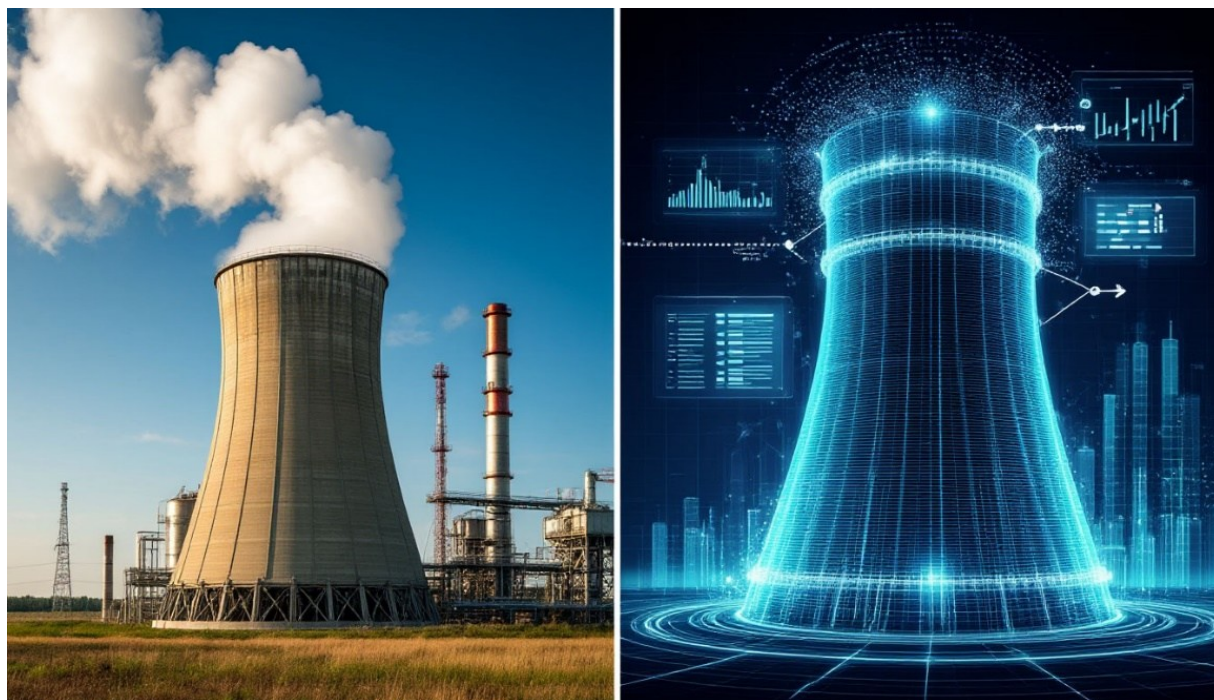


Figure 1. Block diagram of the digital twin architecture

The digital twin of a cooling tower is defined as an adaptive virtual model that is in a state of two-way communication with a physical object through a continuous stream of data and implements the functions of simulation, analysis and forecasting. The architecture of the system is formed by three interrelated layers. The physical layer includes a network of primary sensors installed on the cooling tower. Key parameters are recorded: inlet and outlet water temperatures, dry and wet thermometer temperatures, circulating water flow, vibration acceleration on the fan supports, pressure drop in the irrigation unit, as well as electrical parameters of the drive motors. The virtual layer is the computational core of the system, consisting of deterministic and predictive models. The deterministic heat-hydraulic model is used to calculate the theoretical (nominal) efficiency of a cooling tower under current environmental and load conditions [2]. Predictive degradation models use machine learning techniques to analyze historical trends. The fouling dynamics model correlates the increase in DP with water quality and operating time, predicting the moment when critical hydraulic resistance is reached. The mechanical wear model analyzes the time series and spectral composition of the vibration signal to estimate the remaining life of the bearing assemblies. The integration and visualization layer provides two-way communication: rapid calibration of models with up-to-date data and presentation of forecast results in the form of

specialized dashboards, as well as integration with maintenance management systems to automate the formation of repair requests.

The predictive ability of the digital twin is provided by a set of mathematical models. The fouling efficiency and dynamics assessment model uses heat and material balance equations to calculate theoretical cooling capacity. The deviation of the calculated efficiency indicator from the actual one, determined by sensor data, serves as an integral indicator of the pollution level.

The fouling model, implemented by regression analysis methods or on the basis of recurrent neural networks, predicts the time to reach the threshold value of the DRC, at which flushing becomes technologically necessary and economically justified. The model for predicting mechanical failures of the fan group is based on the analysis of vibration diagnostics. Diagnostic features are identified from the vibration acceleration time series: RMS values, amplitudes at characteristic rotational frequencies and harmonics, as well as statistical indicators. To predict the RUL, the method of extrapolating the trend of the dominant diagnostic feature using adaptive filters or ensembles of regression models is used [3]. The threshold values of the features are determined based on the passport tolerances of the equipment and the analysis of historical data on failures.

The implementation of the digital twin is carried out according to an iterative scheme, starting with a pilot project on a single cooling tower. At the audit stage, the existing instrument base is evaluated, the missing sensors are retrofitted, and retrospective data is collected. The developed models are consistently calibrated and validated.

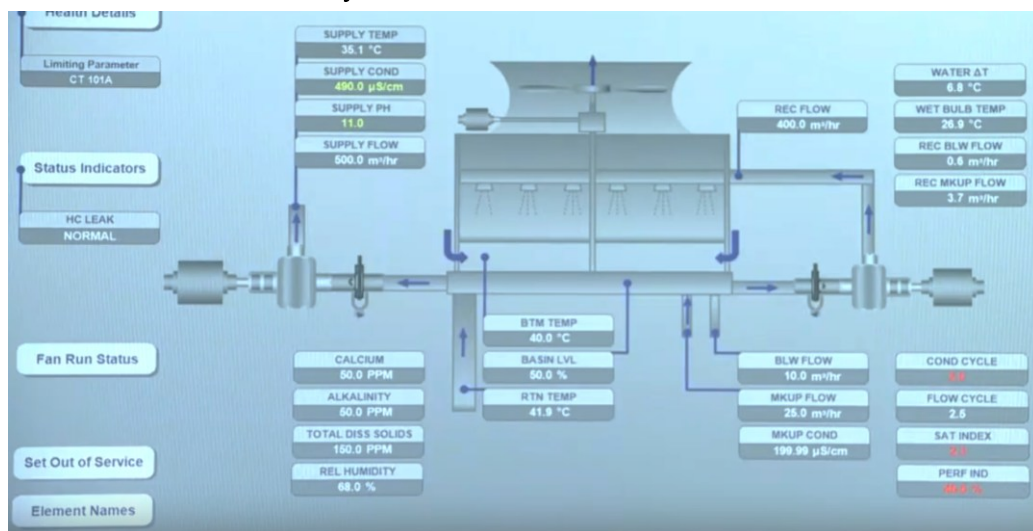


Figure 2. Diagram of a cooling tower with monitoring points

The key result is the formation of predictive scenarios. The flushing optimization algorithm does not provide the staff with a calendar schedule, but with a reasoned recommendation based on the forecast of achieving the DRC, which reduces the number of flushing cycles by 15-25% without reducing the efficiency of heat exchange.

The forecast of the mechanical condition generates warnings about the need to replace the bearing 200-400 hours before the predicted failure, transforming a potentially emergency stop into scheduled repairs. This reduces the cost of urgent spare parts purchases and minimizes downtime. The total economic effect of the implementation for a typical 50 MW cooling tower is manifested in a reduction in operating costs by 10-18% annually due to savings in water, reagents, electricity and repair costs, as well as an increase in equipment availability [4].

The presented methodology for creating a digital twin of a cooling tower demonstrates the effectiveness of using predictive analytics technologies to manage the lifecycle of critical industrial equipment. The integration of physical and mathematical models with operational data flows makes it possible to move from discrete, routine maintenance to continuous condition monitoring and sound predictive planning. Further development is associated with the deepening of models, the use of digital shadows for calibration in conditions of incomplete data, as well as the integration of the cooling tower twin into higher-level digital twins for comprehensive optimization of operating modes.

References

1. Digital doubles of objects in solving control problems / V. A. Minaev, A.V. Mazin, K. B. Zdiruk, L. S. Kulikov // Radio industry. 2019. No. 3. pp. 68-78.
2. Kudryashova, V. A. Digital twins and information technologies in industry / V. A. Kudryashova, A. Yu. Nazarov, M. A. Razakov // Electronic network political journal "Scientific Works of KubSTU". - 2023. – No. 4. – pp. 88-98.
3. Ryzhakov, V. V. digital twins of industrial facilities / V. V. Ryzhakov, O. E. Sizov // Problems of the electric power industry and telecommunications in the North of Russia : Proceedings of the IV International Scientific and Practical Conference, Surgut, April 20-21, 2023. Moscow: Znanie-M Publishing House, 2023, pp. 604-606.
4. Deimundt, A. S. The digital twin of an energy enterprise and its interaction with artificial intelligence modules / A. S. Deimundt // Youth and Science-2023 : Proceedings of the international scientific and practical conference, Petropavlovsk, April 12, 2023. Petropavlovsk: Non-profit Joint-Stock Company "North Kazakhstan University named after Manash Kozybayev", 2023, pp. 224-228.

УДК 004.8

Бугаенко А.П., Тарасова А.П., Интеграция внешних источников данных в процесс генерации текста в больших языковых моделях

Integration of external sources into Large Language Model text generation process

Бугаенко Андрей Павлович,

Магистрант факультета информатики и вычислительной техники,
Национальный исследовательский технологический университет «МИСИС»
Научный руководитель

Тарасова А.П., канд. экон. наук, доцент кафедры управления и права
Поволжский государственный технологический университет

Science adviser: Tarasova A.P., Cand. of Econ. Sc., Asc. Prof. of the Department of Management
and Law,

Volga State Technological University

Bugaenko Andrei Pavlovich,

Graduate Student of the Faculty of Information and Computer Engineering,
National University of Science and Technology «MISIS»

Аннотация. В статье рассматриваются ключевые подходы к интеграции внешних источников данных в процесс генерации текста большими языковыми моделями. Описываются базовые принципы работы, особенности применения и основные вариации RAG, Tool-Use фреймворков и агентных фреймворков. Проводится анализ и сравнение сильных и слабых сторон для каждого подхода.

Ключевые слова: большая языковая модель, большая языковая модель, генерация дополненная извлечением, вызов процедур, автоматические агенты

Abstract. The article discusses key approaches to integrating external data sources into the process of text generation in large language models. The basic principles of operation, features of application and the main variations of RAG, Tool-Use frameworks and agent-based frameworks are described. An analysis and comparison of strengths and weaknesses for each approach is carried out.

Keywords: large language model, large language model, retrieval augmented generation, tool calling, autonomous agent

Рецензент: Мартеха Александр Николаевич – кандидат технических наук, доцент.
ФГБОУ ВО «РГАУ-МСХА им. К.А. Тимирязева»

Введение

После появления в 2018 году ChatGPT [1] большие языковые модели (LLM) перестали быть просто научным инструментом. С каждым годом растет количество пользователей и вариаций и новых версий LLM, а также способов их применения для достижения какого-либо полезного результата [2]. Большую часть времени LLM могли работать только с текстом и имели ограничения связанные механизмами их работы – зависимость от обучающей выборки, а также размера и способа работы со своим контекстным окном.

Традиционным решением описанных выше проблем являлось увеличение размера модели – как количества параметров, так и контекстного окна. Однако такой подход приводит к значительному увеличению требований к количеству обучающих данных, ресурсам на обучение и инференс модели. С помощью изменений в архитектуре можно достичь сопоставимого результата [3], однако они все равно не до конца решают описанные выше проблемы.

В результате начали возникать способы повышения эффективности решения задач LLM, основанные на предоставлении механизмов для взаимодействия с внешними источниками и инструментами, что позволяет решать проблему устаревания данных в модели со временем, снижать размер модели за счет выноса данных вне весов, использовать внешние инструменты для выполнения задач не связанных с генерацией текста. На текущий момент существует три основных метода аугментации генерации LLM, которые будут рассмотрены в этой статье: RAG (Retrieval-Augmented Generation); Tool-Use фреймворки; агентные фреймворки.

Tool-Use фреймворки

RAG (Retrieval-Augmented Generation) является методом, позволяющим включать релевантные данные из внешних источников в результаты генерации LLM. Механизм работы RAG можно разбить на два основных шага – получение данных (Retrieve) и аугментированная генерация (Augmented-Generation). На этапе получения данных проводится поиск наиболее подходящей под запрос информации в базе документов. После чего на этапе аугментированной генерации текст запроса и полученная информация комбинируется для генерации ответа модели [4] Ниже кратко описаны основные вариации развития базового подхода RAG.

Self-RAG – отличается от базового RAG тем, что с помощью ввода специальных токенов (reflection tokens), для генерации которых модель дообучается оценивать необходимость использования источника в тексте и выбирать наилучший дополненный данными результат из сгенерированных [5].

Активный RAG (FLARE) – данный подход генерирует временное следующее предложение в ответе для того, чтобы оценить его на необходимость использования базы знаний для получения актуальной информации [6].

Агентский RAG – данный подход отличается включением активных агентов, способных к динамическому принятию решений, в процесс поиска релевантной информации, что позволяет повысить качество ответа за счет итеративного формирования ответа [7]

Мультимодальный RAG – данный подход расширяет базовый RAG за счет обучения модели использованию баз данных которые содержат не только текстовую информацию, но и изображения, звук и прочие модальности [8].

