Научное издательство
Профессиональный Вестник

Информационные технологии и безопасность

Выпуск №1/2024

Научный журнал для лучших специалистов отрасли. Внутри: оригинальные работы по ИИ, анализу данных, облачным технологиям и инновациям в сфере ИТ.

ИНДЕКСАЦИИ ЖУРНАЛА

Google Scholar · International Scientific Indexing · Academic Resource Index ResearchBib · Научная электронная библиотека eLIBRARY.RU · CyberLeninka · CiteFactor Academic Scientific Journals

ISSN 3100-444X

support@professionalbulletinpublisher.com

professionalbulletinpublisher.com/

Brasov, Sat Sanpetru, Comuna Sanpetru, Str. Sfintii Constantin si Elena, nr. 6

Scientific publishing house
Professional Bulletin

Information Technology and Security

Issue №1/2024

A scientific journal for the best specialists in the industry. Inside: original works on AI, data analysis, cloud technologies and IT innovations.

INDEXATIONS

Google Scholar · International Scientific Indexing · Academic Resource Index ResearchBib · Научная электронная библиотека eLIBRARY.RU · CyberLeninka · CiteFactor Academic Scientific Journals

ISSN 3100-444X

support@professionalbulletinpublisher.com

professionalbulletinpublisher.com/

Brasov, Sat Sanpetru, Comuna Sanpetru, Str. Sfintii Constantin si Elena, nr. 6

Научное издательство «Профессиональный вестник»

Журнал «Профессиональный вестник. Информационные технологии и безопасность»

Профессиональный вестник. Информационные технологии и безопасность – профессиональное научное издание. Публикация в нем рекомендована практикам и исследователям, которые стремятся найти решения для реальных задач и поделиться своим опытом с профессиональным сообществом. Публикация в журнале подходит для тех специалистов, кто работает и активно развивает передовые ИТ-решения, такие как технологии ИИ, блокчейна, больших данных и другие.

Журнал рецензирует все входящие материалы. Рецензирование – двойное слепое, осуществляется внутренними и внешними рецензентами издательства. Статьи индексируются во множестве международных научных баз, доступ к базе данных журнала открыт для любого читателя. Публикация журнала происходит 4 раза в год.

Сайт издательства: https://www.professionalbulletinpublisher.com/

Выпуск № 1/2024

Жудец Брашов, Румыния

The scientific publishing house «Professional Bulletin»
Journal «Professional Bulletin. Information Technology and Security»

Professional Bulletin. Information Technology and Security is a professional scientific journal. The publication in it is recommended to practitioners and researchers who seek to find solutions to real-world problems and share their experiences with the professional community. The publication in journal is suitable for those specialists who work and actively develop advanced IT solutions, such as AI, blockchain, big data technologies and others.

The journal reviews all incoming materials. The review is double-blind, carried out by internal and external reviewers of the publishing house. Articles are indexed in a variety of international scientific databases, and access to the journal's database is open to any reader. Publication in the journal takes place 4 times a year.

Publisher's website: https://www.professionalbulletinpublisher.com/

Issue №1/2024
Brasov County, Romania

Содержание выпуска

Contents

УДК 004.8

АНАЛИЗ ВОЗМОЖНОСТЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МЕДИЦИНСКОЙ ДИАГНОСТИКЕ

Кириллов С.М.
бакалавр, Дальневосточный федеральный университет
(Владивосток, Россия)

ANALYSIS OF ARTIFICIAL INTELLIGENCE CAPABILITIES IN MEDICAL DIAGNOSTICS

Kirillov S.
bachelor's degree, Far Eastern Federal University
(Vladivostok, Russia)

Аннотация

В статье представлен анализ возможностей и вызовов применения искусственного интеллекта (ИИ) в медицинской диагностике. Рассмотрены основные направления использования ИИ, включая обработку текстовой информации, анализ медицинских изображений и предсказание клинических исходов. В ходе работы было отмечено, что автоматизация диагностики позволяет повысить скорость и точность медицинских заключений, однако существуют значительные вызовы. К ним относятся вопросы интерпретируемости, конфиденциальности данных и возможные этические аспекты. Выявлены перспективы для дальнейшего развития ИИ в медицине, такие как совершенствование моделей и подготовка специалистов для работы с новыми технологиями. Данная статья подчеркивает значимость ИИ в медицине и необходимость тщательного подхода к его внедрению.

Ключевые слова: искусственный интеллект, медицинская диагностика, интерпретируемость, обработка данных, безопасность данных.

Abstract

The article presents an analysis of the capabilities and challenges associated with the use of artificial intelligence (AI) in medical diagnostics. Key applications of AI, such as text information processing, medical image analysis, and clinical outcome prediction, are examined. The study indicates that diagnostic automation enhances the speed and accuracy of medical decisions, yet there are significant challenges, including interpretability, data confidentiality, and potential ethical issues. Prospects for further AI development in medicine are identified, including model improvements and specialist training for working with new technologies. This article highlights the importance of AI in medicine and the need for a careful approach to its integration.

Keywords: artificial intelligence, medical diagnostics, interpretability, data processing, data security.

Введение

Искусственный интеллект (ИИ) активно внедряется в различные отрасли, включая медицину, где его потенциал в диагностических процессах вызывает особый интерес. Современные алгоритмы ИИ способны обрабатывать огромные объемы данных и выявлять паттерны, недоступные человеческому восприятию, что открывает новые возможности для раннего выявления и мониторинга заболеваний. Тем не менее, для полноценного внедрения

ИИ в клиническую практику требуется четкое понимание его возможностей и ограничений, а также специфических условий применения в медицинских учреждениях.

В последние годы появились многочисленные исследования, посвященные применению методов машинного обучения и глубокого обучения в диагностике таких заболеваний, как онкология, кардиология, нейрология и др. Эти методы, основанные на анализе изображений, сигналов и других клинических данных, демонстрируют высокую точность и эффективность, однако вопрос надежности и точности остается открытым, особенно в случае сложных клинических случаев. Важным аспектом также является интерпретируемость моделей, так как врачи и пациенты должны понимать, на каких данных основаны прогнозы ИИ.

Цель данной статьи состоит в анализе текущих возможностей и ограничений ИИ в медицинской диагностике, а также в выявлении ключевых направлений для дальнейших исследований и развития.

Основная часть

Развитие методов ИИ в медицинской диагностике значительно расширяет возможности современных технологий в области здравоохранения. Наиболее перспективные направления связаны с обработкой больших массивов данных, анализом медицинских изображений и текстовой информации, а также предсказанием клинических исходов. В результате применения ИИ улучшается точность и скорость диагностики, что позволяет медицинским специалистам быстрее принимать обоснованные решения на основе более полных и структурированных данных [1].

Первым важным направлением является использование ИИ для анализа текстовых данных, содержащихся в электронных медицинских картах [2]. Автоматический анализ таких данных позволяет ускорить извлечение информации о симптомах, назначениях и диагнозах. Например, ИИ может анализировать записи врача и автоматически выделять ключевые симптомы, на основании которых формируются первичные диагностические выводы. Это особенно важно для структурирования больших объемов данных, поступающих от пациентов, и для автоматизации первичного анализа перед консультацией врача. Практическое применение такого подхода позволяет значительно сократить время на сбор анамнеза и повысить точность диагностики.

```
import spacy
from spacy.matcher import Matcher

# Загрузка модели для анализа русского текста
nlp = spacy.load("ru_core_news_md")

# Определение шаблона для обнаружения симптомов
matcher = Matcher(nlp.vocab)
pattern = [{"LOWER": "боль"}, {"LOWER": "в"}, {"LOWER": "груди"}]
matcher.add("SYMPTOM_DETECTION", [pattern])

# Текст медицинской записи
medical_text = "Пациент жалуется на боль в груди и повышение давления."

# Применение модели к тексту и поиск шаблонов
doc = nlp(medical_text)
matches = matcher(doc)

# Вывод найденных симптомов
for match_id, start, end in matches:
    span = doc[start:end]
    print(f"Симптом обнаружен: {span.text}")
```

Этот код анализирует текст медицинских записей и выделяет указанные симптомы, автоматизируя процесс выявления ключевой информации из больших объемов данных. В реальных условиях программа может быть расширена для распознавания различных медицинских терминов, что обеспечит более детализированный анализ симптоматики [3].

Второе направление применения ИИ связано с анализом медицинских изображений, таких как рентген, компьютерная томография и магнитно-резонансная томография. Свёрточные нейронные сети (СНС), используемые для обработки изображений, показывают высокую точность в диагностике заболеваний. Например, алгоритмы на основе СНС могут автоматически сегментировать снимки, выделяя подозрительные области, которые могут указывать на новообразования или другие патологии. Внедрение таких технологий позволяет автоматизировать процессы первичного скрининга и ускоряет выявление заболеваний, что особенно важно для ранней диагностики онкологических заболеваний [4].

Кроме того, ИИ применяется для предсказания клинических исходов на основе анализа структурированных данных пациентов. Используя большие базы данных с медицинской информацией, алгоритмы ИИ могут прогнозировать вероятность развития заболеваний и прогнозировать исходы лечения. Например, на основе данных о пациенте, таких как возраст, анамнез, показатели анализов и других факторов, система может оценить риск сердечно-сосудистых осложнений или рецидивов заболеваний. Это позволяет врачам персонализировать подход к лечению и улучшать прогнозирование исходов.

На рисунке 1 представлена сравнительная точность моделей искусственного интеллекта при диагностике заболеваний в разных медицинских областях. Высокая точность наблюдается в онкологии, что подтверждает возможности ИИ в сложных диагностических задачах.

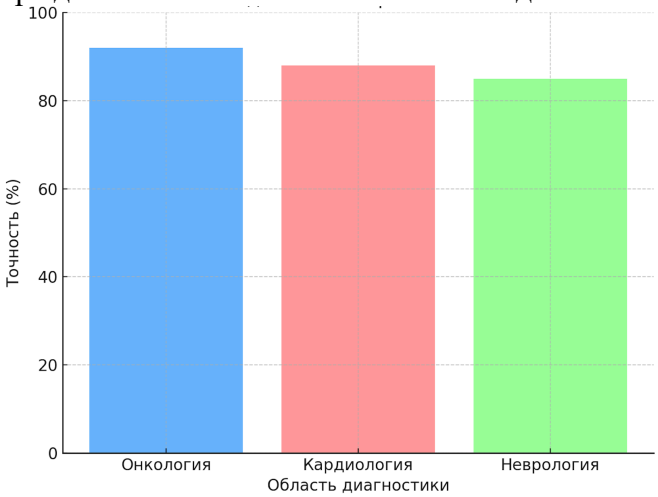

Рисунок 1. Точность ИИ в диагностике различных заболеваний

Из диаграммы видно, что наиболее высокая точность достигается в онкологии (92%), что связано с большим количеством доступных данных для обучения моделей. Однако в других областях, таких как кардиология (88%) и неврология (85%), точность несколько ниже, что подчеркивает необходимость улучшения алгоритмов и данных в этих направлениях.

Наконец, перспективным направлением является использование ИИ для анализа сигналов, таких как электрокардиограмма и электроэнцефалограмма. Системы, основанные на нейронных сетях, способны выявлять аномалии, указывающие на возможные патологии.

Вызовы и перспективы применения искусственного интеллекта в медицинской диагностике

Несмотря на высокие достижения и возможности, применение ИИ в медицинской диагностике сталкивается с рядом вызовов, которые требуют тщательного рассмотрения. Одним из основных препятствий является необходимость обеспечения точности и надежности моделей, особенно при работе с диагностикой сложных заболеваний. ИИ-модели часто требуют больших объемов качественных данных для обучения, и отсутствие таких данных может снизить точность предсказаний. Для преодоления этой проблемы необходимо

разработать стандарты и протоколы, которые помогут регулировать качество и структуру используемых данных [5].

Второй важный аспект – интерпретируемость решений, принимаемых ИИ. В медицине крайне важно, чтобы диагноз был обоснован и понятен медицинскому персоналу. В отличие от традиционных методов, многие алгоритмы ИИ, особенно глубокого обучения, представляют собой «черный ящик», где трудно определить, какие факторы повлияли на конечное решение. Это создает трудности в понимании модели и может снижать доверие специалистов. Решением может стать развитие методов объяснимого ИИ, способных давать прозрачные и понятные объяснения по каждому результату.

Третьим вызовом является необходимость соблюдения норм конфиденциальности и безопасности данных пациентов. Медицинские данные относятся к особо чувствительным категориям, и любые утечки могут привести к серьезным последствиям. Использование ИИ требует хранения и обработки больших массивов данных, что увеличивает риски. Для обеспечения безопасности требуется внедрение методов шифрования и строгих правил хранения данных, а также постоянный мониторинг систем на предмет уязвимостей.

Кроме того, значительное внимание следует уделить этическим вопросам, связанным с применением ИИ. Применение алгоритмов ИИ в медицине вызывает опасения, связанные с возможностью возникновения предвзятости в предсказаниях, что может повлиять на качество медицинского обслуживания [6]. Например, если алгоритм обучен на данных, которые не учитывают разнообразие в популяции, он может давать неверные результаты для определенных групп пациентов. Решением может стать обязательное тестирование ИИ на данных, отражающих демографическое разнообразие, чтобы исключить предвзятость в предсказаниях.

На рисунке 2 показано распределение основных вызовов, с которыми сталкиваются разработчики и медицинские учреждения при внедрении технологий искусственного интеллекта. Интерпретируемость моделей остается ключевым барьером.

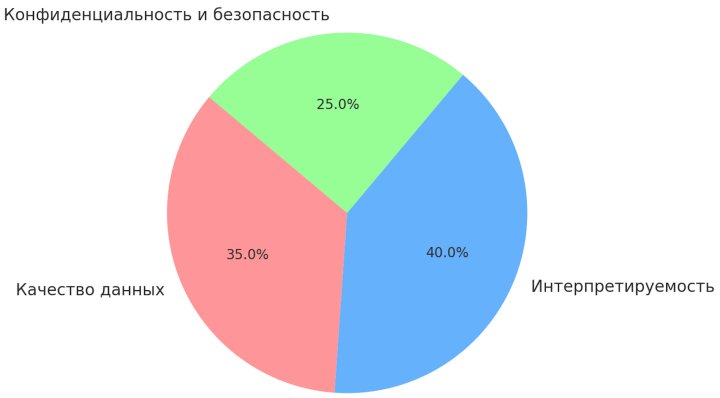

Рисунок 2. Основные вызовы внедрения ИИ в здравоохранении

Согласно данным, наиболее серьезным вызовом является интерпретируемость моделей (40%), так как врачи нуждаются в объяснимости выводов, сделанных ИИ. Качество данных занимает второе место (35%), указывая на необходимость стандартизации данных. Вопросы конфиденциальности и безопасности (25%) требуют создания новых технологий защиты информации для минимизации рисков.

Наконец, интеграция ИИ в существующую систему здравоохранения требует подготовки специалистов, которые смогут работать с новыми технологиями и анализировать результаты, предлагаемые ИИ. Для этого важно развивать образовательные программы и курсы повышения квалификации, которые позволят врачам и техническим специалистам эффективно взаимодействовать с системами ИИ. Подготовка специалистов к работе с ИИ также способствует более точной и быстрой интеграции технологий в медицинскую практику.

