

УДК 004.896

И.А. Митин, И.Б. Елистратова

Сибирский государственный университет телекоммуникаций и информатики,
Новосибирск, email: mitin85@list.ru

СТРАТЕГИЧЕСКАЯ ИНТЕГРАЦИЯ НЕЙРОННЫХ СЕТЕЙ В СОВРЕМЕННЫЕ ДАТА-ЦЕНТРЫ

Ключевые слова: автоматизация; архитектурные инновации; виртуализация; энергоэффективность; искусственный интеллект; ИТ-инфраструктура; машинное обучение; оптимизация; рекуррентные нейронные сети; системы охлаждения; управление ресурсами; центры обработки данных.

В статье представлен сравнительный анализ инновационных методологий управления центрами обработки данных с использованием искусственного интеллекта и машинного обучения. Целью исследования является выявление отличительных характеристик и ограничений, связанных с готовностью современных центров обработки данных к интеграции существующих архитектурных инноваций. Проведен обзор недавно опубликованных исследований, описывающих использование различных типов и архитектур нейронных сетей для повышения эффективности управления центрами обработки данных. Особое внимание уделяется стратегической интеграции нейронных сетей в системы охлаждения и энергопотребления, а также использованию рекуррентных нейронных сетей и методов обучения с подкреплением для оптимизации процессов управления ресурсами. Выводы автора исследования подчеркивают значимость комплексной модернизации ИТ-инфраструктуры для успешного внедрения передовых технологий искусственного интеллекта и машинного обучения в современные центры обработки данных.

I.A. Mitin, I.B. Elistratova

Siberian State University of Telecommunications and Informatics, Novosibirsk,
email: mitin85@list.ru

STRATEGIC INTEGRATION OF NEURAL NETWORKS INTO MODERN DATA CENTRES

Keywords: automation; architectural innovation; virtualisation; energy efficiency; artificial intelligence; IT infrastructure; machine learning; optimisation; recurrent neural networks; cooling systems; resource management; data centres.

The paper presents a comparative analysis of innovative methodologies for data centre management using artificial intelligence and machine learning. The aim of the study is to identify the distinguishing characteristics and limitations associated with the readiness of modern data centres to integrate existing architectural innovations. A review of recently published studies describing the use of different types and architectures of neural networks to enhance data centre management. Particular attention is given to the strategic integration of neural networks into cooling and energy management systems, and the use of recurrent neural networks and reinforcement learning techniques to optimise resource management processes. The author's findings highlight the importance of comprehensive IT infrastructure modernisation for the successful implementation of advanced artificial intelligence and machine learning technologies in modern data centres.

Внедрение передовых технологий искусственного интеллекта (ИИ) в центры обработки данных (ЦОД) знаменует собой новую эру в оптимизации ресурсов и повышении производительности. Интеграция ИИ в операции ЦОД, основанная на использовании передовых алгоритмов машинного обучения, предиктивной аналитики и автономных систем, повышает гибкость и оперативность ИТ-инфраструктуры. Преимуще-

ства интеграции ИИ – снижение операционных расходов, стратегические выгоды от инноваций и поддержка развития ИТ.

Системы ИИ, осуществляющие мониторинг и анализ операционных показателей в реальном времени, предсказывают потенциальные сбои и неэффективность оборудования, сокращая время простоя и затраты на обслуживание. Эта способность к прогнозированию эконо-

мит финансовые ресурсы и продлевает срок службы компонентов инфраструктуры, обеспечивая высокую отдачу от капитальных вложений.

Стратегическое развертывание ИИ превращает ЦОД в ключевые узлы экосистемы для развития ИТ-возможностей. Используя потенциал ИИ, ЦОД максимизируют производительность и рентабельность, укрепляя свою роль в современной ИТ-инфраструктуре и стимулируя инновации и устойчивое развитие.