Графовый RAG – данный подход использует LLM для того, чтобы создать граф знаний, где вершины соответствуют ключевым вхождениям в базе знаний, а ребра соответствуют семантическим отношениями между этими корпусами. Созданный граф разделяется на подгруппы по смысловой близости вершин, смысловое содержание которых обобщается с помощью LLM, и затем используется при формировании ответа [9].

HyDE (Hypithetical Document Embeddings) – данный подход вместо поиска информации по базе данных на основе чистого запроса от пользователя сначала генерирует предполагаемый ответ или документ. После чего векторный эмбединг сгенерированного текста используется для поиска на основе метрик схожести векторов. [10].

REALM (Retrieval-Augmented Language Model) – в данном подходе этап получения данных интегрирован напрямую в обучаемую языковую модель с помощью использования масок. Внедрение извлечения данных в архитектуру позволяет повысить эффективность работы модели за счет того, что текст с самого начала генерируется с учетом необходимости использования источников [11].

RETRO (Retrieval-Enhanced Transformer) – данный подход модифицирует стандартную архитектуру трансформера, добавляя к ней базу данных в форме большого количества токенов, которые находятся в замороженном состоянии и добавляются в ответ по мере генерации [12].

RAG находит широкое применение во многих задачах обработки текста, где важно качество и точность получаемых данных. Это подтверждается тем, что RAG системы широко используются в QA системах, связанных с финансами, медициной и технической информацией – областями, требующими повышенной точности при генерации ответов [13].

Из потенциально отрицательных сторон RAG выделяют необходимость в создании внешней зависимости от программного интерфейса приложения (API) или базы данных, что приводит к отдельным проблемам связанных с их корректной настройкой и поддержкой работы помимо развертывания LLM. Также наличие внешней API зависимости задач которой является поиск данных означает возможность возникновения существенной задержки, что усложняет интеграцию RAG систем в высоконагруженные сервисы.

Tool-Use фреймворки

Tool-Use фреймворк является способом организации архитектуры LLM и её окружения таким образом, что последняя может программно взаимодействовать с предоставляемыми фреймворком функциями, внешними инструментами и API. Основной Tool-Use архитектуры является наличие мыслительной петли, в которой LLM работает как центр принятия решений, определяющий когда и как использовать набор внешних инструментов, доступных в её окружении. В соответствии с принятыми решениями выбранный инструмент используется с заданными параметрами, после чего на основе результата его работы формируется конечный ответ модели на запрос. В результате Tool-Use позволяет использовать LLM не как статический репозиторий с информацией, а динамический интерфейс, способный выполнять заданный набор действий в цифровой среде [14-16]. В общем случае в процессе Tool-Use выделяют четыре основных этапа:

1. Планирование задач – декомпозиция изначальной задачи на набор подзадач с помощью LLM.
2. Выбор инструментов – поиск подходящих инструментов среди доступных с помощью LLM или RAG-системы.
3. Использование инструментов – вызов выбранных инструментов с помощью Tool сервера.
4. Генерация ответа на запрос – сбор и суммаризация данных полученных в результате выполнения инструментов для формирования ответа на запрос пользователя.

Подход Tool-Use является относительно новым направлением в применении LLM для задач помимо генерации текста. Не смотря на это существует множество вариаций этого подхода для множества инструментов, каждый из которых потенциально уникален для среды и ситуации применения метода [15]. За счет относительно абстрактного подхода к вызову инструментов Tool-Use фреймворки позволяют использовать широкий набор внешних инструментов для получения необходимых актуальных данных и последующих действий над ними, если это необходимо. При этом Tool-Use в среднем проще и быстрее чем RAG, поскольку не требуется создания специализированной внешней базы знаний и пайплайна для интеграции информации из него в ответ модели, вместо этого вызывается внешний инструмент. Также за счет использования внешнего инструмента результаты его работы являются относительно детерминированными, так как представляются из себя структурированный результат вызова внешнего API. При этом необходимо учитывать, что Tool-Use ограничен в

понимании контекста, так как не использует эмбединг ответов в запросе аналогично тому как это делается в RAG [14-16].

Агентские фреймворки

Агентный фреймворк представляет из себя систему, которая использует LLM для того, чтобы на основе поставленной сложной задачи составлять план действий и выполнять набор действий по этому плану для достижения поставленного результата [17-19]. В отличие от стандартного взаимодействия в форме запрос-ответ агент может генерировать и обрабатывать множество запросов и ответов для решения задачи в цикле пока задача не будет решена. Также в рамках этого цикла агент может взаимодействовать с набором внешних инструментов и источников данных [17]. Принцип работы агента следующий:

1. Планирование – LLM обрабатывает сложную цель, поставленную пользователем, с целью разделения её на набор простых подзадач.
2. Выполнение – агент выполняет каждый поставленный шаг с помощью выбора и использования инструментов.
3. Наблюдение – агент получает результат работы от инструмента.
4. Повтор – LLM проверяет новое состояние с учетом полученных результатов, и если оно удовлетворяет поставленной задаче - заканчивает выполнение, иначе модифицирует план и начинает цикл с начала.

В общем случае агенты понятие очень широкое, поскольку охватывает все системы, где LLM используется для декомпозиции и планирования задач с последующей передачей их на выполнение внешним агентам в форме других LLM или инструментов [17]. Среди общей массы по похожему функционалу выделяют следующие типы агентов:

1. Общие – общий вид агента, не заточенный на выполнение каких-либо конкретных задач.
2. Инструментальные – агент, обученный управлению набором инструментов для решения узкоспециализированной задачи.
3. Симуляционные – агент, существующий и автономно действующий в некой симулированной среде, обычно используется не напрямую для решения какой-либо задачи, а для извлечения и анализа паттернов поведения формирующихся под воздействием среды.
4. Воплощенные (embodied) – агент, управляющий каким-либо телесным воплощением в виртуальной или реальной среде.

5. Игровые – агент, заточенный на существование и деятельность в какой-либо игре либо для прохождения, либо с целью предоставления реалистичного опыта для игроков.
6. Web – агент, основной задачей которого является автоматизация задач взаимодействия с веб-технологиями.
7. Ассистенты – агент, созданный для выполнения роли секретаря для одного или нескольких пользователей.

Помимо разделения агентов по функционалу также существует разделение по реализации поскольку создать систему-агента можно несколькими разными способами. Среди подходов к реализации агентов выделяют следующие группы:

1. Action Execution (LangChain Tools, LlamaIndex) – агент использует одну или несколько LLM для выбора и использования одного или нескольких доступных ей инструментов для выполнения запроса пользователя.
2. Plan and Execute (BabyAGI, AutoGPT) – агент оперирует в автономном режиме в цикле Размышление-Действие-Наблюдение для планирования и выполнения декомпозированной задачи, заданной пользователем.
3. Multi-Agent (CrewAI, AutoGen) – представляет из себя ансамбль агентов, коммуницирующих и делегирующих между собой задачи для достижения поставленной цели.

Использование агентов позволяет автономно решать сложные задачи, в том числе по поиску и обработке информации, без человеческого вмешательства. Также появляется возможность изменения количества, качества и прочих параметров набора инструментов, используемых моделью, для повышения эффективности решаемой задачи [17-19]. При этом стоит учитывать, что использование агентов ведет к нескольким последовательным вызовам LLM вместо одного, что делает скорость ответа меньше, а стоимость использования таких систем больше. Также использование внешних инструментов помимо описанных выше преимуществ влечет за собой определенную непредсказуемость в поведении модели и сложность в поиске причины возникающих проблем [17-19].

Сравнение подходов

Выбор между RAG, Tool-Use и Agent Frameworks в качестве инструмента для аугментации генерации LLM с помощью внешних источников в основном является решением, основанным на требованиях к задаче и ограничениях ресурсов.

RAG в силу своих свойств предоставляет качественные результаты при работе со статической, контролируемой базой знаний. Его использование позволяет уменьшить

галлюцинации для четко определенных корпусов. В отличие от них, Tool-Use и агентные фреймворки подключаются к динамическим данным в реальном времени, что это создает зависимость от надежности и задержки внешнего API.

RAG является наименее сложной в реализации и работает как единый, усовершенствованный этап генерации. Tool-Use фреймворк добавляет уровень планирования и выполнения. Агентные фреймворки являются наиболее сложными, требующими сложных циклов рассуждений и планирования, но они обеспечивают максимальную автономию для решения открытых целей.

Tool-Use предназначен для обеспечения точности – вызова калькулятора для математики или базы данных для конкретного запроса. RAG предоставляет более широкие, контекстуальные знания из набора документов. Агенты организуют и то, и другое, используя инструменты для обеспечения точности и извлечения контекста, как у RAG, для достижения широкой цели.

Заключение

В большинстве случаев RAG будет являться оптимальным решением для расширения LLM с помощью информации из внешних источников. Применение Tool-Use фреймворкой имеет смысл, когда точность конкретных данных в режиме реального времени имеет первостепенное значение. Агентные фреймворки предназначены для решения сложных многоступенчатых задач по получению и обработке данных в автономном режиме, требующих одновременной работы нескольких инструментов, однако необходимо учитывать, что они требуют наиболее высоких затрат на разработку и вычисления.

Библиографический список

1. T. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan D, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, V. Herbert Ariel, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei. Language Models are Few-Shot Learners // Advances in Neural Information Processing Systems. – 2020. № 33. – С. 1877-1901.
2. M. Nestor, F. Loredana, P. Raymond, G. Yolanda, P. Vanessa, K. Njenga, C. Emily, R. Anka, B. Erik, E. John, L. Katrina, L. Terah, M. James, C. Juan Niebles, S. Yoav, W. Russell, W. Tobi, H. Armin, S. Lapo, B. Julia Lotufo, R. Alexandra, A. Shi, O. Sukrut. The AI Index 2025 Annual Report // AI Index Steering Committee, Institute for Human-Centered AI. – Stanford, CA. – 2025.

3. A. Liu, B. Feng, B. Xue, B. Wang, B. Wu, C. Lu, C. Zhao, C. Deng, C. Zhang, C. Ruan, D. Dai, D. Guo, D. Yang, D. Chen, D. Ji, E. Li, F. Lin, F. Dai, and F. Luo ... Z. Pan. DeepSeek-V3 Technical Report // ArXiv. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/abs/2412.19437> (дата обращения 8.11.2025)
4. P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-tau Yih, T. Rocktäschel, S. Riedel, D. Kiela. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks // ArXiv – 2020. [Электронный ресурс]. URL: <https://arxiv.org/abs/2005.11401> (дата обращения: 1.11.2025)
5. A. Asai, Z. Wu, Y. Wang, A. Sil, H. Hajishirzi. Self-RAG: Learning to Retrieve, Generate, and Critique through Self-Reflection // ArXiv. – 2023. [Электронный ресурс]. URL: <https://arxiv.org/abs/2310.11511> (дата обращения: 3.11.2025)
6. Z. Jiang, F.F. Xu, L. Gao, Z. Sun, Q. Liu, J. Dwivedi-Yu, Y. Yang, J. Callan, G. Neubig. Active Retrieval Augmented Generation // ArXiv. – 2023. [Электронный ресурс]. URL: <https://arxiv.org/abs/2305.06983> (дата обращения: 1.11.2025)
7. A. Singh, A. Ehtesham, S. Kumar, T. Talaei Khoei. Agentic Retrieval-Augmented Generation: A Survey on Agentic RAG // ArXiv. – 2025. [Электронный ресурс]. URL: <https://arxiv.org/abs/2501.09136> (дата обращения: 5.11.2025)
8. Peng Xia, Kangyu Zhu, Haoran Li, Tianze Wang, Weijia Shi, Sheng Wang, Linjun Zhang, James Zou, Huaxiu Yao. MMed-RAG: Versatile Multimodal RAG System for Medical Vision Language Models // ArXiv. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/abs/2410.13085> (дата обращения: 3.11.2025)
9. D. Edge, H., N. Cheng, J. Bradley, A. Chao, A. Mody, S. Truitt, D. Metropolitansky, R. Osazuwa Ness, J. Larson. From Local to Global: A Graph RAG Approach to Query-Focused Summarization // ArXiv. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/abs/2404.16130> (дата обращения: 8.11.2025)
10. L. Gao, X. Ma, J. Lin, J. Callan. Precise Zero-Shot Dense Retrieval without Relevance Labels // ArXiv. – 2022. [Электронный ресурс]. URL: <https://arxiv.org/abs/2212.10496> (дата обращения: 7.11.2025)
11. K. Guu, K. Lee, Z. Tung, P. Pasupat, M.W. Chang. REALM: Retrieval-Augmented Language Model Pre-Training // ArXiv. – 2020. [Электронный ресурс]. URL: <https://arxiv.org/abs/2002.08909> (дата обращения: 6.11.2025)
12. S. Borgeaud, A. Mensch, J. Hoffmann, T. Cai, E. Rutherford, K. Millican, G. van den Driessche, J.B. Lespiau, B. Damoc, A. Clark, D. Casas, A. Guy, J. Menick, R. Ring, T. Hennigan, S. Huang, L. Maggiore, C. Jones, A. Cassirer, A. Brock, M. Paganini, G. Irving, O. Vinyals, S. Osindero, K. Simonyan, J.W. Rae, E. Elsen, L. Sifre. Improving language models by

retrieving from trillions of tokens // ArXiv. – 2021. [Электронный ресурс]. URL: <https://arxiv.org/abs/2112.04426> (дата обращения: 7.11.2025)

13. M. Arslan, H. Ghanem, S. Munawar, C. Cruz. A Survey on RAG with LLMs // Procedia Computer Science Volume. – 2024. – Т. 246. – С. 3781-3790.