Заключение

Внедрение искусственного интеллекта в медицинскую диагностику открывает новые возможности для улучшения качества медицинских услуг и сокращения времени на

постановку диагноза. ИИ способен автоматизировать ряд процессов, включая обработку текстовых данных, анализ изображений и предсказание клинических исходов, что помогает медицинским специалистам принимать более обоснованные решения. Тем не менее, использование этих технологий требует разработки надежных моделей и соблюдения ряда стандартов для минимизации рисков.

Одной из ключевых задач на пути развития ИИ в медицине остается повышение интерпретируемости и объяснимости моделей, что особенно важно для доверия со стороны медицинского персонала и пациентов. Решения, предлагаемые ИИ, должны быть прозрачными и понятными, чтобы специалисты могли принимать их во внимание в клинической практике. Кроме того, соблюдение конфиденциальности и безопасности данных требует постоянного контроля и улучшения инфраструктуры для работы с медицинской информацией.

Для полноценного интегрирования ИИ в систему здравоохранения необходимо также уделить внимание подготовке квалифицированных кадров. Образовательные программы для врачей и технических специалистов помогут им эффективно работать с ИИ-системами и максимально использовать их потенциал в диагностике. Таким образом, ИИ представляет собой мощный инструмент для медицины, способный изменить подход к диагностике и лечению, но его внедрение должно быть осознанным и тщательно спланированным.

Список литературы

1. Хамидов О.А., Гайбуллаев Ш.О., Хакимов М.Б. Обзор методов обработки изображений для диагностики патологии головного мозга: проблемы и возможности // Journal of new century innovations. 2022. Т. 10. №5. С. 181-195.
2. Манкибаев Б.С. Основные направления внедрения искусственного интеллекта в медицине // Наука, образование и культура. 2019. №3(37). С. 69-71.
3. Вахрушева М.А. Искусственный интеллект // Интеллектуальный потенциал XXI века: ступени познания. 2011. №6. С. 162-166.
4. Фершт В.М., Латкин А.П., Иванова В.Н. Современные подходы к использованию искусственного интеллекта в медицине // Территория новых возможностей. Вестник Владивостокского государственного университета экономики и сервиса. 2020. Т. 12. №1. С. 121-130.
5. Литвин А.А., Стома И.О., Шаршакова Т.М., Румовская С.Б., Ковалев А.А. Новые возможности искусственного интеллекта в медицине: описательный обзор // Проблемы здоровья и экологии. 2024. Т. 21. №1. С. 7-17.
6. Холопов А.А., Козырева В.И., Тихонова О.В. Роль искусственного интеллекта в медицине // Наука и образование: актуальные вопросы теории и практики: Материалы III Международной научно-методической конференции, посвященной 50-летию Самарского государственного университета путей сообщения (Самара, 21-22 марта 2023 г.). Оренбург: ОрИПС-филиала СамГУПС. 2023. С. 179.
7. Алексеева М.Г., Зубов А.И., Новиков М.Ю. Искусственный интеллект в медицине // Международный научно-исследовательский журнал. 2022. №7-2(121). С. 10-13.

UDC 004.9

POTENTIAL AND PROSPECTS OF BLOCKCHAIN USAGE IN THE FINANCIAL SECTOR

Belyakova I.
specialist's degree, Northern (Arctic) Federal University
(Arkhangelsk, Russia)

ПОТЕНЦИАЛ И ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ БЛОКЧЕЙН В ФИНАНСОВОМ СЕКТОРЕ

Белякова И.П.
специалист, Северный (Арктический) федеральный университет
(Архангельск, Россия)

Abstract

The article explores the possibilities and challenges of using blockchain technology in the financial sector. It describes the main applications of blockchain, such as international transfers, identity verification, and business process automation through smart contracts. The study highlights blockchain's advantages, including decentralization, increased transparency, and reduced operational costs. Major challenges, such as scalability, security, and regulatory compliance, are identified as limiting factors for the widespread adoption of this technology in traditional financial structures. In conclusion, the article emphasizes that successful blockchain integration in the financial sector requires the development of security standards and adaptation to legal requirements.

Keywords: blockchain, finance, smart contracts, data security, scalability.

Аннотация

Статья посвящена исследованию возможностей и вызовов использования блокчейн-технологии в финансовом секторе. Описаны основные направления применения блокчейн, такие как международные переводы, верификация личности и автоматизация бизнес-процессов с помощью смарт-контрактов. Рассмотрены преимущества блокчейн, включая децентрализацию, повышение прозрачности и снижение операционных издержек. Выявлены основные вызовы, такие как масштабируемость, безопасность и соблюдение нормативных требований, которые ограничивают широкомасштабное внедрение технологии в традиционные финансовые структуры. В заключении подчеркивается, что успешная интеграция блокчейн в финансовый сектор требует разработки стандартов безопасности и адаптации к правовым требованиям.

Ключевые слова: блокчейн, финансы, смарт-контракты, безопасность данных, масштабируемость.

Introduction

The blockchain technology, initially used exclusively for cryptocurrency transactions, has acquired a much broader application scope in recent years, particularly in the financial sector. Blockchain offers a decentralized, secure, and transparent system for transaction recording that eliminates the need for intermediaries. Such a system has become especially relevant in the context of digitalizing financial processes and seeking more efficient ways to manage financial assets. Numerous organizations and banking institutions are actively exploring blockchain integration into their operations, aiming to reduce costs and enhance transaction security.

Significant attention to this technology is also due to its ability to ensure data immutability and prevent transaction manipulation. This is achieved through cryptographic algorithms, creating a resilient, tamper-resistant data management system. In the financial sphere, such advantages can foster increased trust in processes and reduce the level of financial crime. Blockchain integration in the banking sector also opens up opportunities for improved client interactions, simplifying data verification and accelerating processes related to credit issuance and identity verification.

The purpose of this article is to explore the potential of blockchain in the financial sector, analyze existing and possible applications, and identify the main challenges associated with integrating this technology into traditional financial systems.

Main part

Applying blockchain technology in the financial sector opens up significant prospects for improving existing processes such as international transfers, identity verification, and asset management. Using a distributed ledger enables a transparent and tamper-resistant system where each transaction is recorded without the possibility of subsequent modification. This not only reduces the time needed to process operations but also minimizes the risk of human error and data manipulation – an aspect that is critically important for financial institutions [1].

One of the most promising applications of blockchain is international money transfers. In the traditional banking system, such transfers can take several days and involve high fees. Blockchain allows for near-instant transfers with minimal fees. Figure 1 illustrates the international transfer process using both the traditional banking system and blockchain technology.

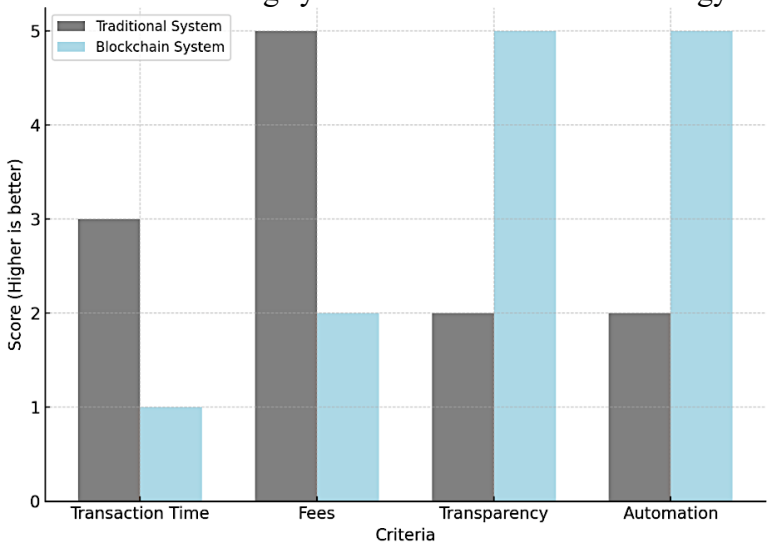

Figure 1. Comparison of international transfer processing in traditional and blockchain systems

The figure demonstrates the differences in processing international transfers between a traditional system, which includes several intermediaries, and a blockchain system, where transactions occur directly between participants. This greatly reduces costs and shortens processing time [2].

Additionally, blockchain also supports the use of smart contracts—algorithmic contracts that execute predefined terms without third-party involvement. In the financial sector, this can include the automation of credit operations, such as automatic loan issuance and payment tracking, which reduces operational costs and minimizes the risk of errors. Smart contracts eliminate a significant number of intermediaries, making processes more efficient and faster.

Using blockchain for identity verification is also an emerging trend in the financial sector. Blockchain enables secure and immutable client data storage, which simplifies verification processes and reduces fraud risks [3]. Implementing a distributed ledger system for data verification expedites identity verification during account opening, loan issuance, or other operations.

Advantages and challenges of implementing blockchain technology in the financial sector

One of blockchain's key advantages is decentralization, which makes the system more resilient and reduces dependence on central authorities. In the financial sector, this opens up new possibilities for creating independent payment systems where transaction control is distributed among network

participants. This approach increases system reliability and reduces risks associated with human error or abuse of power. This is particularly important for data protection and financial information, making blockchain attractive to financial institutions.

Blockchain also enables the use of smart contracts, a key technological element for automating business processes. Smart contracts are programmable agreements that automatically execute under specific conditions, reducing the need for intermediaries and minimizing the risk of unforeseen contract changes [4]. In the financial sector, smart contracts can be valuable for automatic payment calculation and distribution, especially in insurance and lending.

Data transparency, provided by blockchain, is another significant advantage for the financial sector. In a blockchain system, each transaction is accessible to all network participants, which eliminates unauthorized changes and minimizes fraud risk. This feature is particularly useful for organizations that work with large volumes of data and require a transparent system where every transaction is traceable. For example, blockchain can ensure complete transparency and data protection during audits and transaction verifications.

However, despite these advantages, blockchain faces several challenges when implemented in traditional financial structures. One key challenge is scalability: as the network grows and transaction volume increases, processing times may lengthen, which could limit blockchain's application in large-scale projects [5, 6]. Financial organizations considering blockchain must factor in the potential increase in data processing time and plan additional solutions to maintain operational speed.

Another important aspect is regulatory compliance, as blockchain lacks a central regulatory body. Financial institutions must adapt their processes to regulatory requirements that can vary depending on the country and the specific blockchain application area. This adds complexity to implementing this technology in large banking structures, where each transaction must comply with regulatory standards. In some cases, this could be a significant barrier to blockchain use in financial organizations.

Finally, security remains a significant concern when using blockchain [7-9]. Although this technology is highly resistant to hacking, blockchain networks are still susceptible to certain types of attacks. For example, a 51% attack (when a majority of the network is controlled by a single entity) could allow attackers to alter the data ledger. Therefore, financial organizations must consider the possibility of such threats and develop additional security protocols to protect their data and financial information.

Conclusion

Blockchain technology in the financial sector offers unique opportunities to enhance operational efficiency, security, and transparency. Blockchain's decentralized structure eliminates the need for intermediaries and reduces costs, thus optimizing processes like international transfers and data verification. The smart contract technology opens possibilities for automating financial operations, significantly simplifying and accelerating business processes.

However, along with its advantages, blockchain technology faces significant challenges that limit its large-scale application. The primary issues include scalability, security, and regulatory compliance. Financial institutions must consider these factors when integrating blockchain into their systems to minimize possible risks and protect client data.

Thus, blockchain's potential in the financial sector is substantial, but its realization requires a cautious and balanced approach. Successful technology implementation necessitates further research and the development of security standards, as well as training specialists who can effectively apply this technology in practical work.

References

1. Khuryanova I.V., Abibullaev M.S. Opportunities and prospects for using blockchain in insurance // Scientific Notes of the Crimean Federal University named after V.I. Vernadsky. Economics and Management. 2018. Vol. 4. No.3. P. 101-107.

2. Grebenskina A., Zubarev A. Prospects for using smart contracts in the financial sector // Economic Development of Russia. 2018. Vol. 25. No.12. P. 32-43.

3. Ushakov D.S., Podolskaya T.V., Sysoeva A.A. Analysis of the potential of blockchain technology in the modern global economy // State and Municipal Administration. Scientific Notes. 2019. No.1. P. 151-160.
4. Ilyasov A.A., Lebedeva M.E. Blockchain technology and cryptocurrencies in cross-border capital movement: areas of application, potential, and risks // Scientific Notes of the International Banking Institute. 2018. No.1. P. 108-119.
5. Landsman A.Ya., Zakharov S.P., Garaeva A.F. Introduction of new technologies in the banking sector using blockchain // Audit. 2019. No.10. P. 41-43.
6. Apotova N.V., Korolev O.L., Krulikovsky A.P. Analysis of the impact of blockchain technology on the financial system // π-Economy. 2017. Vol. 10. No.6. P. 31-39.
7. Platonov V.V., Spiridonov G.I. Problems and prospects for the use of blockchain technologies in enterprise activities // Bulletin of the St. Petersburg State University of Economics. 2021. No.3(129). P. 102-109.
8. Guo Y., Liang C. Blockchain application and outlook in the banking industry // Financial innovation. 2016. Vol. 2. P. 1-12.
9. Javaid M., Haleem A., Singh R.P., Suman R., Khan S. A review of Blockchain Technology applications for financial services // BenchCouncil Transactions on Benchmarks, Standards and Evaluations. 2022. Vol. 2. No.3. P. 100073.

UDC 004.657

BIG DATA PROCESSING METHODS IN CONSUMER BEHAVIOR ANALYSIS

Mustafayev T.
master's degree, Baku State University (Baku, Azerbaijan)

МЕТОДЫ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ ПРИ АНАЛИЗЕ ПОВЕДЕНИЯ ПОТРЕБИТЕЛЕЙ

Мустафаев Т.М.
магистр, Бакинский государственный университет (Баку, Азербайджан)

Abstract

This article examines modern big data processing methods applied to consumer behavior analysis. The primary focus is on machine learning and deep learning techniques for consumer segmentation, text data analysis, and result visualization. Advantages of using these methods in marketing research, such as increased prediction accuracy and cost optimization, are discussed. The challenges associated with data quality and the need for skilled professionals are also highlighted. It is concluded that applying big data processing methods allows companies to adapt marketing strategies and enhance customer service, providing a long-term competitive advantage.

Keywords: big data, consumer behavior, machine learning, segmentation, marketing.

Аннотация

В данной статье рассматриваются современные методы обработки больших данных, используемые для анализа потребительского поведения. Основное внимание уделяется применению машинного обучения и глубокого обучения для сегментации потребителей, анализа текстовых данных и визуализации результатов. Рассматриваются преимущества использования этих методов в маркетинговых исследованиях, такие как повышение точности прогнозов и оптимизация затрат. Также обсуждаются вызовы, связанные с качеством исходных данных и необходимостью высокой квалификации специалистов. Выявлено, что использование методов обработки больших данных позволяет компаниям адаптировать маркетинговые стратегии и улучшать клиентский сервис, обеспечивая долгосрочное конкурентное преимущество.