Цель исследования

Определить степень готовности современных центров обработки данных к интеграции архитектурных инноваций на основе искусственного интеллекта и машинного обучения, выявить связанные с этим характеристики и ограничения.

Материал и методы исследования

Приведен сравнительный анализ инновационных методологий управления центрами обработки данных с использованием искусственного интеллекта и машинного обучения, для этого выполнен обзор текущих исследований по применению различных архитектур нейронных сетей для оптимизации элементов инфраструктуры ЦОД; в заключении автор делает выводы о готовности современных ЦОД к внедрению этих технологий.

Результаты исследования

Особенности внедрения искусственного интеллекта в дата-центре

Внедрение ИИ в систему управления дата-центром представляет собой серьезную и все более актуальную задачу для компаний, стремящихся повысить эффективность своих ИТ-ресурсов. Однако этот процесс сопряжен со значительными затратами, которые включают в себя не только необходимое оборудование, но и развитие требуемых навыков персонала. Ниже приводится краткий обзор основных категорий затрат, связанных с внедрением ИИ в центрах обработки данных, согласно результатам последних исследований.

Наиболее существенная категория расходов – это расходы на приобретение оборудования. Внедрение алгоритмов ИИ требует использования надежных вычислительных кластеров, которые требуют значительных ресурсов для охлаждения и электропитания, а также надежных и оперативных сетевых подключений.[1].

Нехватка должным образом подготовленного персонала, владеющего технологиями ИИ, требует инвестиций в программы обучения и развития сотрудников. Это включает в себя финансовые расходы, связанные с курсами, программами обучения и другими образовательными мероприятиями, направленными на повышение уровня знаний и навыков сотрудников.[2].

Эксплуатационные расходы включают в себя затраты на электроэнергию, обслуживание оборудования и поддержку инфраструктуры. Снизить эти расходы можно путем оптимизации процессов управления ЦОД, например, за счет внедрения алгоритмов распределения запросов на обслуживание с учетом стоимости электроэнергии и скорости выполнения запроса, что описано в работе [3].

Обзор инноваций в сфере управления дата-центром с использованием ИИ

Стратегическая интеграция нейронных сетей в архитектуру современных дата-центров ознаменовала новую эру интеллектуальных систем, которые существенно отличаются от традиционных подходов к управлению инфраструктурой.

Ярким примером инноваций в применении нейронных сетей является оптимизация систем охлаждения с применением рекуррентных нейронных сетей, для обучения которых используются данные вычислительной гидродинамики (CFD – Computational Fluid Dynamics). Традиционные методы охлаждения расходуют ресурсы неэффективно, поскольку они не учитывают динамическую природу нагрузки. Рекуррентные нейронные сети обеспечивают адаптивное охлаждение, которое учитывает изменения в нагрузках в реальном времени, экономит энергию и снижает нагрузку на системы охлаждения.

Интеграция глубокого обучения с RNN

В своем исследовании Янг и др. (2019) [4] предлагают решение в области облачных вычислений, повышающее устойчивость и эффективность центров обработки данных с помощью архитектуры «зеленого облака» на основе ИИ. Эта архитектура использует рекуррентные нейронные сети (RNN) для динамического регулирования ресурсов и оптимизации энергопотребления.

Используемые в архитектуре «зеленого облака» RNN обрабатывают последовательные данные, что позволяет оптимально прогнозировать переменные нагрузки и потребности в ресурсах в реальном времени. Янг и др. (2019) утверждают, что RNN особенно полезны для приложений, требующих анализа исторических данных для предсказания будущих событий. [4]

Интеграция методов обучения с подкреплением

В предлагаемом подходе предлагается использовать методы обучения с подкреплением, которые помогают системе планирования учиться через взаимодействие с использованием механизма вознаграждения. Это позволяет облачной системе принимать решения, постепенно снижая энергопотребление и улучшая общую производительность.

Математическая модель

Два важных уравнения (1,2) демонстрируют основную функциональность RNN в архитектуре решения [4].