14. X. Weikai, H. Chengrui, G. Shen, S. Shuo. LLM-Based Agents for Tool Learning: A Survey // Data Science and Engineering. – 2025. – С. 2364-1541.

15. C. Qu, S. Dai, X. Wei, H. Cai, S. Wang, D. Yin, J. Xu, J.R. Wen. Tool Learning with Large Language Models: A Survey // ArXiv. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/abs/2405.17935> (дата обращения: 3.11.2025)

16. Z. Shen. LLM With Tools: A Survey // ArXiv. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/abs/2409.18807> (дата обращения: 8.11.2025)

17. L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X., Y. Lin, W. Xin Zhao, Z. Wei, J.R. Wen. A Survey on Large Language Model based Autonomous Agents // ArXiv. – 2023. [Электронный ресурс]. URL: <https://arxiv.org/abs/2308.11432> (дата обращения: 8.11.2025)

18. J. Luo, W. Zhang, Y. Yuan, Y. Zhao, J. Yang, Y. Gu, B. Wu, B. Chen, Z. Qiao, Q. Long, R. Tu, X. Luo, W. Ju, Z. Xiao, Y. Wang, M. Xiao, C. Liu, J. Yuan, S. Zhang, Y. Jin, F. Zhang, X. Wu, H. Zhao, D. Tao, P.S. Yu, M. Zhang. Large Language Model Agent: A Survey on Methodology, Applications and Challenges // ArXiv. – 2025. [Электронный ресурс]. URL: <https://arxiv.org/abs/2503.21460> (дата обращения: 3.11.2025)

19. T. Guo, X. Chen, Y. Wang, R. Chang, S. Pei, N.V. Chawla, O. Wiest, X. Zhang. Large Language Model based Multi-Agents: A Survey of Progress and Challenges // ArXiv. – 2024. [Электронный ресурс]. URL: <https://arxiv.org/abs/2402.01680> (дата обращения: 3.11.2025)

УДК 004.896:620.11

Литвинова А.В., Чернова Д.В. Машинное обучение и разработка эффективных промышленных материалов

Machine learning and the development of efficient industrial materials

Литвинова Александра Владимировна,

Ст. преподаватель кафедры промышленного дизайна и медиатехнологий,
Санкт-Петербургский государственный университет промышленных технологий и дизайна.

Высшая школа технологии и энергетики

Чернова Дарья Вячеславовна,

Студент кафедры промышленного дизайна и медиатехнологий,
Санкт-Петербургский государственный университет промышленных технологий и дизайна.

Высшая школа технологии и энергетики

Litvinova Alexandra Vladimirovna

Senior Lecturer of the Department of Industrial Design and Media Technologies

St. Petersburg State University of Industrial Technologies and Design.

Higher School of Technology and Energy

Darya Vyacheslavovna Chernova,

Student of the Department of Industrial Design and Media Technologies,

St. Petersburg State University of Industrial Technologies and Design.

Higher School of Technology and Energy

Abstract. The article is dedicated to studying the role of machine learning (ML) in the selection of materials for engineering and industrial tasks. The advantages of ML approaches are discussed, such as increased prediction accuracy, reduction of time and material costs, elimination of errors, and improvement of product performance characteristics. In addition, the main challenges hindering the large-scale implementation of ML in materials engineering are outlined: lack of historical data, complexity of interpreting model outputs, and high computational resource requirements.

Keywords: machine learning, materials selection, prediction accuracy, computational resources, historical database, industrial technologies.

Аннотация. Статья посвящена изучению роли машинного обучения (ML) в подборе материалов для инженерных и промышленных задач. Рассматриваются преимущества ML-подходов, такие как повышенная точность прогнозирования, сокращение временных и материальных затрат, устранение ошибок и улучшение эксплуатационных характеристик изделий. Помимо этого обозначены главные проблемы, затрудняющие широкомасштабное внедрение ML в инженерию материалов: недостаток исторических данных, сложность интерпретации выводов моделей и высокие требования к вычислительным ресурсам.

Ключевые слова: машинное обучение, подбор материалов, точность прогнозирования, вычислительные ресурсы, историческая база данных, промышленные технологии.

Рецензент: Мартеха Александр Николаевич – кандидат технических наук, доцент.

ФГБОУ ВО «РГАУ-МСХА им. К.А. Тимирязева»

Правильность выбора материалов для конкретных инженерных и конструкторских задач приобрела первостепенную важность. От правильности выбора зависят ключевые показатели изделия: прочность, вес, износостойчивость,

безопасность и, конечно же, цена. Классические подходы к решению данной задачи основывались преимущественно на опыте и интуиции инженера, однако подобные подходы имели существенные недостатки: долгий цикл тестирования, отсутствие учета множества факторов одновременно и значительные материальные затраты.

С развитием технологий машинного обучения появилась возможность кардинально изменить ситуацию. Теперь стало возможным быстро и эффективно обрабатывать огромные объемы данных, получать точные прогнозы и рекомендации по материалам, минимизируя человеческий фактор и увеличивая скорость выхода готовых изделий на рынок.

Инструменты машинного обучения представляют собой совокупность специализированных алгоритмов и методов, предназначенных для автоматического извлечения полезных закономерностей из данных. Они позволяют выявить зависимости, провести классификацию и сделать прогнозы, значительно упрощая многие рутинные задачи, стоящие перед специалистами разных областей [1].

Среди основных инструментов машинного обучения выделяется линейная и логистическая регрессия, которая используется для выявления прямой зависимости между переменными. Она широко применяется в материаловедении для описания зависимостей между различными характеристиками материала (например, плотностью, прочностью, температурой плавления) и итоговыми техническими параметрами изделия (вес, сопротивление нагрузкам). Логистическая регрессия подходит для бинарных ситуаций («пригоден-непригоден»), что полезно при оценке соответствия материала заявленным характеристикам и техническим требованиям.

Следующий инструмент машинного обучения – деревья решений и случайные леса. Они применяются для разделения набора данных на классы согласно определенным правилам. Каждый узел дерева задаёт условие, которое помогает принять решение о принадлежности элемента к определенному классу. Данный инструмент особенно полезен при выборе материалов, поскольку позволяет систематизировать многочисленные признаки и учитывать особенности внешней среды и режимов эксплуатации. Случайные леса (Random Forests) объединяют несколько деревьев решений, каждое из которых построено на случайно выбранных подмножествах признаков исходных данных. Такой подход повышает стабильность и точность моделей, предотвращая переобучение.

Нейронные сети как инструмент имитируют структуру нервной системы человека, состоящую из отдельных узлов (нейронов), соединенных друг с другом. Простейшие нейронные сети (однослойные перцептроны) подходят для простых задач, тогда как

многослойные нейронные сети обладают большими возможностями и могут находить сложные нелинейные зависимости. Нейронные сети часто используют для моделирования деформаций, разрушения материалов и расчета структурных напряжений. Их способность обобщать позволяет строить надежные прогностические модели даже при наличии неполных данных.

Глубокие нейронные сети отличаются наличием многих слоев и возможностью автоматически извлекать иерархические признаки из сырых данных. Эти сети демонстрируют отличные результаты в задачах, связанных с обработкой изображений, звуком и текстом, что актуально при анализе микроструктур материалов и диагностике дефектов. Глубокая сверточная нейронная сеть способна проанализировать фотографии поверхности материала и выявить дефекты или аномалии, невидимые человеческому глазу. Подобные технологии активно внедряются в автомобилестроении и авиации для контроля качества продукции.

Существуют методы, которые группируют объекты по сходству признаков, выделяя группы схожих материалов или элементов. Такие методы полезны при сортировке данных и предварительной обработке перед построением моделей [2].

Классификация направлена на разделение объектов на заранее известные категории; это удобно при отборе материалов по установленным спецификациям. Типичным примером классификации является выделение подходящих марок стали для строительства мостов, учитывая такие параметры, как предел прочности, вязкость и химический состав.

Удачным считается подход генеративных состязательных сетей (GAN). Эта группа моделей состоит из двух конкурирующих частей: генератора и дискриминатора. Генератор учится создавать реалистичные образцы, пытаясь обмануть дискриминатор, который стремится отличить реальные образцы от поддельных. Подобный подход эффективен при генерации синтетических образцов материалов, ранее неизвестных, однако потенциально выгодных для использования.

Применение машинного обучения в инженерии и промышленном дизайне потенциально выгодно. Оно сокращает время на принятие решений о материале, повышает качества прогнозов по свойствам материалов, оптимизирует расходы ресурсов и снижение общих затрат, уменьшает число ошибок и дефектов.

Примером успешного применения машинного обучения стал случай компании Airbus, которая с помощью нейронных сетей смогла выбрать сверхлегкий углепластиковый композит для фюзеляжа пассажирского лайнера A350XWB, что снизило общую массу самолета на 2 тонны и увеличило дальность полета.

В целом, машинное обучение открыло новые горизонты в сфере подбора материалов для различных сфер промышленности, строительства и технологий. Стоит рассмотреть ключевые преимущества использования машинного обучения в подборе материалов.

Одно из важнейших преимуществ машинного обучения — это возможность точного прогнозирования свойств материалов и их поведения в различных условиях эксплуатации. Традиционные методы часто основываются на ограниченном объеме данных и личном опыте эксперта, что может привести к ошибкам и неправильным решениям. Машинное обучение позволяет оценивать гораздо большее количество переменных и обнаруживать скрытые закономерности, недоступные простым наблюдениям [2;3].

Например, компания Rolls-Royce применяла алгоритмы машинного обучения для подбора материалов лопаток турбин реактивных двигателей. После тщательного анализа множества факторов, включая температуру, механические напряжения и усталость металла, специалисты смогли разработать новый композитный материал, который обеспечил прирост ресурса двигателей на 20% и снизил потребление топлива.

Еще одно важное преимущество машинного обучения — существенное сокращение времени и затрат на проведение тестов и эксперименты. Ранее для проверки материала требовалось проводить лабораторные испытания, что было дорого и долго. Машинное обучение позволяет быстро моделировать различные сценарии и подбирать оптимальный материал, минуя этапы дорогостоящего тестирования. К примеру, американская компания Ford Motor Company применила машинное обучение для быстрого подбора материалов кузовных панелей автомобиля Mustang. В ходе анализа оказалось, что замена обычного алюминия на специально разработанный сплав позволила снизить вес автомобиля на 10 кг, что привело к увеличению динамики разгона и экономии топлива.

Алгоритмы машинного обучения минимизируют вероятность ошибок, возникающих из-за субъективных предпочтений и ограничений человеческих способностей. Подбор материалов, основанный на результатах анализа огромных объемов данных, гарантирует высокое качество продукции и минимальное количество отказов. Например, крупнейший производитель самолетов Boeing использовал машинное обучение для подбора подходящего сплава для крыльев коммерческих авиалайнеров. Программа сравнила сотни различных сплавов, учла нагрузку крыла, климатические условия эксплуатации и физические характеристики материала,

подобрал оптимальное решение, которое повысило прочность конструкции и сократило количество аварийных происшествий.

Также машинное обучение позволяет индивидуально подходить к каждому проекту, учитывая конкретные нужды клиента и внешние условия эксплуатации. Например, для одних клиентов важны высокие показатели теплоустойчивости, другим необходим легкий вес, третьим важна низкая стоимость. Алгоритм может быстро подобрать подходящий материал, балансируя эти запросы. Концерн BMW разработал гибридный автомобиль i3, оснащённый батареями с оптимальным материалом катода. Алгоритм машинного обучения подобрал такое сочетание материалов, которое позволило добиться максимальной емкости аккумулятора при минимальной массе, что сделало автомобиль лидером по запасу хода среди электрических авто премиум-класса.

Возможность учитывать экологические факторы при выборе материалов – один из положительных аспектов машинного обучения является. Например, алгоритм может выбрать наименее токсичный материал или произвести расчеты, направленные на минимизацию выбросов углекислого газа при производстве. Например, Nike использовала машинное обучение для выбора экологичного материала обуви. Компания получила подробную оценку по нескольким параметрам, включая воздействие на окружающую среду и углеродный след. В результате Nike представила коллекцию спортивной обуви с минимальным воздействием на природу, получив признание от организаций защиты экологии.

Алгоритмы машинного обучения работают и с огромными наборами данных, создавая уникальную базу знаний о поведении материалов в различных условиях. Это даёт возможность специалистам оперативно реагировать на изменения требований рынка и осваивать новые рынки сбыта. К примеру, китайская компания Huawei применила машинное обучение для анализа сотен материалов корпусов смартфонов. Новый материал, подобранный алгоритмом, оказался устойчивым к повреждениям, влагозащитным и сохранял отличную передачу сигнала. Благодаря этому смартфоны Huawei получили высокую оценку потребителей и заняли лидирующую позицию на рынке мобильных телефонов.

Рассмотрев преимущества, можно сказать, что машинное обучение в подборе материалов играет важную роль для создания оптимального решения, которое будет отвечать по всем требованиям надёжность, получит весомое конкурентное преимущество.

Машинное обучение продолжает демонстрировать выдающиеся результаты в различных областях, становясь незаменимым инструментом для компаний и исследователей. Далее представлены яркие примеры успешного применения ML в бизнесе, медицине, транспорте и других секторах, подробно раскрывающие механизмы успеха и ценность этих решений.