Ключевые слова: большие данные, потребительское поведение, машинное обучение, сегментация, маркетинг.

Introduction

Modern big data processing technologies play a crucial role in analyzing consumer behavior, providing companies with the ability to understand and predict customer preferences based on large volumes of heterogeneous information. Consumer data can include structured information, such as purchase history, as well as unstructured data from social networks, reviews, and other sources, making analysis particularly complex. As a result, it becomes essential for companies to apply advanced methods and algorithms that enable efficient processing of large datasets and the identification of patterns in consumer behavior.

Big data processing methods not only accelerate analysis processes but also provide more accurate forecasts through machine learning (ML) and deep learning (DL) algorithms. These technologies help uncover hidden relationships in data, predict user preferences, and adapt marketing

strategies to individual needs. For instance, ML algorithms can analyze textual data from social media, extracting keywords and sentiments associated with a particular product or brand.

The purpose of this article is to review the methods of big data processing used in analyzing consumer behavior and assess their effectiveness. The article will analyze modern approaches, including ML and DL, as well as techniques for data segmentation and clustering that help marketing specialists gain deeper insights into target audience behavior and make more informed business decisions.

Main part

Machine learning methods play a central role in processing and analyzing consumer data. One popular approach is to use clustering algorithms to segment users based on various characteristics, such as preferences, demographic data, and purchase frequency. For example, the K-means algorithm allows for partitioning data into clusters, grouping users with similar characteristics. Figure 1 shows an example of using clustering for consumer segmentation [1].

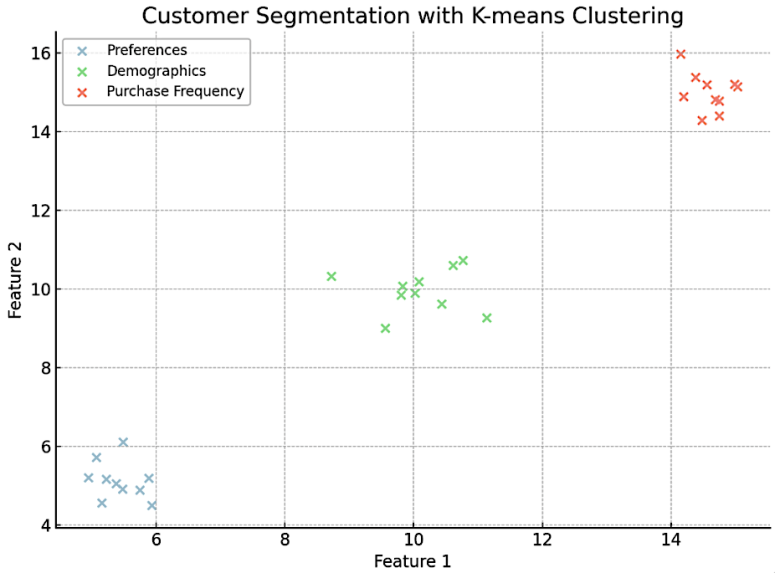

Figure 1. Consumer segmentation using K-means method

The figure illustrates the clustering process, where users are distributed into groups based on similar characteristics. Each group is highlighted in a specific color, allowing for a visual separation of consumers by preferences and other criteria. In addition to clustering, ML enables the application of text analysis methods for processing user reviews and comments. This includes algorithms such as topic modeling and sentiment analysis [2]. Topic modeling (e.g., the LDA algorithm) allows for identifying the main themes discussed in reviews, helping companies understand which aspects of products are most important to customers. Sentiment analysis, on the other hand, provides an assessment of the overall sentiment of the audience towards a product or brand.

Another data processing method is deep learning, which is especially effective for working with images and unstructured data. For instance, in analyzing consumer behavior, DL can be used for face and object recognition in photos, enabling the identification of user preferences. DL algorithms, such as convolutional neural networks, are employed in image analysis to uncover interests and preferences related to specific products or services.

A crucial stage in analyzing big data is the cleaning and preprocessing of information since data from different sources often contains errors, duplicates, and gaps. Processing methods include normalization and standardization of data, removing duplicate records, and filling in missing values. These actions are necessary to enhance the accuracy of ML and DL models, as the quality of the source data directly impacts the final analysis results [3].

To visualize data and present analysis results, a wide range of tools and methods is used. One approach is to create graphs and charts that visually represent the patterns discovered in the data. Table 1 shows the main stages of data analysis when applying ML and DL in marketing research.

Table 1 [4]

Stages of data analysis in marketing research using ML and DL

Stage	Description
Data collection	Obtaining data from various sources, such as social networks, CRM systems, online platforms
Data cleaning	Removing duplicate records, processing missing values
Data processing	Normalizing, standardizing data, and preparing it for analysis
Application of ML and DL	Using algorithms to identify patterns and segment consumers
Visualization and interpretation	Creating graphs and charts to present results

These methods and stages of data analysis enable companies to gain deeper insights into customer needs, improve product quality, and adapt marketing strategies based on data obtained from various sources [5-7].

Prospects for using big data processing methods in marketing

The use of big data processing methods provides marketers with the opportunity to more accurately tailor strategies and offer personalized solutions for each client. By analyzing vast amounts of data, companies can identify unique patterns in consumer behavior, anticipating their future needs. This not only enhances customer satisfaction but also increases business profitability through targeted offers and optimized advertising.

Another important aspect is optimizing marketing expenses. Machine learning and deep learning methods allow for predicting the effectiveness of various advertising channels, identifying the most productive strategies. Thus, companies can optimize their budgets by allocating resources to the most profitable market segments [8]. This is especially crucial in conditions of high competition and rapidly changing customer needs, where the accuracy of marketing decisions plays a key role.

Finally, the development of data processing technologies opens up prospects for a deeper understanding of the impact of external factors on consumer behavior. For example, algorithms can take into account seasonal fluctuations, economic changes, and social trends that affect demand. This enables companies to anticipate potential shifts in audience interests and adapt their offerings in advance, remaining competitive in the long term [9, 10].

Conclusion

Big data processing methods play a critically important role in analyzing and predicting consumer behavior, helping companies adapt their marketing strategies to changing audience preferences. The application of machine learning and deep learning technologies allows for uncovering hidden relationships and predicting future customer behavior, providing marketers with a significant advantage in a highly competitive environment. The results of data analysis help not only to better understand customer needs but also to optimize advertising costs.

Despite the clear advantages, the implementation of big data processing methods in marketing research requires significant resources and skills in working with analytical tools and algorithms. Additionally, the quality of the source data remains a determining factor for the successful application of machine learning and deep learning, as errors and gaps in the data can negatively affect the accuracy of results. However, with proper preparation and adherence to standards, companies can successfully leverage data to gain competitive advantages.

The use of big data methods to analyze consumer behavior represents a promising direction for marketing development. Further research in this area will contribute to the creation of more accurate and robust analysis models, allowing for a deeper exploration of customer behavior and the adaptation of offerings in line with their expectations.

References

1. Checharin E.E. Big data: big problems // Prospects for Science and Education. 2016. No.3(21). P. 7-11.
2. Bolbakov R.G. Big data in information sciences // Educational Resources and Technologies. 2017. No.1(18). P. 30-35.
3. Martyschenko N.S. Methodological support for analyzing consumer behavior in the regional tourism market // Proceedings of the Far Eastern Federal University. Economics and Management. 2005. No.4. P. 19-31.
4. Nazarenko M.A. Transformation of the concept of "quality" in the era of big data // Quality Management Methods. 2016. No.7. P. 36-41.
5. Lobanov A.A. Big data: processing issues // Herald of MGTU MIREA "MSTU MIREA HERALD". 2014. No.3. P. 50.
6. Maltsyeva S.V., Lazarev V.V. Marketing analytics in the field of e-business based on big data // Information Technologies in Design and Production. 2015. No.1. P. 62-67.
7. Gorelova A.A. Big data and directions for their use in marketing // Current Problems of Humanities and Natural Sciences. 2017. No.4-2. P. 11-16.
8. Ivanchenco O.V. Intelligent analysis of big data in the development of relationship marketing in the banking sector // Regional Problems of Economic Transformation. 2019. No.10(108). P. 283-288.
9. Zhang C., Tan T. The impact of big data analysis on consumer behavior // Journal of Physics: Conference Series. IOP Publishing. 2020. Vol. 1544. No.1. P. 012165.
10. Khade A.A. Performing customer behavior analysis using big data analytics // Procedia computer science. 2016. Vol. 79. P. 986-992.

UDC 004.77

INTERNET OF THINGS TECHNOLOGIES FOR INDUSTRIAL PROCESS OPTIMIZATION

Jargalova D.
bachelor's degree, Mirzo Ulughbek National University of Uzbekistan (Tashkent, Uzbekistan)

ТЕХНОЛОГИИ ИНТЕРНЕТА ВЕЩЕЙ ДЛЯ ОПТИМИЗАЦИИ ПРОМЫШЛЕННЫХ ПРОЦЕССОВ

Жаргалова Д.Т.
бакалавр, Узбекистанский национальный университет имени Мирзо Улугбека (Ташкент, Узбекистан)

Abstract

The article explores the potential of Internet of Things (IoT) technologies for optimizing industrial processes. Key IoT advantages, such as real-time data monitoring, enhanced equipment condition control, and predictive maintenance, are described. Emphasis is placed on predictive maintenance using machine learning methods to forecast failures and schedule preventive actions. Additionally, energy efficiency improvements and inventory management optimization enabled by IoT are discussed. The conclusion highlights IoT's importance for industrial digital transformation and improving business competitiveness.

Keywords: Internet of Things, industry, predictive maintenance, energy efficiency, process optimization.

Аннотация

В статье рассматриваются возможности использования технологий Интернета вещей (IoT) для оптимизации промышленных процессов. Описаны основные преимущества IoT, такие как мониторинг данных в реальном времени, улучшение контроля состояния оборудования и поддержка прогнозного обслуживания. Особое внимание уделяется прогнозному обслуживанию, основанному на методах машинного обучения, позволяющему предсказывать неисправности и проводить профилактические работы. Также обсуждаются аспекты повышения энергоэффективности и оптимизации управления запасами с помощью IoT. В заключении подчеркивается значимость IoT для цифровой трансформации промышленности и повышения конкурентоспособности предприятий.

Ключевые слова: Интернет вещей, промышленность, прогнозное обслуживание, энергоэффективность, оптимизация процессов.

Introduction

Internet of Things (IoT) technologies are increasingly being applied in industry, providing opportunities for automation, monitoring, and analysis of production processes. IoT enables devices to exchange data, forming an interconnected network that helps optimize production lines, equipment management, and quality control processes. This technology is becoming the foundation for creating "smart" manufacturing systems capable of responding rapidly to real-time changes, thereby enhancing efficiency and reducing costs. One of the key advantages of IoT in the industrial sector is the ability to continuously collect and analyze data on equipment conditions and production processes. By using sensors and other IoT devices, parameters such as temperature, humidity, machinery speed, and many other indicators that can affect product quality can be monitored. This

data is used to create predictive models that help prevent breakdowns and production downtime. Thus, IoT contributes to the development of predictive maintenance systems that minimize risks and extend equipment lifespan.

The aim of this article is to explore ways to apply IoT for optimizing industrial processes and analyzing their effectiveness. The article will examine examples of IoT usage in manufacturing, as well as methods that facilitate automation, enhance management accuracy, and improve product quality control.

Main part

The implementation of IoT in manufacturing enterprises allows for real-time data collection regarding equipment performance and production processes. This is achieved by installing sensors and connected devices on key equipment elements that record parameters such as temperature, vibrations, wear level, and other indicators. For instance, vibration sensors can detect early signs of bearing or motor failure, allowing the enterprise to conduct preventive maintenance and avoid downtime. IoT systems enable the monitoring of key parameters of production equipment, such as temperature, vibration, noise level, and pressure.

Figure 1 shows the growth of IoT device usage in industrial applications over the past five years. This demonstrates the increasing adoption of IoT technologies in manufacturing processes.

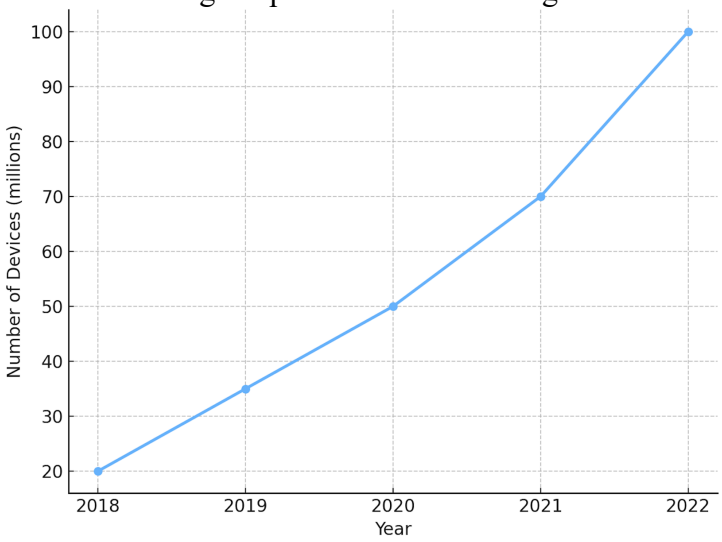

Figure 1. Growth of IoT device usage in industry

As shown in Figure 1, the number of IoT devices used in industrial settings has grown significantly, reaching 100 million by 2022. This trend reflects the increasing reliance on IoT for data collection and analysis.

Temperature sensors help monitor equipment, preventing overheating, which is especially important for reliably operating machinery. Vibration sensors allow for the timely detection of potential failures associated with mechanical wear of components [1]. Noise level control helps comply with safety standards, which is crucial for ensuring comfortable working conditions. Pressure is also continuously monitored to maintain optimal production conditions and exclude deviations that may affect product quality. IoT collects this data in real-time, allowing enterprises to monitor and analyze key performance indicators of production processes.

Another important function of IoT is supporting predictive maintenance systems based on historical data analysis. Predictive maintenance allows for early identification of potential equipment issues and timely preventive actions. Such systems utilize machine learning algorithms to analyze accumulated data and predict the likelihood of failures. This helps reduce repair costs and minimize downtime, which is particularly important for industrial enterprises with continuous production cycles [2]. Additionally, IoT facilitates the optimization of inventory management for raw materials and components. Connected devices can track the number of materials available in real-time and automatically generate orders when supplies approach minimum levels. This reduces the likelihood of production disruptions due to material shortages and allows for more efficient purchasing planning, eliminating excess storage costs.

Another significant area is enhancing energy efficiency. IoT allows for the monitoring of electricity consumption at each stage of production and identifies areas with the highest costs. For example, data analysis can optimize the operating schedule of energy-intensive equipment, reducing energy consumption during low-demand periods [3]. This approach helps lower electricity costs and aligns with modern environmental standards. The use of IoT also significantly improves product quality control systems. By installing sensors on production lines, it is possible to track product parameters in real-time, such as dimensions, weight, temperature, or density, and promptly identify deviations from standards. Implementing such systems allows for the timely detection and correction of defects at early stages, thereby improving the quality of the finished products and reducing the amount of waste.