Если рассмотреть другие сферы применения машинного обучения, то стоит отметить ее влияние на медицину. Так один из самых наглядных примеров – диагностика злокачественных новообразований кожи. Группа ученых из Стэнфордского университета создала алгоритм на основе свёрточных нейронных сетей (CNN), обучив его на обширном наборе изображений родинок и опухолей. Результат превзошёл ожидание: программа оказалась столь же эффективной, как и опытные дерматологи, диагностируя меланому с точностью до 95%.

Уже как несколько лет люди по всему миру поражаются беспилотными автомобилями, а ведь это всё алгоритмы машинного обучения, которые непрерывно анализируют окружающее пространство, принимая решения о торможении, повороте, перестроении и управлении скоростью. Так, Tesla Motors сделала огромный шаг вперёд в развитии беспилотных автомобилей, представив технологию Autopilot. Постоянное обновление моделей позволяет Tesla улучшать алгоритмы на основании опыта всех владельцев, собирая статистику и исправляя ошибки. Автопроизводителю удалось резко снизить количество ДТП, вызванных человеческими ошибками, сделав вождение безопасным и комфортным

СберБанк России реализовал крупный проект по внедрению машинного обучения для выдачи кредитов малым и средним предпринимателям. В последние годы Сбербанк России сосредоточился на внедрении передовых технологий машинного обучения (ML) для повышения эффективности работы с клиентами малого и среднего бизнеса (МСБ). Главной задачей было создание надежного и быстрого кредитного конвейера, способного обеспечивать точное принятие решений по выдаче займов, минимизируя кредитные риски и повышая лояльность клиентов.

До внедрения ML-кредитования работники Сбербанка пользовались стандартными системами скоринга, анализирующими небольшой набор данных заемщика (доход, возраст, семейное положение и т.д.). Но такой подход имел много минусов: устаревшие данные, субъективность сотрудников, невозможность учета нестандартных обстоятельств, влияющих на платежеспособность предпринимателя.

Решение пришло с появлением новых технологий. Сбербанк начал собирать и обрабатывать огромный объем данных о каждом клиенте МСБ, используя внутренние и

сторонние источники. К примеру, помимо кредитной истории анализировались бухгалтерская отчетность, налоговые декларации, обороты по счету, отраслевые индексы, региональные особенности, динамика доходов, расходы на персонал и аренду, платежи контрагентам и многое другое. Помимо анализа данных о компаниях, Сбербанк внедрил возможность прямого подключения банковских счетов клиентов к системе оценки. В режиме реального времени происходит мониторинг оборотов, платежей и прочих финансовых потоков, что дополнительно повышает точность оценки финансового здоровья компании.

Уже спустя полгода после полного запуска системы Сбербанк получил первые позитивные результаты:

- Зафиксировано снижение дефолта по кредитам на 15%. Это произошло благодаря своевременному выявлению неблагополучных заемщиков и повышению точности принятия решений.
- Процесс подачи заявки и одобрения займа сократился с нескольких рабочих дней до считанных часов. В особо срочных случаях клиенту предоставляется кредит за полчаса.
- Через год после внедрения ML-алгоритма количество одобренных кредитов выросло на 20%. Это связано с тем, что появились возможности одобрять кредиты там, где раньше принимался отрицательный вердикт из-за недостатка информации.
- Согласно опросам, клиенты стали чаще отмечать простоту и удобство получения кредита в банке, доверяя его автоматизированной системе оценки.

Опыт Сбербанка свидетельствует о высокой эффективности применения машинного обучения в банковской сфере. Технология позволяет значительно повысить точность оценки рисков, сократить операционные расходы и увеличить количество обслуживаемых клиентов. Важно отметить, что подобный подход помогает банкам перейти от статичной схемы анализа заемщиков к активной работе с большим количеством переменных, что улучшает качество предоставляемых услуг и укрепляет позиции на рынке.

Также российская компания «Яндекс» давно осознала потенциал машинного обучения и активно внедряет его во многие свои сервисы, улучшая качество взаимодействия с пользователями и расширяя собственные возможности. Основной упор сделан на персонализацию услуг, повышение релевантности поиска и продвижение рекламных кампаний.

Поисковая система Яндекс всегда находилась в авангарде борьбы за первенство на российском рынке поиска. Одним из первых шагов к доминированию стала реализация собственного поисковика, который был настроен на русский язык и понимал запросы пользователей лучше зарубежных аналогов. Позже Яндекс перешел к глубокой персонализации поиска, когда каждый запрос пользователя привязывается к профилю его предпочтений.

Что изменилось с приходом машинного обучения: раньше результаты поиска выстраивались по классической формуле ранжирования документов (PageRank, BM25 и другие алгоритмы). Теперь же поисковая система дополняется моделями машинного обучения, которые анализируют поведенческие сигналы пользователя (клики, скроллинг, возврат к поиску), личные предпочтения и даже эмоциональную окраску текста. Пользователю предлагаются персонализированные результаты, что делает поиск более удобным и быстрым. Например, если пользователь ищет рецепт блюда и недавно смотрел кулинарные шоу, система покажет рецепты с учётом любимых ингредиентов и кухни.

Пользователей интернета беспокоит не только качество поиска, но и защита от мошенничества и опасных сайтов. Яндекс внедрил в свои сервисы алгоритмы машинного обучения для выявления угроз и блокировки нежелательной информации. Например, антиспам-система анализирует письма электронной почты, фильтруя их по признакам мошенничества (например, использование шаблонных фраз, фальшивых адресов отправителей, сомнительные вложения).

Также Яндекс активно развивает рекламный бизнес, предлагая бизнесу возможность размещать контекстную рекламу. Система "Директ" использует машинное обучение для подбора целевой аудитории, настраивая рекламную кампанию на нужную аудиторию. Кроме того, предлагается отдельная услуга "Яндекс Аудитории", где с помощью ML проводится сегментация аудитории и подготовка профилей пользователей для размещения рекламы.

Дополнительно Яндекс предложил своим партнёрам и рекламодателям уникальное решение — технологию Look-alike, которая на основе данных о существующих покупателях находит похожих пользователей, тем самым увеличивая конверсию рекламных кампаний.

Яндекс развивается, так, голосовой помощник Алиса стал ярким примером интеграции машинного обучения в повседневную жизнь россиян. Эта система позволяет общаться с компьютером естественным языком, запрашивая помощь в выполнении повседневных задач, таких как заказ такси, поиск рецепта или

консультации по расписанию общественного транспорта. Особенностью Алисы является постоянное обучение на диалогах с пользователями. Чем больше общается Алиса, тем умнее она становится. Её ответы становятся понятнее, точнее и индивидуальнее. Алгоритмы машинного обучения анализируют каждое общение, выделяя шаблоны и создавая универсальный опыт взаимодействия.

Сервис Яндекс.Карты и Навигатор также активно используют машинное обучение. С каждым днём система учёта дорожных ситуаций совершенствуется: загруженность улиц, пробочные зоны, оптимальные маршруты — всё это определяется на основе живых данных с телефона пользователя. Смартфон определяет ваше местоположение и движение, делится информацией о дорожной ситуации, а алгоритм машинного обучения принимает решение, какую дорогу вам предложить.

Применяя ML в своём бизнесе, компания вышла на лидерские позиции в области поиска, рекламы, геоуслуг и персональных ассистентов. При этом не забывает о постоянной доработке алгоритмов, усиливая их новыми данными и функциональностью.

Также стоит рассмотреть Газпром нефть. Это одна из крупнейших российских нефтяных компаний, включила в свою стратегию интенсивное использование технологий машинного обучения для повышения эффективности добычи и переработки углеводородов. Внедрение этих технологий позволило значительно усовершенствовать процессы разведки месторождений, добычу нефти и её последующую переработку.

Нефтяная отрасль традиционно сопряжена с большими рисками и неопределённостями. Цена ошибки в добыче или переработке крайне высока, поэтому важно использовать современные технологии для повышения точности прогнозов и снижения потерь.

Процесс начинается с огромного массива данных, поступающего от приборов и датчиков, установленных на месторождениях и нефтеперерабатывающих заводах. Вся информация консолидируется в единую базу данных, доступную для обработки и анализа.

Алгоритм машинного обучения проходит несколько стадий: сбор данных, подготовка данных, построение модели, оценка и проверка и эксплуатация [4].

Несколько успешных инициатив в Газпромнефти заслуживают отдельного внимания:

1. Операция мониторинга скважины: Используется ML для оценки технического состояния скважин и выявления отклонений от нормативных параметров. Регулярный

мониторинг позволяет заранее планировать техническое обслуживание и ремонт, избегая непредвиденных остановок.

2. Прогноз запасов: Специальные алгоритмы ML позволили компании прогнозировать остаток нефти на конкретном участке месторождения, что помогает оптимизировать программу освоения и рационально распределять ресурсы.

3. Анализ производства на НПЗ: Внедрена система прогнозирования загрузки завода и его энергоэффективности, что существенно снизило энерго- и ресурсопотери [5].

Благодаря правильному управлению процессом добычи и правильной оценке оставшихся запасов нефти, производительность ряда месторождений выросла на 10-15%. Раннее предупреждение о выходе из строя оборудования и правильное распределение ремонтных бригад привели к сокращению простоев и снижению капитальных вложений на ремонт и восстановление. Логистика и доставка нефти оптимизирована на основе ML-предсказаний, что позволило минимизировать запасы и повысить мобильность флота танкеров.

Газпром нефть выступает хорошим примером того, как российские компании могут успешно использовать технологии машинного обучения для повышения эффективности и сокращения затрат.

Машинное обучение активно используется и в Mail.Ru Group (ныне VK Group) – одна из крупнейших российских IT-компаний, владеющей популярными площадками и сервисами, такими как социальные сети ВКонтакте и Одноклассники, почтовая служба Mail.ru, игровой портал My.Games и другие проекты. Компания активно занимается внедрением технологий машинного обучения для улучшения качества своих сервисов, повышения уровня удобства пользования и усиления конкуренции на рынке.

Началом активного внедрения ML в Mail.Ru Group можно считать начало 2010-х годов, когда компания приступила к созданию собственной инфраструктуры для обработки больших данных и анализа пользовательской активности. Первым успешным проектом стало внедрение ML для персонализации ленты новостей социальной сети ВКонтакте. Благодаря обучению на больших объемах данных, алгоритм научился показывать пользователям наиболее интересующие их посты, что увеличило вовлечение аудитории и среднее время пребывания на сайте.

Следующим этапом стало развитие механизмов рекомендаций контента и друзей в социальных сетях. Команда Mail.Ru Group разработала модели, которые анализировали привычки пользователей, круг общения и предпочитаемый контент, создавая персональные рекомендации для подписчиков.

Позже компания перешла к использованию ML для решения задач анти-спама и борьбы с фишингом. В почте Mail.ru был разработан собственный антифрод-инструмент, который ежедневно анализирует миллионы писем, выявляя подозрительные сообщения и защищая пользователей от мошенничества.

Одной из главных задач Mail.Ru Group остаётся защита пользователей от несанкционированного доступа и фишинга. Ежедневно отправляются миллиарды сообщений, значительная часть которых носит мошеннический характер. ML-системы анализируют данные писем, выискивая характерные признаки злоумышленников (определённая структура писем, IP-адреса, регулярность отправки и другие признаки), предотвращая атаки и обеспечивая комфортное пользование сервисом.

Компания также активно использует ML для таргетированной рекламы и персонализации услуг. Сервис myTarget, принадлежащий Mail.Ru Group, собирает и анализирует данные пользователей (интересы, активность, предпочтения), предоставляя рекламодателям возможность точно нацеливаться на целевую аудиторию. Это значительно повышает эффективность рекламных кампаний и увеличивает доход компании.

Социальные сети ВКонтакте и Одноклассники активно пользуются технологиями ML для рекомендаций контента, улучшения функций поиска и чата, персонализации интерфейса и борьбы с дезинформацией. Одно из последних новшеств — внедрение голосового помощника Алиса, умеющего вести осмысленную беседу и помогать пользователям решать бытовые задачи.

Сервисы Delivery Club и Яндекс.Еда (совместный проект Mail.Ru Group и Яндекса) активно применяют ML для прогнозирования времени доставки, оптимизации маршрутов и информирования пользователей о статусе заказа. Благодаря постоянным обновлениям и детализированному анализу данных удаётся сократить время доставки и повысить уровень удовлетворения клиентов.

Сейчас Mail.Ru Group активно проводит исследования в области искусственного интеллекта и машинного обучения, ставя перед собой амбициозные цели:

- Улучшить качество распознавания речи и развивать голосовых ассистентов;
- Совершенствовать алгоритмы антифрода и усилить защиту пользователей;
- Улучшить алгоритмы рекомендаций контента и рекламу, делая их более полезными и интересными для пользователей [1].

Таким образом, Mail.Ru Group является отличным примером успешной интеграции машинного обучения в деятельность крупной компании, что подтвердило

правильность выбранной стратегии и показало большие перспективы для дальнейшего развития.

Эти примеры убедительно свидетельствуют о том, что машинное обучение становится важнейшим драйвером роста и конкурентоспособности в самых разных отраслях. Применение ML в настоящее время находится на начальной стадии, и впереди ждут ещё более интересные открытия и технологии, меняющие привычную жизнь и бизнес навсегда [4].

Хотя машинное обучение обладает огромным потенциалом в области подбора материалов и проектирования, существуют серьезные препятствия, мешающие полному раскрытию его возможностей. Многие организации сталкиваются с нехваткой качественного и достаточного количества данных для обучения моделей машинного обучения. Без полноценных и репрезентативных наборов данных трудно построить надежную модель, способную давать полезные рекомендации. Особенно остро эта проблема проявляется в узкоспециализированных областях, таких как проектирование уникальных материалов для космической промышленности или медицинского оборудования. Когда реальные данные труднодоступны, возможно создание синтетических наборов данных с помощью специальных генераторов, таких как Generative Adversarial Networks (GAN). Хотя синтетические данные не заменят реальный мир, они могут стать полезным дополнением к существующим наборам.

Некоторые алгоритмы машинного обучения, такие как глубокие нейронные сети, сложно объяснить. Почему именно эта комбинация материалов рекомендована системой? Часто остается неясным. Невозможность объяснить выводы препятствует принятию решений и вызывает недоверие со стороны специалистов. Создаются интерпретируемые модели, такие как деревья решений или ансамбли случайных лесов, позволяющие легче проследить ход рассуждений и обосновать финальное решение.

Работа с крупными моделями машинного обучения требует значительных вычислительных ресурсов, которые далеко не каждая организация может позволить себе приобрести или арендовать. Затраты на обучение моделей, обработку больших объемов данных и поддержание инфраструктуры приводят к финансовым проблемам и замедлению темпов внедрения ML. В данном случае использование облачных платформ (Google Cloud, Яндекс Диск) позволяет организациям масштабировать вычислительные мощности в зависимости от потребностей, оплачивая ровно столько, сколько необходимо. Организациям больше не требуется закупать собственное железо.

Различные проблемы, возникающие при внедрении машинного обучения в материаловедение и инженерию, вполне решаемы при грамотном подходе.

Современная наука и технический прогресс предоставляют разнообразные инструменты и подходы для преодоления указанных препятствий, делая машинное обучение доступным и эффективным решением для практических задач.

В будущем мы можем ожидать дальнейшее усиление роли машинного обучения в процессах подбора материалов, так как накапливаются большие наборы данных, улучшается доступность мощных вычислительных ресурсов и повышается зрелость интерпретируемых моделей. Это откроет путь к качественно новому уровню инженерного дела и создаст предпосылки для следующего этапа технологического развития человечества.

Машинное обучение становится ключевым инструментом в совершенствовании методов разработки промышленных материалов, значительно расширяя возможности инженеров и ученых. В данной статье рассмотрены теоретические основы и практические аспекты применения машинного обучения для оптимизации характеристик материалов, повышения их долговечности и снижения экономических затрат. Оно позволяет эффективно обрабатывать огромные объемы данных, выявляя тонкие закономерности и скрытые зависимости, недоступные традиционному анализу. Благодаря машинному обучению создаются инновационные материалы с уникальными физическими и химическими свойствами, способные выдерживать экстремальные условия эксплуатации [4].

Важнейшие задачи, такие как выбор оптимального состава материала, прогнозирование долговечности и анализ усталостных повреждений, успешно решаются с помощью алгоритмов машинного обучения. В перспективе ожидается дальнейшее распространение машинного обучения в промышленности, что приведёт к появлению новых поколений материалов с беспрецедентными свойствами, способствующими устойчивому развитию и росту экономического благосостояния.

Библиографический список

1. Федотова А.В. Искусственный интеллект как фокус развития цифровой экономики: теоретические и практические аспекты // Московский экономический журнал. 2021. № 6. С. 476-487.
2. Антонов А. В. Методы машинного обучения в задачах инженерного анализа материалов / А. В. Антонов, С. В. Иванов // Журнал вычислительной математики и математической физики. — 2019. — Т. 59, № 4. — С. 631–642.
3. Черемисин Д. Г., Мкртчян В. Р. МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ// Журнал Символ Науки. — 2023. — 24 с.

4. Романов И.А. Машинное обучение как конкурентное преимущество предприятия // Московский экономический журнал. – 2022. – №3. – С. 541. \

5. Использование машинного обучения для оптимизации производственных процессов обучения в нефтегазовой отрасли / Ахтямова Д.А., Волкова К.Ю., Максимов С.А., Шайгатдарова А.Р. // Управление образованием: теория и практика. — 2023. — №3.

УДК 004.8:336.71

Сорокин Н.А., Леонова Н.Л. Машинное обучение в задачах анализа временных рядов для прогнозирования тенденций финансового рынка

Machine learning in time series analysis tasks for forecasting financial market trends

Леонова Надежда Львовна,

Ст. преподаватель кафедры прикладной математики и информатики,
Санкт-Петербургский государственный университет промышленных технологий и дизайна.
Высшая школа технологии и энергетики

Сорокин Никита Александрович,

Студент кафедры прикладной математики и информатики,
Санкт-Петербургский государственный университет промышленных технологий и
дизайна. Высшая школа технологии и энергетики

Leonova Nadezhda Lvovna,

Senior Lecturer of the Department of Applied Mathematics and Informatics,
St. Petersburg State University of Industrial Technologies and Design.

Higher School of Technology and Energy

Sorokin Nikita Alexandrovich,

Student of the Department of Applied Mathematics and Informatics,
St. Petersburg State University of Industrial Technologies and Design.

Higher School of Technology and Energy

Аннотация. В статье проводится аналитический обзор применения методов машинного обучения для прогнозирования тенденций на финансовом рынке на основе временных рядов. Рассматриваются ключевые особенности финансовых данных – стационарность, высокая волатильность, шум и нелинейные зависимости, которые затрудняют использование традиционных статистических моделей. Проанализированы современные подходы машинного обучения, включая рекуррентные нейронные сети (LSTM, GRU). Рассмотрены как успешные, так и неудачные кейсы внедрения ИИ в трейдинг, подчеркивающие риски переобучения, неустойчивости сигналов и игнорирования рыночных структурных сдвигов. Показано, что машинное обучение является мощным, но не универсальным инструментом, эффективность которого определяется не столько сложностью модели, сколько качеством данных, методологической строгостью и глубоким пониманием финансовой среды.

Ключевые слова: машинное обучение, временные ряды, прогнозирование тенденций, финансовый рынок, LSTM, финансовые временные ряды.

Abstract. The article provides an analytical review of the application of machine learning methods to predict trends in the financial market based on time series. The key features of financial data are considered – stationarity, high volatility, noise and nonlinear dependencies, which make it difficult to use traditional statistical models. Modern machine learning approaches, including recurrent neural networks (LSTM, GRU), are analyzed. Both successful and unsuccessful cases of AI implementation in trading are considered, emphasizing the risks of retraining, signal instability and ignoring market structural shifts. It is shown that machine learning is a powerful but not universal tool, the effectiveness of which is determined not so much by the complexity of the model as by the quality of the data, methodological rigor and a deep understanding of the financial environment.

Keywords: machine learning, time series, trend forecasting, financial market, LSTM, financial time series.

Рецензент: Мартеха Александр Николаевич – кандидат технических наук, доцент.
ФГБОУ ВО «РГАУ-МСХА им. К.А. Тимирязева»

Современные финансовые рынки представляют собой сложные, динамичные системы, в которых принятие своевременных и обоснованных решений напрямую влияет на эффективность инвестиционной и торговой стратегии. В условиях насыщения рынков огромными объёмами данных и ужесточения конкуренции между участниками особую значимость приобретают методы анализа и прогнозирования. Одной из ключевых задач в этой области является анализ временных рядов – последовательностей наблюдений, упорядоченных во времени, – с целью выявления и прогнозирования рыночных тенденций. Традиционные статистические подходы, такие как модели ARIMA или GARCH, несмотря на свою теоретическую обоснованность, зачастую оказываются недостаточными для адекватного описания нелинейных, нестационарных и шумовых характеристик финансовых данных.

В связи с этим в последние годы все больше распространение получают методы машинного обучения (МО), демонстрирующие высокую гибкость в обработке сложных паттернов и способность адаптироваться к изменяющимся условиям рынка. От классических алгоритмов на основе деревьев решений до современных архитектур глубокого обучения – такие модели открывают новые горизонты в прогнозировании цен, оценке рисков и автоматизации торговых решений. Применение машинного обучения для анализа ценовых движений открывает новые горизонты для трейдеров фьючерсов, особенно в контексте краткосрочной торговли. Технологии машинного обучения предоставляют инструменты для анализа больших объемов исторических данных, выявления закономерностей и предсказания будущих движений цен. Это обеспечивает трейдерам существенное конкурентное преимущество на рынке [1]. Однако, несмотря на многообещающие результаты, применение машинного обучения в финансах сопряжено с рядом фундаментальных вызовов: высоким уровнем шума, склонностью к переобучению, проблемами обобщающей способности и этическими рисками, связанными с алгоритмической торговлей.

Финансовые временные ряды – такие как котировки акций, валютные курсы, индексы волатильности или объёмы торгов – обладают рядом свойств, существенно отличающих их от временных рядов в других областях, например в метеорологии, инженерии или демографии. Эти особенности создают серьёзные трудности для построения надёжных и устойчивых прогнозных моделей и во многом определяют выбор методов их анализа.

Во-первых, финансовые временные ряды характеризуются стационарностью. В ценовых рядах каждое значение зависит от длинной истории ценовых уровней, следовательно, они имеют память. Исследователи в большинстве своем анализируют

изменение цены, т.е. работают с приращениями цен (или логарифмами приращений), изменениями доходностей или волатильности. Подобные трансформации временного ряда делают временной ряд стационарным, но при этом удаляют всю память из ценовых последовательностей, потому что строятся на ограниченной длине окна. Долгосрочная зависимость (долгая память или долгосрочная персистентность) является основой для предсказательных свойств моделей, поэтому модель должна содержать некоторое количество памяти для оценки того, как далеко цена удалась от своего ожидаемого значения. Статистические свойства таких рядов – среднее, дисперсия, автокорреляционная структура – могут меняться со временем под влиянием макроэкономических событий, изменений в регулировании, рыночных настроений или технологических сдвигов [1]. Это делает неприменимыми многие классические методы, предполагающие стационарность процесса.

Во-вторых, на финансовых рынках наблюдается высокая волатильность, часто проявляющаяся в виде кластеризации: периоды высокой изменчивости сменяются периодами относительного спокойствия. Эта особенность хорошо описывается моделями типа GARCH, однако она также осложняет прогнозирование, особенно на краткосрочных горизонтах.

В-третьих, финансовые данные содержат значительный уровень шума. Цена в каждый момент времени формируется совокупностью действий множества участников рынка с разными целями, стратегиями и доступом к информации. Однако существующие модели не включают методы устранения шума. В результате большинство эконометрических исследований приходят к ложным выводам, подкрепленным шумом, а не сигналом.

Статистические тесты эффективности прогнозирования модели обычно проводятся путем разделения заданного набора данных на обучающую выборку, используемую для выбора модели и оценки её гиперпараметров, и тестовую выборку, на данных которой модель не обучалась, и которая используется для оценки эффективности обобщающей способности модели. Модель, полученная в результате обучения на данных, в которых имеется корреляция и выбросы, имеет низкое качество прогнозирования.

В рамках данного исследования будут рассмотрены несколько ключевых алгоритмах машинного обучения, которые хорошо показали себя в задачах прогнозирования временных рядов. Одним из наиболее простых и распространенных методов является линейная регрессия. Этот подход позволяет установить линейные зависимости между переменными и часто используется как базовая модель для

сравнения с более сложными методами. Линейная регрессия может быть эффективной при наличии линейных трендов данных, однако ограничения не позволяют сделать реалистичный вывод, когда данные демонстрируют сложные нелинейные зависимости или сезонные колебания [2].

Модели на основе деревьев, такие как Random Forest и особенно градиентный бустинг (XGBoost, LightGBM, CatBoost), стали одними из самых популярных инструментов в прикладном финансовом прогнозировании. Их преимущества включают устойчивость к выбросам, способность автоматически учитывать взаимодействия между признаками и высокую интерпретируемость важности переменных. В отличие от нейросетей, такие модели не требуют нормализации данных и хорошо работают на небольших выборках. Однако они не обрабатывают последовательности напрямую: временная зависимость моделируется исключительно через искусственно созданные лаги и агрегаты [3]. Тем не менее, на практике ансамбли деревьев часто оказываются конкурентоспособными или даже превосходят глубокие модели по устойчивости и скорости обучения.

Для работы с последовательностями естественным выбором стали рекуррентные архитектуры, в особенности LSTM (Long Short-Term Memory) и GRU (Gated Recurrent Unit). Эти модели способны запоминать долгосрочные зависимости и учитывать динамику временного ряда без явного конструирования признаков. LSTM получили широкое распространение в финансах благодаря своей способности моделировать нелинейные временные паттерны и адаптироваться к изменяющейся волатильности. Однако их эффективность сильно зависит от объема данных, качества предобработки и регуляризации, а также подвержена риску переобучения.

Рекуррентная нейронная сеть без сохранения состояния представляет собой разновидность нейронной сети, в которой данные передаются в виде перекрывающихся окон. В конце каждого окна формируется вектор состояния, который затем сбрасывается и не используется дальше (Рис. 1). Эта особенность усложняет запоминание длительных паттернов временного ряда, так как сеть может удерживать информацию только о нескольких десятках временных шагов.