Predictive maintenance using IoT

One of the significant advantages of implementing IoT in industrial production is the opportunity to realize predictive maintenance based on accumulated data analysis [4]. By utilizing data collected from sensors, IoT systems can predict potential equipment failures before they actually occur. This enables optimization of preventive maintenance scheduling and avoidance of unscheduled downtimes that can lead to financial losses. Predictive maintenance is achieved through the application of machine learning methods and statistical analysis. Analyzing historical data regarding equipment conditions helps identify patterns indicating the likelihood of breakdowns. For instance, data regarding temperature and vibrations may signal increased wear of bearings or other moving parts. Such models allow for the planning of necessary maintenance actions in advance, reducing repair costs and extending equipment operating life. Moreover, predictive maintenance aids in resource optimization. Through an automated fault prediction process, enterprises can more effectively allocate labor and materials, avoiding excess inventory and unnecessary repairs. Implementing IoT and predictive models creates a flexible maintenance system that takes into account the current state of equipment and adapts to real production conditions [5].

Challenges and risks of IoT implementation in industrial production

Despite the significant benefits of IoT in industrial production, its implementation faces several challenges and risks that must be considered. One of the main challenges is ensuring data security and protecting against cyber threats. IoT systems collect and transmit vast amounts of data, often involving sensitive information related to production processes. Unauthorized access or data breaches can lead to production disruptions, financial losses, and intellectual property theft. Therefore, implementing robust security measures, such as encryption and secure data transfer protocols, is essential to safeguard IoT infrastructure in industrial environments [6].

Another challenge is the high initial cost associated with IoT deployment. Installing sensors, upgrading equipment, and establishing data storage and analysis systems require significant financial investments, which may be prohibitive for small or medium-sized enterprises. Moreover, ongoing maintenance and updates to IoT infrastructure can also contribute to costs [7, 8]. To address this, companies need to evaluate the return on investment by considering long-term benefits such as reduced downtime, increased productivity, and lower operational costs, which can offset the initial expenses over time [9].

Finally, IoT implementation demands a skilled workforce capable of managing and analyzing data generated by interconnected devices. Training personnel to work with IoT systems and interpret data analytics is critical for maximizing the value of IoT. Additionally, integrating IoT into existing production workflows may require adjustments and adaptations, as legacy systems may not always be compatible with IoT technologies. Addressing these challenges requires comprehensive planning and a phased approach to IoT deployment to minimize disruption and maximize benefits for industrial enterprises.

Conclusion

The implementation of IoT technologies in industrial production opens up significant prospects for improving efficiency and quality of production processes. IoT allows enterprises to collect real-time data, which enhances equipment condition monitoring and provides opportunities for predictive maintenance. This contributes to reduced downtime and lower repair costs, enabling enterprises to

operate more efficiently. Predictive maintenance, supported by IoT, is becoming an integral part of industrial asset management, as data analysis allows for anticipating potential failures and conducting preventive maintenance at the right time. Such technologies, including machine learning methods, ensure equipment longevity and reduce the risk of emergency situations, which is particularly important for high-load enterprises with continuous production cycles. The use of IoT also aids in resource optimization and reduces costs for energy and materials. Monitoring energy consumption metrics and inventory status enables more accurate management of production processes and minimizes expenses. Thus, IoT is a key element facilitating digital transformation and enhancing the competitiveness of industrial enterprises.

References

1. Gordienko E.P. Internet of Things: history of emergence and prospects for application // Actual problems and prospects for the development of transport, industry, and economy of Russia ("TransPromEx-2023"). 2023. P. 40.
2. Petrov V.Yu., Rudashevskaya E.A. Internet of Things technology as a promising modern information technology // Fundamental Research. 2017. No.9-2. P. 471-476.
3. Kupriyanovsky V.P., Namiot D.E., Drozhzhinov V.I., Kupriyanovskaya Yu.V., Ivanov M.O. Internet of Things in industrial enterprises // International Journal of Open Information Technologies. 2016. Vol. 4. No.12. P. 69-78.
4. Cherapanov N.V. Industrial Internet of Things in enterprises // Innovations and Investments. 2019. No.10. P. 151-154.
5. Paskova A.A. Integration of artificial intelligence technologies and industrial Internet of Things // Issues of sustainable development of society. 2020. No.4-2. P. 607-612.
6. Xu S., Chen J., Wu M., Zhao C. E-commerce supply chain process optimization based on whole-process sharing of internet of things identification technology // Computer Modeling in Engineering & Sciences. 2021. Vol. 126. No.2. P. 843-854.
7. Chui M., Loffler M., Roberts R. The internet of things. McKinsey Global Institute. 2010.
8. Huang X. Quality of service optimization in wireless transmission of industrial Internet of Things for intelligent manufacturing // The International Journal of Advanced Manufacturing Technology. 2020. Vol. 107. No.3. P. 1007-1016.
9. Soori M., Arezoo B., Dastres R. Internet of things for smart factories in industry 4.0, a review // Internet of Things and Cyber-Physical Systems. 2023. Vol. 3. P. 192-204.

УДК 004.4

СОВРЕМЕННЫЕ МЕТОДЫ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ВЫСОКОНАГРУЖЕННЫХ СИСТЕМ

Рахматуллаев О.С.
специалист, Ташкентский университет информационных технологий (Ташкент, Узбекистан)

MODERN SOFTWARE DEVELOPMENT METHODS FOR HIGH-LOAD SYSTEMS

Rakhmatullayev O.
specialist's degree, Tashkent University of Information Technologies (Tashkent, Uzbekistan)

Аннотация

В статье исследуются современные методы разработки программного обеспечения для высоконагруженных систем (ВНС), включая микросервисную архитектуру и асинхронное программирование. Описаны преимущества этих подходов, такие как повышение производительности, отказоустойчивость и возможность масштабирования отдельных компонентов системы. Представлены примеры использования асинхронных операций для параллельной обработки данных, что позволяет уменьшить время отклика и повысить стабильность системы. В заключении обсуждается значимость этих методов для создания устойчивых ВНС, способных выдерживать большие объемы данных и справляться с интенсивной нагрузкой.

Ключевые слова: высоконагруженные системы, асинхронное программирование, микросервисы, масштабируемость, отказоустойчивость.

Abstract

The article examines modern software development methods for high-load systems (HLS), including microservice architecture and asynchronous programming. The advantages of these approaches, such as increased performance, fault tolerance, and the scalability of individual system components, are described. Examples of asynchronous operations for parallel data processing are provided, allowing for reduced response time and improved system stability. In conclusion, the significance of these methods for creating resilient HLS capable of handling large data volumes and intense workloads is discussed.

Keywords: high-load systems, asynchronous programming, microservices, scalability, fault tolerance.

Введение

Современные высоконагруженные системы (ВНС) требуют использования эффективных методов разработки программного обеспечения, способных обрабатывать большие объемы данных и обеспечивать стабильность при высокой интенсивности запросов. ВНС охватывают различные сферы, включая финтех, онлайн-торговлю, стриминг и телекоммуникации, где устойчивость и высокая производительность программного обеспечения являются критически важными. Эти системы должны обеспечивать низкую задержку, быстрое время отклика и высокую степень масштабируемости, что обуславливает необходимость использования передовых технологий и подходов к разработке.

В последние годы ключевые методики разработки для ВНС значительно изменились, включив в себя такие подходы, как микросервисная архитектура, асинхронное программирование и использование технологий кэширования. Микросервисы позволяют разделить систему на небольшие независимые модули, каждый из которых отвечает за определенную функцию, что упрощает их разработку, тестирование и масштабирование. Асинхронное программирование, в свою очередь, позволяет оптимизировать обработку запросов, избегая блокировки потоков и улучшая производительность системы при интенсивной нагрузке.

Цель данной статьи – исследовать современные методы разработки программного обеспечения для ВНС, а также проанализировать их преимущества и недостатки. В статье будут рассмотрены ключевые подходы, такие как микросервисная архитектура, асинхронное программирование и использование распределенных баз данных, которые позволяют добиться высокой производительности и стабильности работы программных систем при значительной нагрузке.

Основная часть

Микросервисная архитектура является одной из наиболее распространенных моделей разработки для ВНС. В отличие от монолитных приложений, где все компоненты системы тесно интегрированы, микросервисы представляют собой набор независимых модулей, каждый из которых выполняет одну функцию. Это позволяет легко масштабировать каждый микросервис, повышая гибкость и отказоустойчивость системы. Например, для увеличения производительности можно масштабировать только те микросервисы, которые подвергаются наибольшей нагрузке, не затрагивая остальные части системы [1].

Асинхронное программирование также является важным методом оптимизации производительности в ВНС. Оно позволяет системе обрабатывать несколько задач одновременно, избегая блокировки потоков при выполнении длительных операций, таких как запросы к базе данных или вызовы внешних сервисов [2]. В асинхронном программировании часто используются операторы **async** и **await**, что делает код более гибким и позволяет ускорить обработку данных. Пример кода с использованием асинхронного программирования на Python:

```
import asyncio

async def fetch_data():
    print("Запрос данных...")
    await asyncio.sleep(2)  # Имитация сетевого запроса
    print("Данные получены")
    return {"data": "результат"}

async def process_data():
    print("Обработка данных...")
    data = await fetch_data()
    print("Обработанные данные:", data)

# Запуск асинхронных задач
asyncio.run(process_data())
```

В приведенном примере демонстрируется использование асинхронных функций в Python для обработки данных без блокировки основного потока. Такой подход позволяет системе оставаться отзывчивой, так как выполнение задачи происходит параллельно, что снижает общую нагрузку на систему.

Асинхронность в ВНС

Асинхронное программирование является одной из ключевых технологий, обеспечивающих высокую производительность в высоконагруженных системах (ВНС). Асинхронность позволяет системе выполнять несколько операций одновременно, избегая

блокировки основного потока при выполнении длительных операций. Это особенно важно в случаях, когда ВНС обрабатывает тысячи или миллионы запросов в секунду, и задержка при выполнении одного запроса может вызвать задержку всей системы [3].

Асинхронное программирование основывается на использовании событийного цикла, который распределяет задачи между потоками, обеспечивая выполнение задач параллельно. В то время как синхронные операции ожидают завершения каждого запроса по очереди, асинхронные задачи передают управление, пока ожидают завершения операции (например, запроса к базе данных или выполнения сетевого вызова), тем самым освобождая ресурс для других операций. Это позволяет обрабатывать значительно больше запросов за единицу времени, что особенно важно для ВНС, где требуется поддерживать высокий уровень отклика даже при пиковых нагрузках [4].

Ключевые элементы асинхронного программирования включают операторы **async** и **await** (в языке Python) или их аналоги в других языках, такие как Promise в JavaScript. Эти операторы позволяют задавать асинхронные функции, которые могут приостанавливать свое выполнение, передавая управление другим задачам и возвращаясь к выполнению по завершении операции. Пример ниже иллюстрирует использование этих операторов для выполнения нескольких асинхронных запросов:

```
import asyncio

# Асинхронная функция для имитации запросов
async def fetch_data(endpoint):
    print(f"Запрос к {endpoint} начат")
    await asyncio.sleep(1)  # Имитация задержки
    print(f"Запрос к {endpoint} завершен")
    return f"Данные с {endpoint}"

# Основная асинхронная функция, выполняющая несколько запросов параллельно
async def main():
    tasks = [
        fetch_data("Сервис 1"),
        fetch_data("Сервис 2"),
        fetch_data("Сервис 3")
    ]
    results = await asyncio.gather(*tasks)  # Параллельное выполнение задач
    print("Результаты:", results)

# Запуск
asyncio.run(main())
```

В данном примере несколько асинхронных задач выполняются одновременно, что иллюстрирует, как система может обрабатывать несколько запросов параллельно. Функция asyncio.gather() собирает задачи и запускает их параллельно, ускоряя общий процесс и оптимизируя использование ресурсов [5, 6]. Преимущества асинхронного программирования включают уменьшение времени отклика и повышение пропускной способности системы. Благодаря тому, что асинхронные функции освобождают ресурсы, пока ожидается завершение операции, нагрузка распределяется более равномерно, что снижает риск задержек и отказов. Асинхронность также позволяет гибко масштабировать систему, адаптируясь к увеличению запросов и требованиям пользователей в реальном времени. Использование асинхронного программирования является эффективным подходом для оптимизации ВНС, обеспечивая высокий уровень стабильности и производительности даже в условиях пиковых нагрузок [7].

Заключение

Асинхронное программирование и микросервисная архитектура стали одними из ключевых инструментов разработки ВНС, способных обеспечивать высокую

производительность и отказоустойчивость. Эти технологии позволяют создавать гибкие, легко масштабируемые системы, где каждый компонент может быть автономно развернут и оптимизирован под конкретные задачи. Микросервисный подход также обеспечивает возможность изолированного масштабирования, что позволяет эффективно распределять ресурсы и снижать риски отказа всей системы.

Асинхронное программирование, основанное на использовании событийного цикла, позволяет системе выполнять параллельные операции, избегая блокировки потоков, что повышает пропускную способность и снижает время отклика. Системы, работающие в режиме реального времени, могут обрабатывать множество запросов одновременно, что минимизирует задержки и обеспечивает стабильную производительность даже при пиковых нагрузках. Этот подход особенно актуален для ВНС, где задержки при выполнении запросов могут негативно сказаться на качестве обслуживания.

Таким образом, методы разработки, включающие микросервисную архитектуру, асинхронное программирование и использование распределенных баз данных, формируют основу для построения высоконадежных и производительных систем, способных эффективно обрабатывать большие объемы данных. Внедрение этих технологий позволяет создавать современные ВНС, готовые к работе в условиях высокой нагрузки и удовлетворяющие потребности бизнеса в устойчивых и масштабируемых решениях.

Список литературы

1. Амиров С.Н. Особенности разработки высоконагруженных систем // International journal of open information technologies. 2020. Т. 8. №8. С. 37-45.
2. Матчин В.Т., Плотников С.Б., Цветков В.Я. Работа с информационными ресурсами в высоконагруженных приложениях // Образовательные ресурсы и технологии. 2020. №4 (33). С. 62-72.
3. Сотников О.О. Смена парадигмы разработки программных продуктов с переходом от человеческого труда на искусственный интеллект // Аллея науки. 2020. Т. 2. №5 (44). С. 967.
4. Зиборев А.В. Антипаттерны построения микросервисных приложений в высоконагруженных проектах // Universum: технические науки. 2023. №11-1 (116). С. 29-34.
5. Болодурина И.П., Парфёнов Д.И. Моделирование управляющих потоков данных высоконагруженных информационных систем, построенных на базе облачных вычислений // Интеллект. Инновации. Инвестиции. 2014. №4. С. 93-101.
6. Глумов К.С. Преимущества реактивного подхода в высоконагруженных микросервисных системах // Вестник науки. 2024. Т. 1. №10 (79). С. 410-426.
7. Kryvenchuk Y., Mykalov P., Novytskyi Y., Zakharchuk M., Malynovskyy Y., Řepka M. Analysis of the architecture of distributed systems for the reduction of loading high-load networks // Conference on Computer Science and Information Technologies. Cham : Springer International Publishing. 2019. P. 759-770.