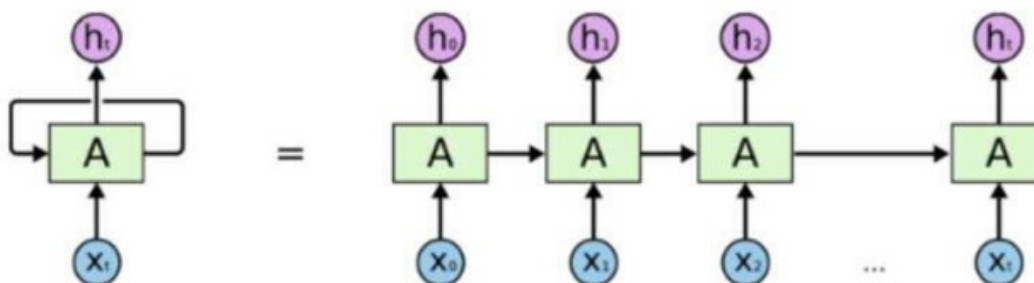


Рисунок 1. Пример работы развернутой рекуррентной нейронной сети (RNN)

Long Short-Term Memory (LSTM) – это метод, который позволяет рекуррентным нейронным сетям запоминать долгосрочные последовательности данных. Нейросеть LSTM обладает более продолжительной памятью, что позволяет ей закреплять паттерны/модели, охватывающие более 100 временных шагов. Несмотря на это, работа с очень длинными последовательностями может представлять трудности [4]. Кроме того, LSTM решает проблему исчезающих градиентов, которая часто возникает при использовании стандартных рекуррентных нейронных сетей и требует внесения определенных изменений. Данные нейронные сети эффективны в задачах классификации, обработки и прогнозирования временных рядов.

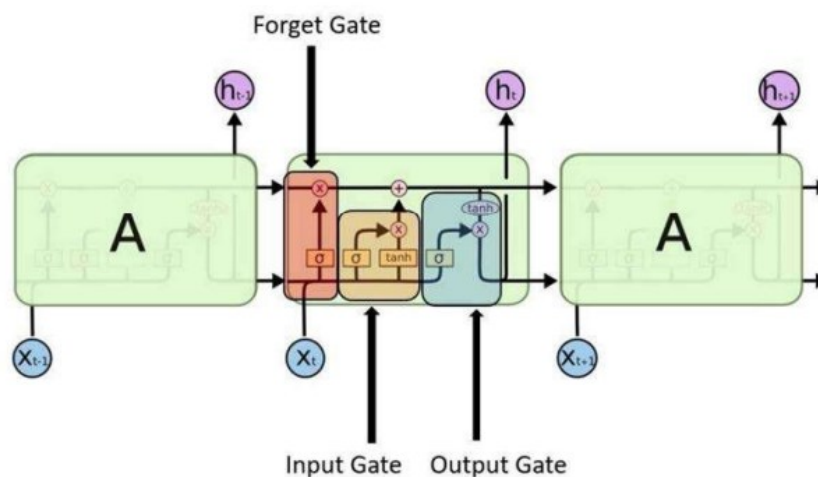


Рисунок 2. Составляющие LSTM

В финансовом секторе нейронные сети активно используются для прогнозирования цен акций, валютных курсов и других финансовых инструментов. Например, компания Goldman Sachs применяет LSTM для анализа исторических данных

о ценах на золото и других активов. Модель обучается на больших объемах данных, включая информацию о предыдущих ценах, объемах торгов и даже новостях. После обучения, такая модель пытается предсказать предстоящие значения цен. В данном случае, такие метрики, как RMSE и MAPE, помогают трейдерам оценивать качество прогнозов. Например, если внедрение LSTM модели приводит к снижению MAPE на 5%, это может значительно повысить прибыльность торговых стратегий [4].

Одним из наиболее известных случаев применения МО в трейдинге является хедж-фонд Renaissance Technologies, который на протяжении десятилетий использует сложные математические и статистические модели, в том числе методы машинного обучения, для генерации альфа-доходности [5].

В академической и прикладной литературе также описаны успешные применения градиентного бустинга для краткосрочного прогнозирования направления движения цены. Например, исследование Dixon et al. (2020) продемонстрировало, что XGBoost, обученный на технических индикаторах и микроструктурных признаках, способен генерировать статистически значимые сигналы на валютном рынке (forex) с положительной информационной долей (Information Ratio). Аналогично, Fischer et al. (2018) показали, что LSTM-сети могут эффективно классифицировать трендовые и контртрендовые режимы на рынке акций S&P 500, что позволяет строить прибыльные торговые стратегии после учёта транзакционных издержек.

Однако многочисленные попытки внедрения МО в трейдинг заканчиваются неудачей. Яркий пример — бум «AI-трейдинговых ботов» в 2017–2021 гг., когда множество стартапов и частных разработчиков обещали сверхдоходы с помощью нейросетей. Большинство таких систем, протестированных независимыми исследователями (например, в работах Bailey et al., 2019), показали, что их прибыльность исчезает после учёта спредов, комиссий и проскальзывания. Более того, в условиях рыночных шоков (например, во время кризиса 2020 года или банковского коллапса в 2023 г.) многие алгоритмические стратегии демонстрировали катастрофическую просадку, так как обучались на «нормальных» данных и не учитывали экстремальных сценариев (concept drift).

Сложность решения задачи прогнозирования роста цен и акций на рынке обусловлена наличием значительного количества факторов, влияющих на стоимость ценных бумаг и приводящих к высокой волатильности и зашумлению данных [5]. Для приведения к стационарности ряд дифференцируют, т.е. аналитики работают с приращениями или логарифмами приращений цен. Это позволяет применить модели,

основанные на автокорреляционной функции (ARMA, ARIMA). Однако дифференцирование ряда удаляет всю память из ценовых последовательностей.

Для того чтобы ряд был стационарным и при этом содержал максимально возможное количество памяти, используют дробное дифференцирование.

Для решения задачи прогнозирования финансовых временных рядов представляется перспективным использовать методы машинного обучения, в частности, нейронные сети, поскольку они обладают тем преимуществом, что способны обобщить и выделить скрытые зависимости между данными, на основе чего прогнозировать или проводить классификацию.

Анализ современных подходов к прогнозированию финансовых временных рядов показывает, что, несмотря на значительный прогресс в методологии и вычислительных возможностях, задача надёжного и устойчивого прогнозирования рыночных тенденций остаётся чрезвычайно сложной. Модели машинного обучения продемонстрировали способность выявлять нелинейные закономерности, адаптироваться к динамике рынка и в отдельных случаях генерировать статистически значимые торговые сигналы. Однако их практическая эффективность сильно ограничена фундаментальными свойствами финансовых данных: высоким уровнем шума, стационарностью, низким соотношением «сигнал/шум» и склонностью к резким структурным сдвигам. Наиболее устойчивые результаты, как правило, достигаются не за счёт сложности архитектуры, а благодаря тщательной предобработке данных.

Библиографический список

1. Половинин Д.А. Применение искусственных нейронных сетей для автоматизированного прогнозирования временных рядов. – М.: Издательство НИУ ВШЭ, 2023. – 150 с.;
2. Лабусов, М.В. Обзор моделей анализа и прогнозирования высокочастотных финансовых временных рядов // Экономика и предпринимательство. 2019. № 6 (107). С. 1256-1258.
3. Моцаков М.А. Алгоритм создания простой нейросети для прогнозирования временных рядов. – СПб.: Издательство СПбГУ, 2022. – 120 с;
4. Сравнение нейросетевых моделей для прогнозирования временных рядов: LSTM, GRU и RNN против ARIMA и ETS. – [Электронный ресурс]. – URL: <https://w.researchgate.net/publication/332123456> (дата обращения: 14.12.2025)
5. Трегуб, И.В. Технический анализ финансовых рынков: учебник. Москва : Финансовый университет, 2013. 224 с.

КИБЕРЭКОНОМИКА И ЭКОНОМИКА ДАННЫХ

УДК 658.78:004

Коновалова В.К., Кириллова В.А. «Легкая» цифровизация как инструмент повышения эффективности логистики

"Light" digitalization as a tool for improving logistics efficiency

Коновалова Вера Константиновна,

Ассистент кафедры менеджмента и права, Санкт-Петербургский государственный университет промышленных технологий и дизайна. Высшая школа технологии и энергетики

Кириллова Валерия Антоновна,

Студент кафедры менеджмента и права, Санкт-Петербургский государственный университет промышленных технологий и дизайна. Высшая школа технологии и энергетики

Konovalova Vera Konstantinovna,

Assistant Professor of the Department of Management and Law, St. Petersburg State University of Industrial Technology and Design. Higher School of Technology and Energy

Kirillova Valeria Antonovna,

Student of the Department of Management and Law, St. Petersburg State University of Industrial Technologies and Design. Higher School of Technology and Energy

Abstract. The article describes the process of optimizing the operation of the recycling and packaging site in the Adamant RC. The study identified three main difficulties: disorganized traffic flows, poor visibility of planning, and a shortage of storage facilities. To solve these problems, a set of measures has been proposed, including a digital accounting system, an information monitor, and rack upgrades. Special attention is paid to evaluating cost-effective IT solutions that have shown a quick payback. The results of the study may be useful for other enterprises seeking to reduce operating costs.

Keywords: warehouse logistics, distribution center, recyclables, recycling containers, process optimization, digitalization, economic efficiency, ROI.

Аннотация. Статья описывает процесс оптимизации работы участка вторсырья и тары в РЦ «Адамант». В ходе исследования были определены три основные трудности: неорганизованность транспортных потоков, слабая видимость планирования и дефицит складских помещений. Для решения этих задач предложен комплекс мер, включающий цифровую систему учёта, информационный монитор и модернизацию стеллажей. Особое внимание уделено оценке экономически выгодных ИТ-решений, показавших быструю окупаемость. Результаты исследования могут быть полезны для других предприятий, стремящихся к снижению операционных расходов.

Ключевые слова: складская логистика, распределительный центр, вторсырьё, оборотная тара, оптимизация процессов, цифровизация, экономическая эффективность, ROI.

Рецензент: Мартеха Александр Николаевич – кандидат технических наук, доцент.
ФГБОУ ВО «РГАУ-МСХА им. К.А. Тимирязева»

В условиях формирования цифровой экономики и динамичного развития технологий четвертого промышленного уклада конкурентоспособность компаний все в большей степени определяется качеством управления их логистическими цепочками. Из сферы вспомогательных операционных издержек логистика трансформируется в стратегический актив, чья эффективность усиливается за счёт внедрения интернета вещей, обработки больших данных, алгоритмов искусственного интеллекта и средств автоматизации. Лидеры рынка последовательно внедряют роботизированные складские комплексы, беспилотный транспорт и распределённые реестры для формирования полностью автоматизированных и прозрачных систем поставок.

Тем не менее, для значительного числа российских предприятий, в частности в сегментах розничной торговли и дистрибуции, комплексная цифровая трансформация сопряжена со значительными трудностями. Высокие капитальные требования, продолжительные сроки окупаемости, необходимость глубинной реорганизации процессов и дефицит квалифицированных кадров формируют серьёзные барьеры для реализации масштабных технологических проектов. В этой связи актуализируется запрос на более адаптивную стратегию — тактическую цифровизацию. В отличие от полномасштабной модернизации инфраструктуры, данный подход ориентирован на адресное решение конкретных операционных проблем с применением простых, быстрых в развёртывании и финансово доступных цифровых инструментов. Речь может идти о специализированных надстройках для действующих систем, мобильных сервисах, интерактивных панелях мониторинга и иных решениях, требующих скромных инвестиций и минимального вмешательства в устоявшиеся рабочие процедуры.

Однако внедрение даже таких «низковесных» решений требует строгой экономической аргументации. На сегодня наблюдается явный недостаток в чётких, прикладных и универсальных методиках, которые позволили бы руководителям и ИТ-архитекторам оперативно и обоснованно оценить потенциальную отдачу от подобных инициатив. Классические подходы к расчёту возврата на инвестиции (ROI), совокупной стоимости владения (TCO) или чистой приведённой стоимости (NPV) зачастую излишне громоздки для анализа небольших поэтапных улучшений и не всегда в полной мере учитывают специфику логистических операций, где экономический эффект может выражаться не только в прямом сокращении расходов, но и в росте скорости, точности и надёжности процессов.

Целью данной работы является создание и практическая проверка структурированного фреймворка (поэтапной методики) для расчёта экономической эффективности «низкозатратных» ИТ-решений, нацеленных на совершенствование

логистических бизнес-процессов. В качестве демонстрационной площадки для апробации методики выбран конкретный операционный узел — участок приёма возвратного сырья и оборотной тары крупного распределительного центра. Анализ его деятельности выявил характерные признаки «узкого места»: неэффективное использование ресурсов, продолжительные технологические простои и низкую производительность труда. [1]

Прикладная ценность исследования состоит в разработке для менеджеров и специалистов по цифровым технологиям структурированного инструментария экономического обоснования ИТ-инвестиций. Данный метод даёт возможность отказаться от интуитивных решений в пользу доказательного управления, основанного на фактах, и подтвердить финансовую обоснованность поступательной, модульной цифровизации в рамках бюджетных ограничений.

Определение экономического эффекта от внедрения информационных систем остаётся одной из наиболее спорных областей на пересечении управленческой, экономической и компьютерной наук. В сфере логистики и управления цепочками создания стоимости эта проблема стоит особенно остро ввиду значительной капиталоемкости отрасли и непрерывно возрастающих требований к операционной эффективности.

Динамика развития методов цифровизации в логистике показывает устойчивую тенденцию смещения акцентов - от концепции «революционного преобразования» к философии постоянных поступательных улучшений. Долгое время доминирующей являлась парадигма «тяжёлой» цифровизации, подразумевающая одномоментную реализацию сложных корпоративных платформ, таких как системы планирования ресурсов предприятия (ERP) и специализированные системы управления складом (WMS). Бесспорно, эти решения формируют технологическую основу современного бизнеса, обеспечивая сквозную интеграцию информации и комплексное управление операциями. Тем не менее, их внедрение связано с фундаментальными трудностями: крупными первоначальными капиталовложениями, долгим и сложным циклом интеграции, потребностью в глубокой реорганизации существующих бизнес-процессов и высокими постоянными расходами на поддержку. Для множества организаций, в особенности для представителей среднего бизнеса или отдельных направлений крупных корпораций, подобные проекты остаются финансово неподъёмными или сопряжёнными с неприемлемым уровнем риска. [2]

Сейчас набирает обороты новый подход к цифровизации, который называют "гибкой цифровизацией" или тактической автоматизацией. Это как бы перенос лучших

идей из мира разработки программ (Agile, Scrum) и бережливого производства (Lean) в корпоративные ИТ. Вместо того чтобы кардинально менять все процессы, как это бывает при "тяжелой" цифровизации, "гибкая" предлагает более спокойный и поэтапный путь. В логистике это означает, что мы не пытаемся переделать все сразу, а концентрируемся на решении конкретных, понятных проблем, которые реально мешают. Главная идея – сделать минимально рабочий продукт (MVP). Это значит, что мы запускаем не идеальное решение, а то, которое уже сейчас помогает, а потом дорабатываем его, слушая, что говорят пользователи. Предпочтение отдается инструментам, которые легко подключаются к уже имеющимся системам (вроде WMS или TMS), а не тем, которые требуют полного демонтажа и замены. Важно, что все происходит быстро – от идеи до запуска проходит всего несколько недель или месяцев, и стоимость таких проектов не такая высокая, как при капитальных вложениях, а скорее относится к текущим расходам. Это делает такие инициативы гораздо более доступными как с финансовой, так и с организационной точки зрения.

К «легким» ИТ-решениям в логистике можно отнести широкий спектр инструментов: от специализированных мобильных приложений для водителей и кладовщиков и веб-дашбордов для визуализации KPI, до систем автоматической идентификации на основе QR-кодов, чат-ботов для обработки запросов и «надстроек», использующих API существующих WMS/Yard-систем для автоматизации рутинных операций. Их общая черта – узкая специализация и относительно низкая сложность внедрения. [3]

Традиционные методы оценки экономической эффективности ИТ-проектов, такие как ROI, NPV, IRR и TCO, изначально создавались для масштабных капитальных вложений. Однако применительно к менее значительным ИТ-решениям эти подходы сталкиваются с серьезными трудностями. Во-первых, они плохо учитывают неосязаемые выгоды (например, улучшение прозрачности, качества данных, удовлетворенности сотрудников или снижение рисков), которые часто являются ключевыми для тактических внедрений. Во-вторых, при тесной интеграции с существующей инфраструктурой сложно точно определить, какие затраты относятся только к проекту, и как именно новая система повлияла на общую экономию. В-третьих, для небольших проектов (с бюджетом в сотни тысяч рублей и быстрой окупаемостью) эти сложные финансовые расчеты становятся излишними. Наконец, классические методы больше подходят для анализа уже завершенных проектов, тогда как на начальном этапе требуется простая и быстрая модель для обоснования запуска. [4]

Таким образом, в научной и практической литературе наблюдается методологический вакуум: отсутствует общепризнанный, простой в применении, но достаточно строгий фреймворк, предназначенный специально для оценки точечных, быстрых и недорогих IT-улучшений в логистике. Существующие подходы либо избыточно громоздки, либо, напротив, чрезмерно упрощены и не учитывают специфику логистических операций, где эффект часто проявляется в сокращении временных, а не только прямых финансовых потерь.

Необходимость цифровизации логистических процессов не вызывает сомнений, однако путь к ней не обязательно лежит через масштабные и дорогостоящие проекты. Концепция «гибкой цифровизации» предлагает жизнеспособную альтернативу, но её распространение сдерживается отсутствием адекватных инструментов экономического обоснования. Существующие классические методы оценки IT-инвестиций требуют адаптации к особенностям «легких» решений, что обуславливает актуальность разработки специализированного практического фреймворка, представленного в следующем разделе.

На основе анализа ограничений традиционных методик и учета специфики тактической цифровизации предлагается шестишаговый практический фреймворк, ориентированный на менеджеров среднего звена, логистов и IT-специалистов предприятий. Его цель - предоставить структурированный и относительно быстрый инструмент для обоснования инвестиций в точечные IT-улучшения с понятным и измеримым результатом.

Методология начинается с количественной диагностики «болевой точки»: выделения узкого места (например, приёмка тары) и оценки операционных потерь (P) по временным, финансовым и качественным метрикам. Полученная сумма потерь служит основой для обоснования проекта.

На втором этапе проводится концептуализация IT-решения как прямого ответа на проблему. Решение описывается в бизнес-терминах через модель «Как было - как будет», с чётким указанием механизма воздействия (например, автоматизация передачи данных для сокращения времени). Результат - текстовое и схематичное описание, связанное с метриками потерь.

Далее следует детализация капитальных затрат (CAPEX) по статьям: разработка ПО, лицензии, оборудование, внедрение, обучение персонала и резервный фонд (10-20%). Итог - сводная таблица с итоговой суммой затрат (C).

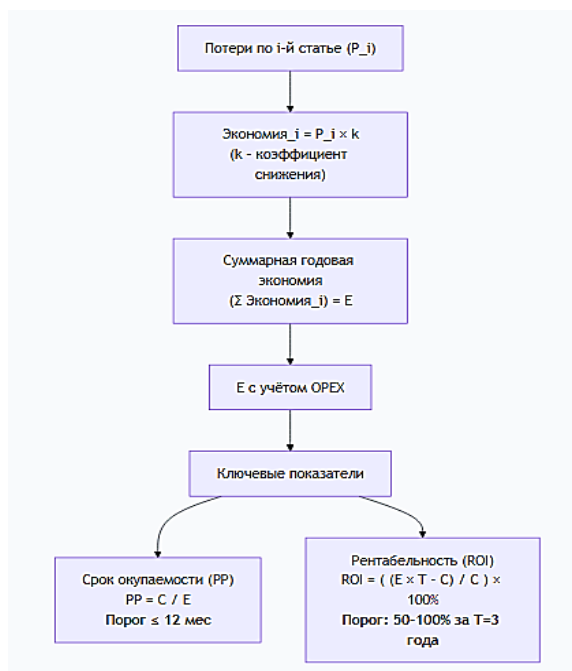


Рисунок 1. Алгоритм расчёта операционной экономии и показателей эффективности

На рисунке представлена последовательность расчёта ключевых показателей проекта. Процесс начинается с оценки потерь по каждой статье (P_i), которые умножаются на прогнозируемый коэффициент снижения k для определения потенциальной экономии. Суммарная годовая экономия (E) корректируется на ежегодные операционные расходы (OPEX) на поддержку решения. На основе полученного значения E и объёма первоначальных инвестиций (C) рассчитываются итоговые метрики: срок окупаемости (PP) и рентабельность инвестиций (ROI) за трёхлетний период (T). Формулы и целевые пороговые значения приведены на схеме. [4]

Заключительный шаг - это консолидация выводов и план внедрения. Если PP и ROI соответствуют порогам, проект признаётся целесообразным. Результаты оформляются в бизнес-заявку, включающую качественные выгоды, анализ рисков с мерами по их снижению и детальный план-график внедрения (roadmap).

Данная методология позволяет системно перейти от интуитивного желания «что-то улучшить с помощью IT» к обоснованному инвестиционному предложению, основанному на данных о текущих потерях, реалистичной оценке затрат и четком прогнозе экономического эффекта. В следующем разделе предложенный фреймворк будет применен к конкретному практическому кейсу оптимизации участка вторсырья распределительного центра.

Для демонстрации работоспособности предложенного фреймворка рассмотрим его пошаговое применение к реальной проблеме, детально исследованной в дипломной работе. Объектом выступает участок приемки вторсырья (ВС) и оборотной тары (ОТ) распределительного центра «Адамант» (ООО «Агроторг»), входящего в группу X5 Retail Group. Проблематика данного участка является классическим примером «узкого места», формирующего негативный каскадный эффект на ключевые логистические KPI всего центра [1].

На рисунке 2 представлена логическая модель, связывающая три выявленные системные проблемы участка ВС и ОТ с их непосредственными операционными последствиями и совокупным финансовым ущербом. Схема наглядно демонстрирует, как дисфункции в координации, планировании и инфраструктуре трансформируются в конкретные потери, снижая эффективность выполнения критически важной функции участка в экономике замкнутого цикла.

А на рисунке 3 представлена детализация операционных потерь участка ВС и ОТ, рассчитанная на основе данных хронометража, системных логов и финансовой отчетности. Суммарный ежемесячный ущерб до оптимизации оценивается в 84.5 тыс. рублей, при этом ключевой статьёй (более 70%) являются убытки от штрафов и конфликтов, вызванные дезорганизацией логистических процессов.

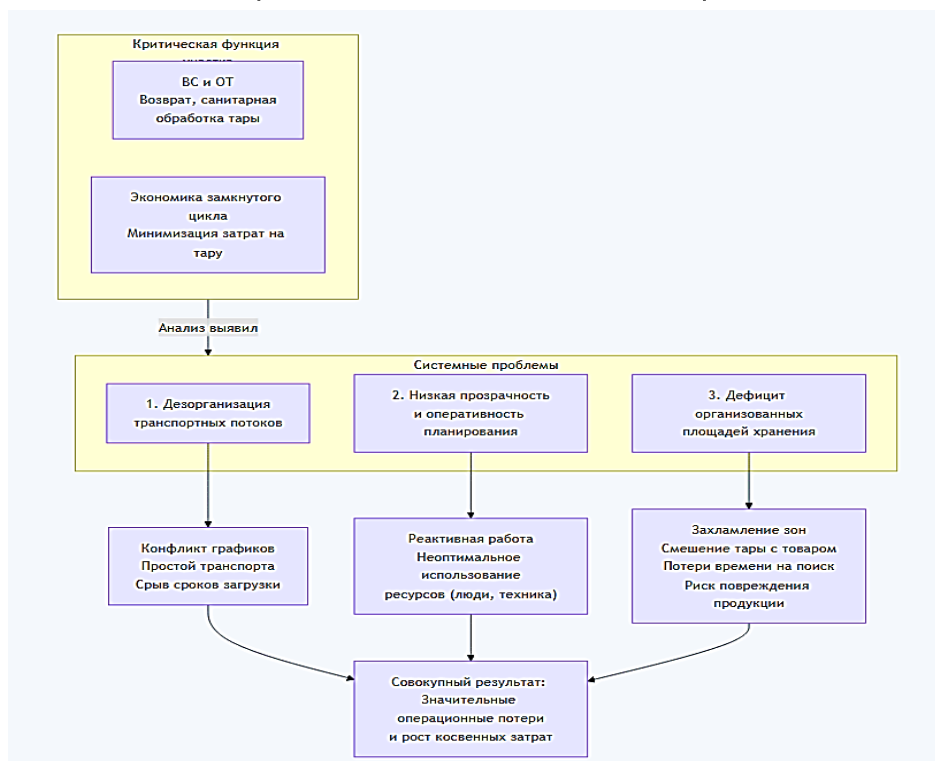


Рисунок 2. Причинно-следственная связь системных проблем участка ВС и ОТ

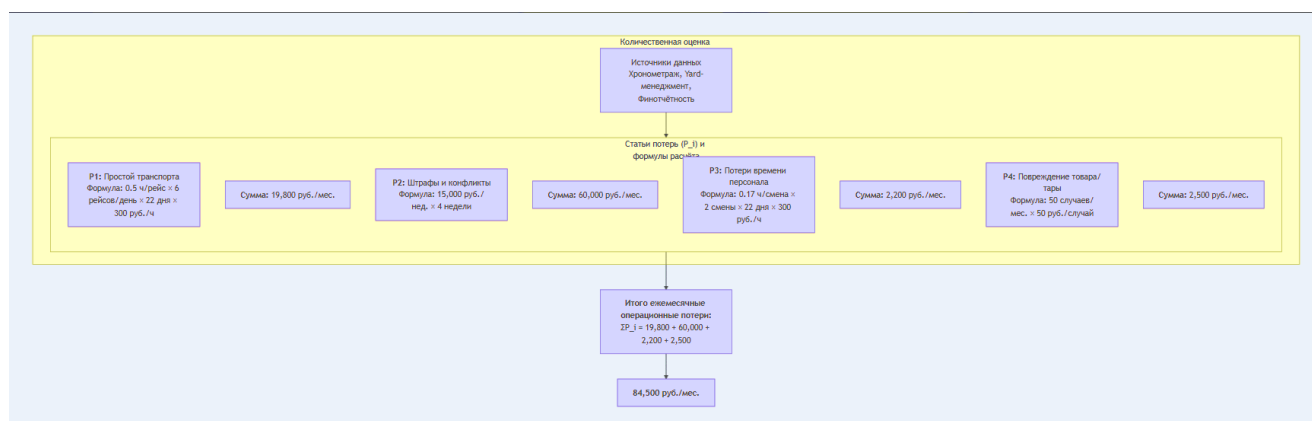


Рисунок 3. Расчёт операционных потерь участка ВС и ОТ по статьям (базовый период)

В ответ на выявленные проблемы был предложен не единый монолитный продукт, а комплекс из трех взаимодополняющих, быстро внедряемых решений, каждое из которых атакует конкретную причину потерь.