УДК 004.056

МОДЕЛИ ПРЕДСКАЗАНИЯ КИБЕРАТАК С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Киселев Р.Г.
магистр, Южный федеральный университет
(Ростов-на-Дону, Россия)

CYBERATTACK PREDICTION MODELS USING MACHINE LEARNING

Kiselev R.
master's degree, Southern Federal University
(Rostov-on-Don, Russia)

Аннотация

В статье рассматриваются методы машинного обучения (МО) для предсказания кибератак и повышения уровня кибербезопасности. Основное внимание уделено алгоритмам классификации и методам обнаружения аномалий, которые позволяют выявлять как известные, так и новые типы атак. Приведены примеры применения классификационных моделей, таких как случайные леса, и методов обнаружения аномалий, таких как изоляционный лес, для анализа сетевой активности. Также обсуждаются перспективы использования глубокого обучения и методов МО с подкреплением для анализа временных рядов и динамических данных. В заключении подчеркивается значимость данных подходов для обеспечения устойчивости информационных систем в условиях постоянно меняющейся киберугрозы.

Ключевые слова: кибератаки, машинное обучение, обнаружение аномалий, классификация, кибербезопасность.

Abstract

The article examines machine learning (ML) methods for predicting cyberattacks and enhancing cybersecurity. The focus is on classification algorithms and anomaly detection techniques that identify both known and unknown types of attacks. Examples of applying classification models, such as random forests, and anomaly detection methods, such as isolation forests, for network activity analysis are provided. The prospects for using deep learning and reinforcement learning methods to analyze time series and dynamic data are also discussed. The conclusion emphasizes the importance of these approaches for maintaining the resilience of information systems amid constantly evolving cyber threats.

Keywords: cyberattacks, machine learning, anomaly detection, classification, cybersecurity.

Введение

С развитием технологий и расширением использования цифровых инфраструктур возрастает и число кибератак, что представляет серьезную угрозу для безопасности данных и информационных систем. Традиционные методы защиты, такие как системы предотвращения вторжений и антивирусные программы, зачастую не способны эффективно противостоять новым видам атак. В условиях, когда злоумышленники используют продвинутые методы и инструменты, необходимы более гибкие и эффективные подходы для защиты данных. Одним из таких методов является использование алгоритмов машинного обучения (МО), которые способны анализировать большие объемы данных, выявлять аномалии и предсказывать потенциальные атаки.

Модели предсказания кибератак, основанные на МО, позволяют обнаруживать подозрительную активность, анализируя сетевые пакеты, логи и события системы безопасности. Алгоритмы, такие как деревья решений, случайные леса и нейронные сети, могут обучаться на исторических данных, чтобы определять паттерны, характерные для кибератак. На основе этих моделей можно прогнозировать вероятность атаки, что позволяет предпринять меры по защите системы до момента реального вторжения. Преимущество МО заключается в способности обнаруживать новые типы атак, что значительно расширяет возможности традиционных методов безопасности.

Цель данной статьи состоит в рассмотрении различных моделей МО, применяемых для предсказания кибератак, а также анализа их эффективности и применимости в различных условиях. В статье будут описаны основные методы, включая алгоритмы классификации и обнаружения аномалий, а также представлены примеры их практического применения для повышения уровня кибербезопасности.

Основная часть

Алгоритмы классификации, такие как логистическая регрессия, деревья решений и случайные леса, часто применяются для определения, принадлежит ли сетевое событие к категории нормальной активности или кибератаки.

На рисунке 1 представлено распределение типов кибератак, с которыми чаще всего сталкиваются информационные системы. Данные подчеркивают разнообразие угроз и необходимость применения алгоритмов машинного обучения для их обнаружения.

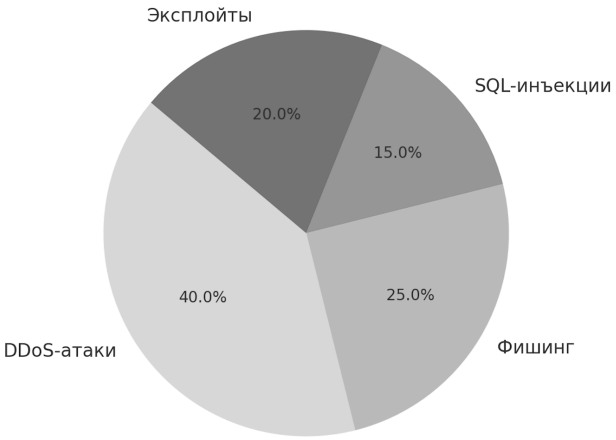

Рисунок 1. Распределение типов кибератак

Как видно из рисунка 1, наиболее частыми типами атак являются DDoS-атаки (40%) и фишинг (25%), что связано с их относительной простотой и высокой эффективностью для злоумышленников. Реже встречаются сложные атаки, такие как SQL-инъекции (15%) и кража данных через эксплойты (20%). Эти данные подтверждают важность разработки универсальных алгоритмов, способных распознавать как стандартные, так и более сложные типы угроз.

Эти модели обучаются на большом объеме данных, содержащих метки классов (например, «атака» и «норма»), что позволяет алгоритму обнаруживать паттерны, характерные для атакующих действий. Примером такого подхода может служить случайный лес, который строит несколько деревьев решений и принимает итоговое решение на основе их совокупного результата [1]. Пример кода на Python для классификации сетевых событий с использованием случайного леса:

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

# Пример данных, где X - признаки, а y - метки классов (атака или норма)
# X может включать такие параметры, как количество запросов в секунду, IP-адрес, тип пакета и т.д.
```

```
X = [[200, 0.5, 10], [100, 0.1, 5], [300, 0.7, 8], [50, 0.2, 3]]
y = ["норма", "атака", "атака", "норма"]

# Разделение данных на обучающую и тестовую выборки
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25, random_state=42)

# Инициализация и обучение модели случайного леса
clf = RandomForestClassifier(n_estimators=10, random_state=42)
clf.fit(X_train, y_train)

# Предсказание на тестовой выборке
y_pred = clf.predict(X_test)

# Вывод точности модели
accuracy = accuracy_score(y_test, y_pred)
print(f"Точность классификации: {accuracy * 100:.2f}%")
```

В данном примере используется модель случайного леса для классификации событий на «атаку» и «норму». Модель обучается на небольшом наборе данных, включающем параметры сетевой активности, и затем применяется для предсказания на тестовой выборке. Функция accuracy_score позволяет оценить, насколько точно модель определяет атаки, что является важным показателем ее эффективности [2]. Использование таких алгоритмов классификации дает возможность системам безопасности автоматически реагировать на потенциальные угрозы, снижая нагрузку на специалистов по кибербезопасности. Случайные леса и другие алгоритмы классификации могут эффективно выявлять кибератаки в условиях высокой динамичности сетевой активности, обеспечивая дополнительный уровень защиты для информационных систем. Методы обнаружения аномалий представляют собой еще один важный инструмент в предсказании кибератак. В отличие от классификационных алгоритмов, которые требуют размеченных данных для обучения, методы обнаружения аномалий работают с неразмеченными данными и позволяют выявлять неизвестные типы атак. Это особенно важно для ситуаций, когда системы безопасности сталкиваются с новыми, ранее неизвестными угрозами [3]. Примером подобного подхода является использование изоляционного леса, который строит несколько деревьев для выявления аномальных записей в данных, что позволяет обнаруживать возможные признаки кибератак. Также важную роль играют ассоциативные правила, которые помогают выявлять последовательности событий, предшествующих кибератакам. Эти алгоритмы позволяют найти связи между различными действиями пользователей или сетевой активностью, что может сигнализировать о подготовке атаки. Например, частое повторение определенных запросов или команд может указывать на подозрительную активность, что позволяет системе безопасности предпринять соответствующие меры до начала атаки.

Глубокое обучение также является перспективным направлением в области предсказания кибератак, особенно когда требуется анализировать большие объемы сложных данных. Нейронные сети способны работать с неструктурированными данными, такими как сетевые журналы и текстовые записи, позволяя системе безопасности находить аномальные паттерны. Например, сверточные нейронные сети можно использовать для анализа сетевых пакетов и поиска подозрительных последовательностей, указывающих на угрозу. Рекуррентные нейронные сети (РНС) особенно полезны для анализа временных рядов сетевой активности. Это дает возможность выявлять последовательные аномалии и связывать их с потенциальной атакой, особенно если атака разворачивается постепенно. Использование РНС позволяет моделировать сложные зависимости между событиями, что делает ее незаменимой в анализе данных с временной составляющей, таких как журналы активности системы [4].

Методы МО с подкреплением также находят применение в предсказании кибератак, позволяя системе обучаться через взаимодействие с реальной средой. Такие модели способны

адаптироваться к быстро меняющимся условиям, что особенно актуально для кибербезопасности, где характер угроз может значительно меняться. Алгоритмы с подкреплением помогают системе оптимизировать свои действия, повышая вероятность обнаружения угроз в режиме реального времени и снижая риски, связанные с возможными атаками.

Применение моделей обнаружения аномалий в предсказании кибератак

Модели обнаружения аномалий играют критически важную роль в выявлении кибератак, так как они позволяют обнаруживать отклонения от нормального поведения системы, указывающие на возможную угрозу. В отличие от классификационных моделей, которые требуют размеченные данные для обучения, методы обнаружения аномалий способны работать с неразмеченными данными, что делает их более универсальными в условиях, когда полный набор данных о кибератаках отсутствует или содержит неизвестные типы угроз. Эти модели подходят для случаев, когда атакующие используют новые техники, еще не встречавшиеся в исторических данных.

Основной принцип работы моделей обнаружения аномалий заключается в создании профиля «нормального» поведения системы на основе сетевых пакетов, активности пользователей или логов системы, после чего любые значительные отклонения от этого профиля классифицируются как аномалии [5]. Например, если в обычной сетевой активности среднее количество запросов к серверу составляет 50 в минуту, а вдруг происходит всплеск до 200 запросов, это может свидетельствовать о начале DDoS-атаки.

Изоляционный лес (Isolation Forest) – это один из алгоритмов, эффективных для обнаружения аномалий в кибербезопасности. Алгоритм строит множество деревьев решений, каждое из которых изолирует случайные точки данных. При этом аномалии (выбросы) изолируются на более низких уровнях дерева, так как для их выделения требуется меньше шагов из-за их редкости или необычности. Такой подход позволяет алгоритму выделять подозрительные события на основе отклонений от нормы без необходимости разметки данных.

Методы обнаружения аномалий обладают способностью выявлять ранее неизвестные типы атак, так как они не ограничены размеченными данными. Однако данные модели чувствительны к качеству исходных данных и профилю «нормального» поведения, на основе которого строятся прогнозы [6]. Например, если модель обучена на данных, содержащих аномалии, они могут быть приняты за норму, что снизит точность предсказаний. Применение моделей обнаружения аномалий для предсказания кибератак позволяет улучшить проактивный подход к безопасности, реагируя на угрозы до их непосредственного проявления [7].

Заключение

Использование алгоритмов МО для предсказания кибератак позволяет повысить эффективность и проактивность систем безопасности. Классификационные алгоритмы, такие как случайные леса, показывают высокую точность в определении известных атак на основе исторических данных, что позволяет системам безопасности оперативно реагировать на угрозы. Эти модели дают возможность автоматизировать обработку и анализ данных, снижая нагрузку на специалистов по безопасности и позволяя оперативно предотвращать потенциальные атаки. Методы обнаружения аномалий также занимают важное место в предсказании кибератак, так как они способны выявлять ранее неизвестные угрозы, что особенно важно в условиях постоянно меняющейся киберугрозы. Такие алгоритмы, как изоляционный лес, помогают выделить отклонения от нормального поведения и предсказать потенциальные угрозы, даже если их паттерны ранее не были зафиксированы в данных. Это позволяет поддерживать высокий уровень безопасности в условиях, когда новые техники атак появляются регулярно.

Список литературы

1. Мишин А.Е., Былевский П.Г. Применение машинного обучения для прогнозирования кибератак // Hi-Hume Journal. 2023. Т. 3. С. 80.

2. Асяев Г.Д., Соколов А.Н. Модели предиктивной защиты информации автоматизированной системы управления водоснабжением на основе временных рядов с использованием технологий машинного обучения // Вестник УрФО. Безопасность в информационной сфере. 2021. №4 (42). С. 39-45.
3. Кечеджиев А.С., Цветкова О.Л., Дубровина А.И. Методика выявления аномалий в данных оценки кибератак с использованием Random Forest и градиентного бустинга в машинном обучении // Вестник Дагестанского государственного технического университета. Технические науки. 2024. Т. 51. №3. С. 72-85.
4. Власенко А.В., Дзьобан П.И., Жук Р.В. Обзор инструментов машинного обучения и их применения в области кибербезопасности // Прикаспийский журнал: управление и высокие технологии. 2020. №1 (49). С. 144-155.
5. Боброва М.В., Мастилин А.Е. Машинное обучение в кибербезопасности // Научные междисциплинарные исследования. 2021. №2. С. 24-29.
6. Котенко И.В., Саенко И.Б., Лаута О.С., Крибель А.М. Методика обнаружения аномалий и кибератак на основе интеграции методов фрактального анализа и машинного обучения // Информатика и автоматизация. 2022. Т. 21. №6. С. 1328-1358.
7. Ibor A.E., Oladeji F.A., Okunoye O.B., Ekabua O.O. Conceptualisation of Cyberattack prediction with deep learning // Cybersecurity. 2020. Vol. 3. P. 1-14.

UDC 658.7

EFFICIENCY OF BLOCKCHAIN PLATFORMS FOR SUPPLY CHAIN MANAGEMENT

Asanova N.
bachelor's degree, Kyrgyz-Russian Slavic University
(Bishkek, Kyrgyzstan)

ЭФФЕКТИВНОСТЬ БЛОКЧЕЙН-ПЛАТФОРМ ДЛЯ УПРАВЛЕНИЯ ЛОГИСТИЧЕСКИМИ ЦЕПОЧКАМИ

Асанова Н.К.
бакалавр, Кыргызско-Российский Славянский университет
(Бишкек, Кыргызстан)

Abstract

The article examines the potential of blockchain platforms for managing and optimizing supply chains. Key benefits of blockchain, such as transparency, reliability, and process automation through smart contracts, are discussed. Successful applications of blockchain by companies like Maersk and Walmart for tracking shipments and ensuring product safety are presented. A Solidity smart contract example demonstrates the automation of logistics operations. The conclusion highlights the significance of blockchain for managing logistics processes, optimizing costs, and securing supply chains.

Keywords: blockchain, logistics, smart contracts, supply chain management, automation.

Аннотация

В статье рассматриваются возможности блокчейн-платформ для управления логистическими цепочками и повышения их эффективности. Описаны основные преимущества блокчейна, включая прозрачность, надежность и автоматизацию процессов с помощью смарт-контрактов. Приведены примеры успешного применения блокчейна компаниями Maersk и Walmart для отслеживания грузов и обеспечения безопасности продуктов. Пример смарт-контракта на языке Solidity иллюстрирует автоматизацию логистических операций. В заключении подчеркивается значимость блокчейна для управления логистическими процессами, оптимизации затрат и обеспечения безопасности цепочек поставок.

Ключевые слова: блокчейн, логистика, смарт-контракты, управление цепочками поставок, автоматизация.

Introduction

The technologies of the Internet of Things (IoT) are increasingly being applied in logistics, providing high levels of transparency and control at every stage from production to the end consumer. With the growing complexity of logistics processes and the involvement of multiple parties such as manufacturers, suppliers, transport companies, and retailers, the need for effective management and monitoring systems has risen. In this context, blockchain technology offers promising opportunities to enhance transparency, reliability, and data security, helping to reduce risks and build trust among participants in the logistics chain.

One of the main advantages of blockchain is its ability to ensure immutability of records, making it a valuable tool for tracking the provenance of goods and preventing counterfeiting.

Blockchain platforms allow real-time recording of data on each stage of the movement of goods, which reduces the risk of errors and simplifies the verification of authenticity. Each transaction in the blockchain is recorded in a distributed ledger, making it accessible to all participants in the logistics chain, thereby increasing transparency and enabling quick detection of violations.

The aim of this article is to explore the effectiveness of blockchain platforms for managing logistics chains. The article will examine the main advantages and challenges of using blockchain in this field, as well as present examples of successful implementations and analyses of results confirming its impact on process optimization and reduction of operational costs.

Main part

Blockchain technology in logistics chain management provides a number of unique opportunities for increasing efficiency and optimizing processes. One of the key aspects is the ability to track the movement of goods in real time. Each participant in the chain can access current information about the location and status of the cargo at any moment, allowing for quicker responses to unforeseen delays or violations of transportation conditions. This is particularly important for perishable goods such as food and pharmaceuticals, where delays can lead to significant losses [1].

The use of smart contracts supported by blockchain platforms presents another advantage for logistics. Smart contracts are self-executing contracts where the terms of the agreement are programmed and automatically executed when certain events occur. In the logistics chain, this can include automatic payment confirmation after the cargo delivery is verified or automatic status updates for the cargo upon passing a specific point in the route [2-4]. Smart contracts reduce the likelihood of errors, simplify interactions between parties, and accelerate transaction execution.

The implementation of blockchain also contributes to lowering operational costs by eliminating intermediaries and automating processes. Traditional logistics management methods often require multiple parties to verify data and use paper documents, increasing costs and processing time. In contrast, blockchain provides a single data repository accessible to all participants, eliminating the need for numerous verifications and duplication of information. This enables companies to reduce administrative process costs and improve efficiency in partner interactions.

Examples of blockchain application in logistics chains

The application of blockchain technology in logistics can be illustrated by companies such as Maersk and Walmart, which actively use blockchain to optimize processes and enhance transparency. In collaboration with IBM, Maersk has implemented the TradeLens blockchain platform, which enables real-time tracking of container shipments and provides a unified information space for all participants in the logistics chain — from manufacturers and suppliers to customs authorities and end customers. TradeLens consolidates transportation data at all stages and allows partners to exchange information promptly, thereby reducing delays and mitigating risks associated with documentation errors [5-7]. Walmart uses blockchain to track the provenance of food products, which is especially important for ensuring food safety and increasing consumer trust. With the blockchain platform developed in collaboration with IBM, Walmart can trace a product's journey from farm to store, allowing for rapid identification of the source of a problem in the event of discovering substandard goods or outbreaks of illness [8, 9]. This significantly reduces the time needed to find a problem and minimizes losses while providing transparency for consumers who want to know the origin and storage conditions of the products. Another example is DHL, which uses blockchain to improve warehouse inventory management and cargo tracking. Blockchain allows for the recording of data on the condition and location of goods, simplifying the inventory process and preventing losses or theft. Blockchain ensures the immutability of records and access to information for all participants in the logistics network, minimizing errors associated with data updates and enabling quick responses to changes [10].

These examples confirm that blockchain can significantly enhance the efficiency of logistics processes, improve transparency, and reduce risks associated with delays, errors, and fraud.

Solidity example code for a smart contract in logistics using Solidity

Smart contracts allow for the automation of processes in logistics chains by executing programmed conditions without the need for third-party involvement. For example, in logistics, a

smart contract can automatically confirm the completion of delivery and transfer payment to the supplier after cargo delivery confirmation [11].

Below is an example of a simple smart contract written in Solidity that models the cargo tracking process and automatic payment transfer upon delivery confirmation.

```
// SPDX-License-Identifier: MIT
pragma solidity ^0.8.0;

contract LogisticsSmartContract {
    address public sender;          // Sender's address
    address public receiver;        // Receiver's address
    address public logisticsCompany; // Logistics company's address
    uint public paymentAmount;      // Payment amount
    bool public isDelivered;        // Delivery status

    // Events for tracking cargo status
    event Shipped();
    event Delivered();
    event PaymentTransferred(address receiver, uint amount);

    constructor(address _receiver, address _logisticsCompany, uint _paymentAmount) {
        sender = msg.sender;
        receiver = _receiver;
        logisticsCompany = _logisticsCompany;
        paymentAmount = _paymentAmount;
        isDelivered = false;
    }

    // Function to ship the cargo
    function ship() public onlySender {
        emit Shipped();
    }

    // Function to confirm delivery by the logistics company
    function confirmDelivery() public onlyLogisticsCompany {
        isDelivered = true;
        emit Delivered();
        transferPayment();
    }

    // Function to transfer payment to the receiver after delivery confirmation
    function transferPayment() internal {
        require(isDelivered, "Delivery not confirmed");
        payable(receiver).transfer(paymentAmount);
        emit PaymentTransferred(receiver, paymentAmount);
    }

    // Restriction for calling function only by the sender
    modifier onlySender() {
        require(msg.sender == sender, "Only sender can call this function");
        _;
    }
```

```
    // Restriction for calling function only by the logistics company
    modifier onlyLogisticsCompany() {
        require(msg.sender == logisticsCompany, "Only logistics company can call this function");
        _;
    }