Решение А: «Цифровой маркер статуса водителя». Модуль для мобильного приложения водителя и веб-интерфейса терминала на въезде, позволяющий водителю при регистрации указать флаг «Требуется сдача вторсырья». Данные в реальном времени интегрируются в систему управления двором «Ярд». Приоритетные рейсы (с временем до загрузки <30 мин.) автоматически направляются на выделенную линию приемки. Механизм воздействия: Ликвидация причины Р1 и Р2 через устранение информационного разрыва и алгоритмическую приоритезацию.

Решение Б: «Информационный монитор оперативной обстановки». Веб-дашборд, выводимый на крупный монитор в зоне работы участка. В реальном времени отображает план прибытия машин с ВС на текущие сутки (из «Ярд»), текущую загрузку линий, статус заполнения зон хранения. Механизм воздействия: Ликвидация причины Р3 через повышение прозрачности и возможности для превентивного планирования работы бригады.

Решение В: «Организация цифрового следа для складских активов». Хотя данное решение включает физический элемент – установку дополнительных стеллажных систем (MS STANDART, 2000×1000×400 мм), его ядром является организационно-технологическое изменение. Каждой секции стеллажа присваивается цифровой идентификатор, а в мобильном приложении кладовщика появляется функция быстрого учета размещения/изъятия тары, привязанная к этой секции. Механизм воздействия: Ликвидация причины Р4 через создание структурированной и учитываемой системы хранения, что также косвенно сокращает Р3.

Расчеты выполнены с консервативными допущениями. Коэффициент снижения потерь (к) для Решений А и Б взят как 0.5 (снижение на 50%), для Решения В – 0.8 (снижение на 80%), исходя из специфики воздействия.

Таблица 1

Детализация капитальных затрат (С)

Категория затрат	Решение А (Маркер статуса)	Решение Б (Монитор)	Решение В (Стеллажи + цифр.учет)	Итого
1. Разработка/доработка ПО	200,000	50,000	20,000	270,000
2. Аппаратное обеспечение	20,000 (терминалы)	15,000 (монитор)	18,920 (стеллажи)	53,920
3. Внедрение и интеграция	50,000	10,000	10,000	70,000
4. Обучение персонала	30,000	15,000	5,000	50,000
5. Резерв (15%)	45,000	13,500	8,088	66,588
ИТОГО CAPEX (С)	345,000	103,500	62,008	510,508

Таблица 2

Расчет годовой операционной экономии (Е)

Статья потерь	Исходные потери (P _i), руб./год	Решение	Кэф. снижения (к)	Вклад в экономии, руб./год
Простой транспорта (P1)	237,600	А	0.5	118,800
Штрафы/конфликты (P2)	720,000	А	0.5	360,000
Потери времени персонала (P3)	26,400	Б	0.5	13,200
Потери от повреждений (P4)	30,000	В	0.8	24,000
Итого экономия до OPEX				516,000
Годовые OPEX (поддержка, 10% от dev-cost)				-27,000
ЧИСТАЯ ГОДОВАЯ ЭКОНОМИЯ (Е)				489,000

Таблица 3.

Ключевые показатели эффективности

Показатель	Формула	Расчет	Значение	Порог	Соответствие
Срок окупаемости (PP)	C / E	510,508 / 489,000	~1.04 года (≈12.5 мес.)	≤ 12 мес.	Пограничное (+0.5 мес.)
ROI за 3 года	$((E3 - C) / C)100\%$	$((489к3 - 510.5к) / 510.5к)100\%$	~187%	≥ 100%	Превышает

Расчеты на реальных данных демонстрируют высокую экономическую эффективность предложенного пакета «легких» IT-решений. При сроке окупаемости чуть более года и ROI на уровне 187% за трёхлетний период каждый вложенный рубль генерирует почти два рубля чистой экономии, что существенно превосходит установленные пороговые значения. Привлекательность проекта усиливают ключевые качественные выгоды: повышение управляемости за счёт создания цифрового потока данных для анализа, снижение операционных рисков и формирование технологического фундамента для будущих улучшений (например, внедрения RFID). Анализ чувствительности подтверждает устойчивость выводов: даже в пессимистичном сценарии (снижение экономии на 20% при росте затрат на 15%) проект сохраняет окупаемость около 1.5 лет и ROI около 100%, что соответствует минимальным требованиям. Таким образом, применение шестишагового фреймворка позволило количественно обосновать, что целевые цифровые решения, направленные на устранение измеренных «болевых точек», способны обеспечить высокую рентабельность и быструю окупаемость, внося существенный вклад в операционную эффективность логистического комплекса.

Практическое применение разработанного фреймворка на кейсе РЦ «Адамант» подтвердило его эффективность и позволило сформулировать ключевые инсайты для тактической цифровизации логистики. Во-первых, успех определяется не сложностью технологии, а фокусом на измеримых операционных KPI, таких как время простоя транспорта или уровень штрафов, где основной эффект достигается за счёт ликвидации информационных разрывов, а не «интеллектуализации». Во-вторых, синергия логически связанных «лёгких» решений (цифровой маркер, мониторинг, организация хранения) создаёт кумулятивный эффект, превосходящий простую сумму отдельных улучшений, что подчёркивает важность системного подхода даже при точечной оптимизации. В-третьих, критическую роль играет организационная

компонента: цифровизация требует параллельной реорганизации физических процессов и рабочих мест, без которой IT-инструменты теряют свою эффективность. Наконец, для операционных менеджеров ключевым критерием остаётся срок окупаемости в пределах одного финансового года - психологически значимый и понятный порог, в то время как высокий ROI служит весомым дополнительным аргументом. [6]

На основе результатов исследования предлагается следующий алгоритм внедрения для практического применения. Цифровое преобразование необходимо начинать с этапа «аудита операционных узких мест», предусматривающего оцифровку текущих процессов и перевод субъективных дискуссий в плоскость объективных количественных показателей, таких как сокращение среднего времени простоя на 25 минут. Для достижения максимального эффекта рекомендуется фокусироваться не на внедрении единичных инструментов, а на формировании комплекса взаимодополняющих «лёгких» решений, нацеленных на устранение одной сквозной операционной проблемы.

Предложенный фреймворк экономического обоснования целесообразно закрепить в качестве стандартной процедуры для оценки тактических IT-проектов с капитальными затратами до 5 млн рублей, что позволит структурировать процесс принятия решений и создать единую терминологическую базу для кросс-функциональных команд. Критически важным элементом успеха является выделение в бюджете проекта отдельной статьи (рекомендуемый резерв до 15%) на сопутствующие организационные изменения и системное обучение персонала, поскольку именно игнорирование этой составляющей чаще всего приводит к неудачным результатам внедрения.

Следует учитывать, что область применения данного подхода имеет определённые ограничения. Методология наиболее релевантна для проектов в сфере дистрибуции с капитальными затратами до 5–7 млн рублей и требует существенной адаптации при переносе в другие отрасли или для масштабных стратегических инициатив.

Перспективы дальнейшего развития подхода видятся в его интеграции с инструментами предиктивной аналитики для перехода от реактивного устранения проблем к проактивному планированию операционных ресурсов. Эволюция методологии может привести к созданию «цифровых двойников» ключевых участков, позволяющих моделировать сценарии изменений и оценивать их потенциальный эффект до начала физической реализации. Кроме того, открытым остаётся вопрос

тиражирования данного фреймворка на оптимизацию других типовых логистических операций, таких как комплектация заказов или управление складскими запасами. Цифровая трансформация логистики не обязательно должна начинаться с многомиллионных контрактов и многолетних проектов. Как демонстрирует кейс РЦ «Адамант», значительный экономический эффект и рост операционной зрелости могут быть достигнуты за счет последовательного, измеримого и финансово обоснованного внедрения «легких» IT-решений. Разработанный фреймворк предоставляет для этого необходимый инструментарий, переводя цифровизацию из области стратегических абстракций в плоскость практических, окупаемых управленческих действий.

Практическое исследование подтвердило рабочую гипотезу: значимый рост эффективности логистических операций достижим не только через масштабные капиталоемкие преобразования, но и посредством точечного внедрения сфокусированных, «низкобюджетных» IT-инструментов. Ключевым сдерживающим фактором для распространения данной практики выступает дефицит доступных и применимых на практике методик расчёта экономической целесообразности, которые были бы одинаково понятны и убедительны для специалистов в области информационных технологий, операционного управления и финансового анализа.

Основным научно-практическим результатом данной работы стала разработка и успешная апробация шестишагового практического фреймворка для оценки экономической эффективности точечных цифровых проектов в логистике. Методология системно охватывает весь цикл – от количественной диагностики «болевого точки» и концептуализации решения до детального расчета затрат, экономии и ключевых показателей окупаемости. Важнейшим элементом фреймворка стало установление четких прагматичных пороговых значений: срок окупаемости (Payback Period) в пределах 12 месяцев и рентабельность инвестиций (ROI) не менее 100% за трехлетний период. Эти критерии формируют объективную основу для принятия управленческих решений. [7]

В ходе апробации на реальном кейсе – оптимизации участка приемки вторсырья и оборотной тары РЦ «Адамант» – фреймворк доказал свою работоспособность и высокую практическую ценность. Расчеты на реальных операционных и финансовых данных показали, что пакет из трех взаимодополняющих решений (цифровой маркер статуса, информационный монитор, организация цифрового следа с модернизацией стеллажей) способен обеспечить срок окупаемости около 1 года и ROI на уровне 187% за 3 года. Эти цифры не только соответствуют, но и существенно превосходят установленные пороги, делая проект экономически безусловно привлекательным.

Анализ чувствительности подтвердил устойчивость выводов даже при консервативных сценариях отклонения по затратам и экономии.

Анализ операционной деятельности позволил сформулировать базовые принципы результативного управления. В сфере логистики качественное улучшение показателей зачастую становится следствием не внедрения ресурсоёмких интеллектуальных платформ, а ликвидации элементарных, но значимых пробелов в информационных потоках, а также автоматизации повседневных коммуникационных процедур. При этом совокупный эффект от набора скоординированных, даже простых решений, существенно превышает результат от отдельных, пусть и технически сложных, усовершенствований.

Важнейший вывод заключается в том, что результативность «низковесной» цифровизации лишь частично (ориентировочно на 30%) обусловлена применяемыми технологическими средствами. Основной вклад (порядка 70%) вносят тщательно выстроенные организационные изменения и комплексная адаптация персонала к новым процессам. Следовательно, описываемая методология трансформирует сам подход к цифровому преобразованию, переориентируя внимание с глобальных стратегических концепций и сложных технических выборов на конкретные, поддающиеся количественной оценке и оперативно окупаемые управленческие шаги. Она предоставляет руководителям практический структурированный механизм для запуска преобразований «снизу вверх», отправной точкой для которых становятся текущие операционные трудности, финансовые последствия которых поддаются точному расчёту уже в настоящий момент. Перспективы дальнейших исследований связаны с расширением области применения фреймворка на другие типовые логистические процессы (приемка товаров, комплектация, инвентаризация), его интеграцией с методами предиктивной аналитики для более точного прогнозирования экономического эффекта, а также с разработкой на его основе отраслевых стандартов и шаблонов для быстрой оценки типовых IT-решений. Внедрение подобных практик на предприятиях станет реальным вкладом в построение гибкой, устойчивой и технологически оснащенной логистической системы, способной отвечать на вызовы современной экономики.

Библиографический список

1. Общая информация о предприятии X5 Group. – URL: <https://www.x5.ru> (дата обращения: 19.02.2025).

2. ГОСТ Р 56906-2016 «Логистика. Термины и определения». – М.: Стандартинформ, 2016. – 16 с. –
3. Голов, Р. С., Агарков, А. П., Мыльник, А. В. Организация производства, экономика и управление. – М.: Дашков и К, 2019. – 858 с
4. Сергеев, В. И. Логистика : учебник для бакалавров. – М.: Юрайт, 2012. – 848с.
5. Методические рекомендации по организации складского хозяйства в розничных торговых сетях. – М.: Минпромторг России, 2020. – 48 с.
6. McKinsey & Company. Цифровая логистика: технологическая гонка набирает обороты // McKinsey Insights. – 2023. – URL: <https://www.mckinsey.com/capabilities/operations/our-insights/digital-logistics-technology-race-gathers-momentum> (дата обращения: 10.12.2025).
7. Цифровая трансформация в цепочках поставок: повышение устойчивости и экологичности // Frontiers in Sustainability. – 2025. – Том 6. – Статья 1584580. – URL: <https://www.frontiersin.org/articles/10.3389/frsus.2025.1584580/full> (дата обращения: 10.12.2025).

Электронное научное издание

КИБЕРНАУКА

НАУЧНЫЙ ЭЛЕКТРОННЫЙ ЖУРНАЛ

№2/2025

По вопросам и замечаниям к изданию, а также предложениям к сотрудничеству обращаться по электронной почте mail@scipro.ru

Подготовлено с авторских оригиналов

ISSN 3033-8131

Усл. печ. л. 3,6

Объем издания 14,6 МВ

Издание: Международный научный электронный журнал «Право, общество, справедливость»
Учредитель: Краснова Н.А.

Издательство Индивидуальный предприниматель Краснова Наталья Александровна
Адрес редакции: Россия, 603000, г. Нижний Новгород пл. М. Горького, 4/2, 4 этаж, офис №1,
Тел.: +7962508702