    // Function to receive funds into the contract balance
    receive() external payable {}
}
```

Conclusion

The application of blockchain technology in logistics chains offers significant advantages such as transparency, reliability, and the potential for process automation. Blockchain allows for the recording of each stage of cargo movement in a distributed ledger, ensuring immutability of records and data accessibility for all participants in the chain. This reduces the risks of loss, errors, and fraud, enhancing the overall efficiency of logistics processes. The use of smart contracts, as illustrated in the example, simplifies interactions between parties by automating operations and minimizing delays associated with data verification and confirmation. Smart contracts ensure the execution of transaction conditions automatically, increasing accuracy and reducing the time spent on administrative procedures. This automation is crucial for speeding up processes and cutting operational costs. Thus, blockchain platforms can significantly transform the approach to managing logistics chains, providing a higher level of transparency, reliability, and efficiency. Successful implementation examples of blockchain, such as the projects by Maersk and Walmart, confirm that distributed ledger technologies can become the foundation for building digital ecosystems where interaction and trust among partners are achieved through automation and data immutability.

References

1. Larin O.N., Bush Yu.D., Nekrutova S.P. Current issues in the application of digital blockchain platforms for transport logistics // Intelligent data analysis and digital economy. 2018. P. 8-22.
2. Larin O.N. Model of interaction of blockchain platforms for logistics ecosystems // Transport: science, technology, management. Scientific information collection. 2021. No.9. P. 26-29.
3. Zolotarev M. Blockchain finds its footing in logistics and supply chains // Transport strategy-XXI century. 2018. No.39. P. 46-47.
4. Bazaka V.V. Blockchain technologies based on "green" logistics // Strategies and tools for economic management: sectoral and regional aspects. 2021. P. 263-265.
5. Nikonova Y.I. Application of blockchain technology in transport and logistics // Journal of Monetary Economics and Management. 2024. No.2. P. 22-24.
6. Madhani P.M. Supply chain transformation with blockchain deployment: enhancing efficiency and effectiveness // IUP Journal of Supply Chain Management. 2021. Vol. 18. No.4. P. 7-32.
7. Park A., Li H. The effect of blockchain technology on supply chain sustainability performances // Sustainability. 2021. Vol. 13. No.4. P. 1726.
8. Keresztes É. R., Kovács I., Horváth A., Zimányi K. Exploratory analysis of blockchain platforms in supply chain management // Economies. 2022. Vol. 10. No.9. P. 206.
9. Esmaeilian B., Sarkis J., Lewis K., Behdad S. Blockchain for the future of sustainable supply chain management in Industry 4.0 // Resources, conservation and recycling. 2020. Vol. 163. P. 105064.
10. Li K., Lee J.Y., Gharehgozli A. Blockchain in food supply chains: a literature review and synthesis analysis of platforms, benefits and challenges // International Journal of Production Research. 2023. Vol. 61. No.11. P. 3527-3546.
11. Rejeb A., Keogh J. G., Treiblmaier H. Leveraging the internet of things and blockchain technology in supply chain management // Future Internet. 2019. Vol. 11. No.7. P. 161.

УДК 004.89

ОБУЧЕНИЕ НЕЙРОННЫХ СЕТЕЙ НА ДАННЫХ ИНТЕРНЕТА ВЕЩЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ТЕХНИЧЕСКИХ ОТКАЗОВ

Гурова Л.В.
Уральский федеральный университет (Екатеринбург, Россия)

TRAINING NEURAL NETWORKS ON INTERNET OF THINGS DATA FOR TECHNICAL FAILURE PREDICTION

Gurova L.
Ural Federal University (Yekaterinburg, Russia)

Аннотация

В статье исследуются возможности применения нейронных сетей (НС) для анализа данных Интернета вещей (IoT) и прогнозирования технических отказов промышленного оборудования. НС позволяют анализировать параметры, такие как вибрации, температура и давление, для выявления аномалий и предсказания возможных неисправностей. Описаны основные этапы подготовки данных и архитектуры НС, включая рекуррентные и сверточные сети. Приведены примеры успешного применения в компаниях Siemens, GE и Hitachi, продемонстрировавшие эффективность предсказательной аналитики на основе данных IoT. Заключение подчеркивает значимость IoT и НС для повышения надежности и снижения затрат.

Ключевые слова: нейронные сети, IoT, прогнозирование отказов, промышленное оборудование, анализ данных.

Abstract

The article explores the potential of neural networks (NNs) for analyzing Internet of Things (IoT) data and predicting technical failures in industrial equipment. NNs analyze parameters such as vibrations, temperature, and pressure to detect anomalies and forecast possible faults. Key stages of data preparation and NN architectures, including recurrent and convolutional networks, are described. Examples of successful applications in companies like Siemens, GE, and Hitachi demonstrate the effectiveness of predictive analytics based on IoT data. The conclusion emphasizes the importance of IoT and NNs in enhancing reliability and reducing costs.

Keywords: neural networks, IoT, failure prediction, industrial equipment, data analysis.

Введение

С развитием технологий Интернета вещей (IoT) увеличивается объем данных, собираемых с подключенных устройств, что предоставляет широкие возможности для анализа и прогнозирования. В промышленном и инженерном секторах такие данные могут использоваться для предсказания технических отказов оборудования, минимизации простоев и улучшения эксплуатационных характеристик. IoT-устройства фиксируют ключевые параметры работы машин и механизмов, такие как температура, давление, вибрации и другие показатели, которые помогают определить потенциальные признаки износа и возможных неисправностей. Традиционные методы диагностики и мониторинга, основанные на регулярных проверках, зачастую не позволяют своевременно выявить возможные проблемы, что может привести к нежелательным простоям и экономическим потерям. В отличие от них, использование нейронных сетей (НС) для анализа IoT-данных позволяет моделировать зависимости между параметрами и прогнозировать отказы оборудования с высокой

точностью. Нейронные сети обучаются на исторических данных, выявляя скрытые взаимосвязи и паттерны, которые могут предсказать технические неисправности до их фактического наступления.

Цель данной статьи заключается в изучении применения нейронных сетей для анализа данных IoT и прогнозирования технических отказов. В статье будут рассмотрены основные подходы к обучению НС на IoT-данных, а также примеры применения моделей в реальных условиях, что позволяет оценить их эффективность и возможность интеграции в существующие системы мониторинга и обслуживания оборудования.

Основная часть

Обучение НС на данных IoT представляет собой многоэтапный процесс, включающий сбор данных, их предварительную обработку, выбор и настройку архитектуры модели, обучение и тестирование. На первом этапе данные, полученные с IoT-устройств, проходят этап очистки и нормализации, что необходимо для исключения аномальных значений и подготовки данных к использованию в моделях НС. Данные могут содержать различные параметры, такие как температура, вибрация, давление, и другие физические показатели, которые отражают состояние оборудования и позволяют выявить начальные признаки износа [1].

На рисунке 1 представлено распределение ключевых параметров, собираемых с IoT-устройств, таких как температура, давление и вибрации. Эти данные составляют основу для анализа состояния оборудования.

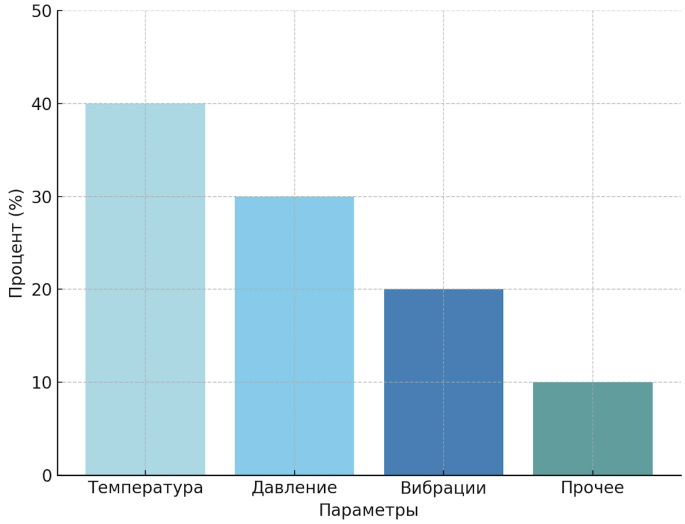

Рисунок 1. Распределение данных с IoT-устройств

Из графика видно, что наибольший объем данных приходится на параметры температуры (40%), что связано с их высокой значимостью для мониторинга оборудования. Давление и вибрации составляют 30% и 20% соответственно, подчеркивая их роль в предсказании технических отказов. Оставшиеся 10% занимают другие параметры, такие как ток или уровень шума. Это распределение данных отражает специфику использования IoT в промышленности.

Выбор архитектуры нейронной сети играет ключевую роль в успехе прогноза отказов. Для анализа временных рядов данных с IoT-устройств используются рекуррентные нейронные сети (РНС), так как они хорошо подходят для обработки последовательностей данных, например, данных о вибрациях или изменениях температуры. В то же время, сверточные нейронные сети (СНС) также могут использоваться в случае, когда необходимо выявить пространственные паттерны в данных. Глубокие НС с несколькими слоями обладают высокой гибкостью и способны анализировать сложные зависимости, что делает их подходящими для задач предсказания отказов на основе сложных IoT-данных. Процесс обучения модели включает настройку гиперпараметров, таких как количество слоев, количество нейронов, коэффициент обучения и параметры регуляризации, чтобы минимизировать переобучение и повысить точность модели [2]. Настройка этих параметров проводится на основе тестирования на отдельном наборе данных, который не использовался в обучении. Модели НС обучаются на исторических данных, содержащих как нормальные состояния оборудования, так и данные,

зарегистрированные перед отказом. Использование метода обратного распространения ошибки позволяет корректировать веса нейронов и улучшать точность прогнозирования.

Примеры успешного внедрения нейронных сетей для прогнозирования технических отказов на основе данных IoT

Одним из ярких примеров использования нейронных сетей и IoT для прогнозирования технических отказов является проект **компании Siemens**, внедрившей технологию предсказательной аналитики в сфере энергетики и промышленного оборудования. На базе данных, собираемых с датчиков IoT, Siemens использует нейронные сети для анализа состояния турбин и генераторов. Эти модели позволяют обнаруживать аномалии в работе оборудования на ранних стадиях, что позволяет снизить вероятность аварий и минимизировать затраты на ремонт. Подход с использованием нейронных сетей продемонстрировал высокий уровень точности предсказаний, позволивший компании сократить время простоя и снизить затраты на эксплуатацию [3-5].

Компания General Electric (GE) также активно использует нейронные сети и IoT в своих решениях для предсказательного обслуживания авиационных двигателей и локомотивов. В рамках платформы GE Predix применяются модели глубокого обучения для анализа данных с подключенных устройств, которые фиксируют состояние ключевых компонентов. Например, система мониторинга авиационных двигателей отслеживает показатели вибрации, температуры и давления, позволяя нейронным сетям выявлять изменения, которые могут сигнализировать о приближающейся неисправности. Этот подход помог GE не только повысить безопасность эксплуатации двигателей, но и значительно сократить непредвиденные расходы на ремонт.

Компания Hitachi внедрила технологию предсказания отказов в свою систему управления промышленными насосами. Используя данные, полученные с IoT-датчиков, нейронные сети анализируют такие параметры, как давление и температура, что позволяет предсказать износ и предупредить возможные неисправности. Применение этого подхода позволило Hitachi продемонстрировать сокращение количества поломок на 40% и оптимизировать обслуживание насосного оборудования, сделав его более экономически эффективным.

Эти примеры свидетельствуют о высокой эффективности применения нейронных сетей в сочетании с IoT для прогнозирования технических отказов, что помогает промышленным компаниям сократить затраты, связанные с простоями, и повысить надежность работы оборудования.

Текущие и перспективные проблемы внедрения нейронных сетей на данных IoT

Несмотря на многочисленные преимущества использования нейронных сетей для анализа данных IoT, существуют определенные проблемы и риски, которые необходимо учитывать при внедрении таких технологий в промышленность. Одной из главных проблем является качество и полнота данных, поступающих с IoT-устройств [6]. Данные могут содержать шум, аномалии или быть неполными, что влияет на точность модели. Например, сбои в передаче данных или ошибки датчиков могут привести к искажениям, что создаст сложности при обучении нейронных сетей. Для повышения надежности анализа требуется тщательная очистка и фильтрация данных перед их использованием в модели.

Кроме того, нейронные сети требуют значительных вычислительных ресурсов и времени для обучения, особенно при анализе больших массивов данных, поступающих с тысяч датчиков. В условиях промышленного производства, где важна оперативность принятия решений, задержки, вызванные обработкой данных, могут негативно сказаться на эффективности процессов [7, 8]. Внедрение современных вычислительных технологий, таких как облачные вычисления или графические процессоры (GPU), позволяет ускорить обучение моделей, но увеличивает затраты на инфраструктуру, что может быть затруднительно для малых и средних предприятий.

Наконец, вопрос интерпретируемости результатов нейронных сетей остаётся важным аспектом. Нейронные сети, особенно глубокие, часто рассматриваются как "черный ящик," и

результаты их работы могут быть сложны для интерпретации пользователями без специальных знаний. Это особенно критично в производственных условиях, где необходимо не только предсказать возможные отказы, но и понять причины их возникновения для предотвращения в будущем. Разработка методов, повышающих интерпретируемость нейронных сетей, таких как метод LIME (Local Interpretable Model-agnostic Explanations) или визуализация активаций слоев, может способствовать более широкому внедрению этих технологий в промышленности [9].

Заключение

Использование НС на данных IoT для прогнозирования технических отказов позволяет промышленным компаниям существенно повысить эффективность управления оборудованием. НС обеспечивают возможность предсказательного анализа на основе данных, поступающих в режиме реального времени, что способствует более точному выявлению потенциальных отказов и снижению времени простоя. Такой подход значительно повышает надежность эксплуатации, снижая риски и экономические издержки, связанные с внеплановыми ремонтами. Применение глубокого обучения и продвинутых архитектур нейронных сетей, таких как РНС и СНС, позволило достичь высокого уровня точности предсказаний. Эти модели могут обрабатывать сложные временные ряды и анализировать скрытые зависимости между параметрами, что делает их незаменимыми в условиях высоких требований к надежности оборудования. Благодаря успешным примерам внедрения, таким как проекты компаний Siemens, GE и Hitachi, становится очевидным, что использование IoT в сочетании с методами искусственного интеллекта может изменить подход к эксплуатации промышленного оборудования. Таким образом, нейронные сети в анализе IoT-данных открывают перспективы для создания полностью автономных систем мониторинга и управления, которые могут функционировать в режиме реального времени, выявляя неисправности еще до их фактического наступления. Это подтверждает актуальность и перспективность данной технологии для широкого внедрения в индустрии, ориентированные на оптимизацию затрат и повышение уровня безопасности.

Список литературы

1. Кожевников А.В., Илатовский И.С., Соловьева О.И. Применение методов машинного обучения в рамках прогнозирования состояния электромеханических систем прокатного производства // Вестник Череповецкого государственного университета. 2017. №1 (76). С. 33-39.
2. Капанский А.А. Современные стратегии использования искусственного интеллекта для предотвращения аварий в технических системах ресурсоснабжения // Вестник Казанского государственного энергетического университета. 2024. Т. 16. №1 (61). С. 38-51.
3. Поляков Р.Н., Крупенин Н.В., Кудрявцев И.Е., Кузькин В.И., Кузькин А.В. Подходы к построению интеллектуальной системы мониторинга, диагностики и прогнозирования технического оборудования // Современные материалы и технологии восстановления и упрочнения деталей промышленного оборудования. 2021. С. 81-85.
4. Елагин В.С., Гребенщикова А.А. Прогнозирование трафика трехмерной сети интернета вещей высокой плотности как многомерного случайного процесса // Труды учебных заведений связи. 2024. Т. 10. №4. С. 38-47.
5. Квонг С.В., Щербаков М.В. Архитектура системы предсказательного технического обслуживания сложных многообъектных систем в концепции Индустрии 4.0 // Программные продукты и системы. 2020. Т. 33. №2. С. 186-194.
6. Yen C.T., Chuang P.C. Application of a neural network integrated with the internet of things sensing technology for 3D printer fault diagnosis // Microsystem Technologies. 2022. Vol. 28. №1. С. 13-23.
7. Wang H., Zhang W., Yang D., Xiang Y. Deep-learning-enabled predictive maintenance in industrial internet of things: methods, applications, and challenges // IEEE Systems Journal. 2022. Vol. 17. №2. P. 2602-2615.

8. Latif S., Zou Z., Idrees Z., Ahmad J. A novel attack detection scheme for the industrial internet of things using a lightweight random neural network // IEEE access. 2020. Vol. 8. P. 89337-89350.
9. Liu H. Application of industrial Internet of things technology in fault diagnosis of food machinery equipment based on neural network // Soft Computing. 2023. Vol. 27. №13. P. 9001-9018.

УДК 004.6

АЛГОРИТМЫ ОБНАРУЖЕНИЯ АНОМАЛИЙ В РАЗНОРОДНЫХ НАБОРАХ ДАННЫХ

Шукуров Н.А.
магистр, Таджикский технический университет имени М.С. Осими (Душанбе, Таджикистан)

ANOMALY DETECTION ALGORITHMS IN HETEROGENEOUS DATASETS

Shukurov N.
master's degree, Tajik Technical University named after M.S. Osimi (Dushanbe, Tajikistan)

Аннотация

Статья посвящена исследованию алгоритмов обнаружения аномалий в разнородных наборах данных. Описаны основные методы, включая статистические подходы, алгоритмы машинного обучения и глубокого обучения, такие как изоляционный лес, автокодировщики и рекуррентные нейронные сети (РНС), а также методы временных рядов. Приведены примеры применения изоляционного леса для выявления аномальных точек данных. Статья подчеркивает актуальность использования этих алгоритмов в задачах предсказания сбоев и повышения безопасности. Заключение акцентирует внимание на важности комбинации различных подходов для достижения высокой точности и эффективности.

Ключевые слова: обнаружение аномалий, разнородные данные, машинное обучение, изоляционный лес, глубокое обучение.

Abstract

The article explores anomaly detection algorithms for heterogeneous data sets. It describes key methods, including statistical approaches, machine learning algorithms, and deep learning techniques such as Isolation Forest, autoencoders, and recurrent neural networks (RNNs), as well as time series methods. An example using Isolation Forest to detect anomalous data points is provided. The article emphasizes the relevance of these algorithms in failure prediction and security enhancement tasks. The conclusion highlights the importance of combining different approaches to achieve high accuracy and efficiency.

Keywords: anomaly detection, heterogeneous data, machine learning, Isolation Forest, deep learning.

Введение

Современные наборы данных, используемые в различных отраслях, часто обладают высокой степенью разнородности, что затрудняет их анализ и обработку стандартными методами. Разнородные данные включают как структурированные, так и неструктурированные данные, которые могут поступать из множества источников, таких как социальные сети, датчики, финансовые системы и веб-ресурсы. Из-за сложности этих данных и их различной природы возникает потребность в эффективных алгоритмах, способных обнаруживать аномалии и отклонения от нормы, которые могут сигнализировать о проблемах в работе системы, безопасности или данных.

Обнаружение аномалий в разнородных данных важно для различных применений, таких как предсказание неисправностей в производственных системах, обнаружение мошенничества

в банковской сфере, анализ поведения пользователей и мониторинг состояния инфраструктуры. Традиционные методы анализа данных, ориентированные на однородные наборы, часто не справляются с задачей анализа разнородных источников, что повышает актуальность разработки специализированных алгоритмов. Эти алгоритмы способны учитывать сложные взаимосвязи между параметрами, которые отличаются по структуре и временным характеристикам.

Цель данной статьи заключается в рассмотрении существующих алгоритмов обнаружения аномалий, применяемых к разнородным наборам данных, а также в анализе их эффективности и возможностей. В статье будут описаны различные подходы, включая методы машинного обучения и глубокого обучения, а также подходы, основанные на анализе временных рядов и статистическом моделировании.

Основная часть. Основные алгоритмы обнаружения аномалий

Обнаружение аномалий в разнородных наборах данных требует применения различных подходов, которые могут учитывать как структурированные, так и неструктурированные данные. Наиболее эффективные методы включают статистические подходы, алгоритмы машинного обучения (МО) и глубокого обучения, а также методы, основанные на временных рядах [1].

Статистические методы

Статистические методы являются одним из классических подходов к обнаружению аномалий и основываются на анализе распределений данных. Например, метод межквартильного размаха (IQR) позволяет выявлять аномалии в числовых данных, сравнивая значения с межквартильным диапазоном. Другим методом является использование гипотезы нормального распределения, где аномалии определяются как значения, находящиеся за пределами нескольких стандартных отклонений от среднего. Статистические методы эффективны для анализа небольших и однородных данных, но часто имеют ограничения при работе с разнородными наборами, где распределение данных может быть неоднородным и непредсказуемым.

Методы МО

МО-алгоритмы, такие как метод опорных векторов (SVM) и случайный лес (Random Forest), применяются для обнаружения аномалий путем обучения модели на исторических данных. SVM, например, выделяет гиперплоскость, которая разделяет нормальные и аномальные точки в пространстве признаков. Случайный лес строит несколько деревьев решений для анализа данных, и аномальные точки определяются на основе частоты их классификации в качестве «выбросов». МО позволяет использовать как супервизорные, так и несупервизорные алгоритмы, что делает их гибкими в использовании для различных наборов данных. Эти методы хорошо работают с разнородными наборами данных, но для достижения высокой точности требуют достаточно больших объемов данных для обучения [2].

Изоляционные методы

Изоляционный лес (Isolation Forest) — метод, специально разработанный для обнаружения аномалий. Он строит несколько деревьев, где данные, которые легче изолируются, считаются аномалиями. Этот метод эффективен для анализа высокоразмерных и разнородных данных, поскольку он не требует гипотезы о распределении данных и работает быстро на больших объемах. Изоляционный лес особенно подходит для применения в наборах данных, где аномалии представляют собой редкие и значительно отличающиеся от нормы объекты [3].

Глубокое обучение

Методы глубокого обучения, такие как автокодировщики и генеративно-состязательные сети (GAN), обеспечивают высокую гибкость в анализе сложных и разнородных данных. Автокодировщики представляют собой нейронные сети, обучающиеся воспроизводить входные данные на выходе. При этом они эффективно кодируют нормальные данные, но «проваливаются» на аномальных, что делает их обнаружение возможным на основе высокой ошибки восстановления. GAN, с другой стороны, создают синтетические данные и

сравнивают их с реальными, что позволяет выявлять аномалии как отклонения от синтетической нормы. Глубокое обучение эффективно для работы с разнородными данными, особенно в случаях, когда набор данных обладает сложной внутренней структурой и большим объемом [4, 5].

Методы анализа временных рядов

Для данных, которые включают временные зависимости, такие как сенсорные данные или финансовые транзакции, используются методы анализа временных рядов. Рекуррентные нейронные сети (РНС), такие как LSTM, позволяют моделировать временные зависимости и выявлять аномалии на основе отклонений от предсказанных значений. Также используются методы, такие как ARIMA и STL, которые выявляют аномалии путем анализа изменений и выбросов в данных на временной шкале. Методы временных рядов особенно полезны для прогнозирования изменений во времени и обнаружения аномалий, связанных с сезонностью и трендами [6].

Пример использования изоляционного леса для обнаружения аномалий

Изоляционный лес (Isolation Forest) – один из популярных методов для обнаружения аномалий, особенно в разнородных и высокоразмерных наборах данных. Этот алгоритм работает на основе принципа изоляции: данные, которые легче изолировать, считаются аномальными [7]. Изоляционный лес строит несколько деревьев, и если точка данных изолируется на ранних уровнях дерева, то она считается выбросом. Ниже приведен пример кода на Python, демонстрирующий использование изоляционного леса для обнаружения аномалий. Библиотека sklearn предоставляет готовую реализацию этого алгоритма.

```
from sklearn.ensemble import IsolationForest
import numpy as np

# Пример данных, содержащих нормальные точки и выбросы
data = np.array([[10, 5], [12, 6], [11, 5], [9, 5], [15, 10], [100, 2], [20, 1], [10, 4], [11, 6], [110, 2]])

# Инициализация модели изоляционного леса с указанием доли выбросов
iso_forest = IsolationForest(contamination=0.2, random_state=42)
iso_forest.fit(data)

# Предсказание аномалий (-1 - аномалия, 1 - нормальная точка)
predictions = iso_forest.predict(data)
print("Результаты предсказания:", predictions)

# Вывод результатов
for i, prediction in enumerate(predictions):
    if prediction == -1:
        print(f"Точка {data[i]} считается аномалией")
    else:
        print(f"Точка {data[i]} считается нормальной")
```

Описание работы примера

Подготовка данных: Массив data содержит нормальные точки и несколько выбросов, которые мы хотим обнаружить.

Создание модели: Создается модель IsolationForest, где параметр contamination=0.2 указывает на ожидаемую долю выбросов в данных (в данном случае 20%).

Предсказание аномалий: Метод predict() возвращает массив, где -1 указывает на аномалию, а 1 — на нормальную точку. Таким образом, алгоритм определяет выбросы, изолируя их на ранних уровнях дерева.

Вывод результата: Каждая точка данных проверяется на принадлежность к норме или аномалии, и результаты выводятся на экран.

Интерпретация результатов

В данном примере изоляционный лес обнаруживает точки, которые отличаются от нормальных значений, таких как (100, 2) и (110, 2). Эти точки имеют значения, которые значительно отклоняются от остальных точек, что и делает их аномалиями. Использование изоляционного леса позволяет эффективно выделять выбросы даже в разнородных наборах данных, делая его подходящим инструментом для различных прикладных задач анализа данных [8].

Заключение

Внедрение алгоритмов обнаружения аномалий в разнородные наборы данных предоставляет организациям возможность автоматизировать процессы мониторинга, повышая надежность работы систем и предотвращая потенциальные сбои и нарушения безопасности. Обзор основных алгоритмов показал, что статистические методы и методы МО, такие как изоляционный лес и методы глубокого обучения, могут быть адаптированы для анализа данных различной структуры и временных характеристик. Однако выбор оптимального метода зависит от конкретного типа данных и задачи. Методы, такие как изоляционный лес и автокодировщики, доказали свою эффективность для работы с высокоразмерными и разнородными данными. Эти алгоритмы обеспечивают высокую гибкость и масштабируемость, что позволяет использовать их в реальных условиях, где часто требуется оперативное обнаружение аномалий. Кроме того, применение временных моделей, таких как рекуррентные нейронные сети, позволяет учитывать временные зависимости, которые особенно важны для мониторинга динамических систем. Таким образом, комбинирование различных алгоритмов, таких как статистические подходы, МО и методы временных рядов, дает наиболее точные результаты и позволяет создавать устойчивые решения для обнаружения аномалий. Эти технологии обеспечивают новый уровень безопасности и контроля, позволяя компаниям своевременно реагировать на потенциальные угрозы и оптимизировать работу своих систем.

Список литературы

1. Багутдинов Р.А. Подход к обработке, классификации и обнаружению новых классов и аномалий в разнородных и разномасштабных потоках данных // Вестник Дагестанского государственного технического университета. Технические науки. 2018. Т. 45. №3. С. 85-93.
2. Чесноков М.Ю. Поиск аномалий во временных рядах на основе ансамблей алгоритмов DBSCAN // Искусственный интеллект и принятие решений. 2018. №1. С. 99-107.
3. Багутдинов Р.А., Саргсян Н.А., Красноплахтыч М.А. Аналитика, инструменты и интеллектуальный анализ больших разнородных и разномасштабных данных // Экономика. Информатика. 2020. Т. 47. №4. С. 792-802.
4. Вишератин А.А., Мухина К.Д., Струков А.М. Разработка интеллектуальной платформы мониторинга состояния городской среды на основе разнородных данных // Альманах научных работ молодых ученых Университета ИТМО. 2019. С. 146-150.
5. Щеглеватых Р.В., Сысоев А.С. Математическая модель обнаружения аномальных наблюдений с использованием анализа чувствительности нейронной сети // Моделирование, оптимизация и информационные технологии. 2020. Т. 8. №1. С. 14-15.
6. Симанков В.С., Колодий А.С. Подход к построению архитектуры интеллектуальной системы обнаружения и устранения сетевых аномалий // Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. 2012. №4 (110). С. 191-196.
7. Dang Y., Wang B., Brant R., Zhang Z., Alqallaf M., Wu Z. Anomaly detection for data streams in large-scale distributed heterogeneous computing environments // ICMLG2017 5th International Conference on Management Leadership and Governance. 2017. P. 121.
8. Zhu J., Ding C., Tian Y., Pang G. Anomaly heterogeneity learning for open-set supervised anomaly detection // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024. P. 17616-17626.

UDC 004.056

DATA PROTECTION IN CLOUD SYSTEMS USING MACHINE LEARNING-BASED ENCRYPTION

Davydova A.
bachelor's degree, Siberian Federal University (Krasnoyarsk, Russia)

ЗАЩИТА ДАННЫХ В ОБЛАЧНЫХ СИСТЕМАХ С ИСПОЛЬЗОВАНИЕМ ШИФРОВАНИЯ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Давыдова А.С.
бакалавр, Сибирский федеральный университет (Красноярск, Россия)

Abstract

This article explores modern data encryption methods in cloud systems using machine learning (ML) algorithms. The main approaches include adaptive encryption models that adjust data protection processes based on the data type and volume. Special attention is given to ML-based anomaly detection methods, enabling prompt identification of potential threats and preventing data leaks. The use of hybrid encryption methods, combining symmetric and asymmetric encryption, ensures high security levels and optimized computational costs. Emphasis is placed on ML's role in enhancing encryption key distribution and risk prediction processes. Practical recommendations for implementing proposed solutions aimed at improving cloud system reliability are presented. The analysis shows that the use of ML in cloud data encryption significantly enhances data security, minimizes risks, and boosts cloud technology performance.

Keywords: data encryption, cloud systems, machine learning, anomaly detection, hybrid encryption.

Аннотация

В статье рассматриваются современные методы шифрования данных в облачных системах с использованием алгоритмов машинного обучения (МО). Приведены основные подходы, включающие адаптивные модели шифрования, позволяющие настраивать процесс защиты данных в зависимости от их типа и объема. Особое внимание уделено методам обнаружения аномалий на основе МО, которые позволяют оперативно выявлять потенциальные угрозы и предотвращать утечки информации. Рассматривается применение гибридных методов шифрования, сочетающих симметричное и асимметричное шифрование, обеспечивающих высокий уровень безопасности и оптимизацию вычислительных затрат. Подчеркивается значимость МО для улучшения процессов распределения ключей шифрования и прогнозирования рисков. Представлены рекомендации по практическому применению предложенных решений, направленных на повышение надежности облачных систем. Проведенный анализ демонстрирует, что использование МО в шифровании данных в облаке значительно улучшает защиту данных, минимизирует риски и повышает производительность облачных технологий.

Ключевые слова: шифрование данных, облачные системы, машинное обучение, обнаружение аномалий, гибридное шифрование.

Introduction

Modern cloud systems, which provide data storage and processing, offer companies and users convenient access to information, significantly expanding the possibilities for interaction with digital

resources. However, as the use of cloud technologies grows, so do the risks associated with data security, especially the threats of unauthorized access, theft, and data leaks. In this context, one of the most in-demand methods for protecting information in cloud systems is data encryption, which secures information from potential attacks and guarantees its confidentiality. At the same time, traditional encryption methods often encounter difficulties when working with large volumes of data typical for cloud systems. This leads to the need for developing new approaches that can not only ensure protection but also optimize the encryption and decryption processes. Machine learning (ML) opens up opportunities for creating adaptive encryption methods that can automatically adjust to the type and volume of data, as well as the parameters affecting security and performance. Such methods allow for minimizing computational costs and improving security metrics [1].

The purpose of this article is to explore approaches to data encryption in cloud systems using ML algorithms. The article will analyze modern ML-based methods for encrypting and decrypting data, as well as provide examples and recommendations for their application to enhance security levels in cloud environments.

Main part. Application of ML for optimizing encryption

Traditional encryption methods, while reliable, often face limitations when working with large volumes of data, which is characteristic of cloud systems. To overcome these challenges, ML offers adaptive encryption models that can adjust the encryption process based on the specifics and volume of data [2, 3]. Machine learning algorithms, such as neural networks, can be trained on analyzing data streams and selecting the most effective parameters for encryption, thereby reducing computational resource costs and increasing performance.

Figure 1 illustrates the relationship between the volume of data and the time required for encryption. This demonstrates how larger datasets typically lead to longer encryption times, highlighting the need for optimization.

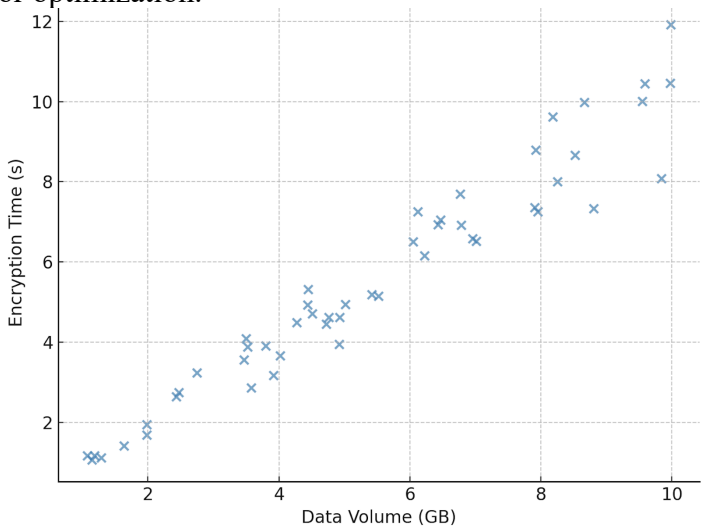

Figure 1. Dependency between data volume and encryption time

As shown in Figure 1, there is a clear trend indicating that as the data volume increases, the encryption time rises proportionally. This emphasizes the importance of adaptive ML-based encryption methods, which can minimize computational costs and enhance performance. By analyzing data streams and dynamically adjusting encryption parameters, these methods can significantly improve efficiency, even for large-scale cloud systems. ML can also be used to improve the distribution of encryption keys in cloud systems. ML methods allow for forecasting temporal patterns of data usage, which aids in timely key updates and reduces the risk of attacks. For example, clustering methods help categorize data streams, highlighting data that requires a higher level of protection. This optimizes resource usage by distributing the load on encryption systems.

Protection from attacks through anomaly analysis

ML is also applied to detect anomalies in data streams, enabling prompt identification of potential threats. For instance, anomaly detection algorithms can analyze network traffic and user behavior, identifying unusual actions that may signal attempts at hacking or unauthorized access.

Algorithms such as isolation forests and autoencoders analyze habitual activity patterns, and when deviations occur, checks and additional security measures are triggered [4]. Using ML in these processes allows the system to self-learn from incoming data and improve the accuracy of threat detection, adapting to changing conditions. This approach enhances the reliability of the system and minimizes the risks of data loss by applying preventive measures based on analysis results.

Application of hybrid encryption methods for cloud data

Hybrid encryption methods, which combine the advantages of symmetric and asymmetric approaches, are also successfully implemented in cloud systems. For instance, during data transmission, asymmetric methods may be used for secure key transfer, while symmetric algorithms encrypt the content, which reduces processing time and maintains a high level of security [5]. Combined with ML, hybrid methods can adapt to the type of data and transmission conditions.

The hybrid approach using ML allows for selecting the most appropriate encryption method for specific tasks, such as protecting databases, file storage, or messaging systems. ML analyzes the volume, nature, and frequency of data transmission and can predict the best parameters for encryption, thereby maintaining a balance between security and performance.

Application of encryption in cloud systems

Data encryption in cloud systems is one of the key methods of ensuring security, as it protects data from unauthorized access at all stages of its storage and transmission [6]. Data stored in the cloud can be subject to various attacks, including interception during transmission and attacks on storage servers. Encrypting data at both the server and client levels helps prevent information leaks, as even if attackers gain access to encrypted data without the decryption key, it remains unreadable.

Several stages play an important role in encryption. First, encrypting data before it is sent to the cloud, known as client-side encryption, ensures that the information is protected before it leaves the user's device. Thus, only encrypted data reaches the cloud, and neither the cloud provider nor third parties can access the original data [7]. Second, server-side encryption is applied when storing data in the cloud, protecting it from unauthorized access within the cloud infrastructure itself. The third stage is encrypting data during transmission, which protects data from interception when exchanged between the client and server or between servers.

Cloud systems widely employ encryption methods such as symmetric and asymmetric encryption, each having its own advantages and disadvantages [8]. For example, symmetric encryption uses a single key for both encrypting and decrypting data, making it faster but less secure in terms of key storage. Asymmetric encryption uses two different keys – a public key and a private key – enhancing security but requiring greater computational costs.

Conclusion

Thus, data encryption in cloud systems using ML methods represents an innovative approach to ensuring information security, which is particularly relevant given the growing volume of data and heightened risks of cyberattacks. ML not only optimizes encryption processes but also adapts them to dynamic conditions, reducing computational costs and improving system performance. The application of adaptive encryption models and anomaly detection technologies based on ML enables timely identification of suspicious activity, significantly reducing the likelihood of unauthorized access and data leaks. ML methods can be used to create hybrid solutions that combine the best features of symmetric and asymmetric encryption algorithms, ensuring a high level of confidentiality and reliability.

Therefore, the combination of traditional encryption methods and modern ML approaches allows for the creation of a multi-layered data protection system in the cloud that can withstand modern threats. Further research and implementation of these methods will help ensure data security and increase user trust in cloud technologies, opening new prospects for their use in various industries.

References

1. Krasov A.V., Shterenberg S.I., Fakhrutdinov R.M., Ryzhakov D.V., Pestov I.E. Analysis of information security of an enterprise based on collecting user data from open sources and monitoring

information resources using machine learning // T-Comm – Telecommunications and Transport. 2018. Vol. 12. No.10. P. 36-40.
2. Kalinine M.O., Shterenberg S.I. Analysis of information security of an enterprise based on monitoring information resources using machine learning // Intelligent technologies in transport. 2018. No.3(15). P. 47-54.
3. Vavilova A.S. Overview of existing mechanisms for ensuring data confidentiality in machine learning-based systems // Innovations. Science. Education. 2021. No.36. P. 1159-1167.
4. Babenko L.K., Shumilin A.S., Alexeyev D.M. Algorithm for ensuring the protection of confidential data in a cloud medical information system // News of the Southern Federal University. Technical sciences. 2021. No.5(222). P. 120-134.
5. Kuznetsov I.A. Security and confidentiality of data in mobile applications developed using machine learning technologies // Cold Science. 2024. No.2. P. 5-13.
6. Bespalova N.V., Nechaev S.V. Ensuring information security in cloud storage // Security issues. 2023. No.2. P. 19-26.
7. Vlasov R.S., Lukashik E.P., Osipyan V.O. Development of a mechanism for secure communication based on machine learning methods // Caspian Journal: Management and High Technologies. 2020. No.2(50). P. 108-117.
8. Angapov V.D. Overview of modern cloud platforms for machine learning purposes // Problems of modern science and education. 2023. No.7(185). P. 5-17.