Уравнения обновления состояния и вывода RNN:

$$h_t = \sigma(W_h x_t + U_h h_{t-1} + b_h), \quad (1)$$

$$o_t = \sigma_y(W_y h_t + b_y). \quad (2)$$

где h_t – текущее состояние сети в момент времени t ;

x_t – входное значение в момент времени t ;

o_t – выходное значение в момент времени t ;

W_h, U_h – весовые матрицы для входного значения и предыдущего состояния;

b_h, b_y – смещающие параметры bias;

$\sigma, \sigma_{y\sigma}$ – функции активации, обычно сигмоида или tanh.

Уравнения отражают рекуррентную природу сети, где каждое состояние зависит от предыдущего состояния и текущего входного значения, сохраняя память о предыдущих шагах. Уравнение обновления состояния описывает, как входные данные и предыдущая информация объединяются для обновления состояния сети, в то время как уравнение вывода использует текущее состояние для генерации выхода сети, оказывая влияние на последующие действия и решения. Эти уравнения объясняют, как сеть обновляет состояния и выводы на основе входных данных и предыдущих состояний, поддерживает память на протяжении временных шагов и обеспечивает лучшую производительность.

Авторы детально анализируют функции каждого параметра, демонстрируя сложные и эффективные вычисления, лежащие в основе их системы на базе ИИ. Для поиска оптимального решения задачи распределения ресурсов используется алгоритм Дискретной Оптимизации Роя Частиц (DPSO – Discrete Particle Swarm Optimization) [4].

Алгоритм DPSO демонстрирует высокую эффективность в решении задач комбинаторной оптимизации, тем не менее, поиск оптимального решения может затянуться из-за изменчивости среды. Инициализированный рой частиц служит оптимизированным решением, что значительно сокращает время поиска.

Интеграция RNN с CFD

Исследование, опубликованное в Journal of Machine and Computing (2023) [5], рассматривает интеграцию RNN с CFD для оптимизации систем охлаждения в дата-центрах. Целью исследования является повышение энергоэффективности и улучшение управления через разработку прогностической модели, адаптирующейся к изменяющимся условиям эксплуатации.

В предлагаемом подходе для моделирования среды центра обработки данных используются передовые методы CFD. В результате моделирования генерируются наборы данных, охватывающие целый ряд сценариев, включая различные вычислительные нагрузки и изменения в системах охлаждения. Сбор данных позволяет определить приоритеты воздушных потоков и температурных градиентов, обеспечивая надежную основу для последующего применения машинного обучения.

RNN выбраны за их способность обрабатывать последовательные данные и распознавать временные паттерны, что делает их подходящими для предсказания будущих состояний на основе исторических данных. Обучение RNN на данных, созданных с помощью CFD, позволяет разрабатывать модели, прогнозирующие динамику температуры и воздушных потоков в центре обработки данных. Эти прогнозы позволяют корректировать работу систем охлаждения, снижая риск перегрева и оптимизируя энергопотребление [5].

Результаты исследования показывают, что интеграция RNN с CFD существенно повышает эффективность охлаждения. Предсказательные модели обеспечивают целенаправленное охлаждение, предотвращая избыточное энергопотребление и повышая общую устойчивость работы центров обработки данных. Этот подход не только оптимизирует энергопотребление, но и создает прецедент для применения аналогичных методов в других сложных системах, требующих динамического управления окружающей средой.

Архитектура решения демонстрирует значительное улучшение эффективности и надежности работы систем охлаждения в ЦОД [5]. Применение ИИ и машинного обучения (МО) позволяет минимизировать энергетические потери, поддерживать оптимальные температурные условия и продлевать срок службы оборудования, что способствует устойчивому развитию и снижению эксплуатационных затрат.

Модель GRLA

Генетический алгоритм обучения с подкреплением (GRLA – Genetic

Reinforcement Learning Algorithm) [6], разработанный исследователями факультета компьютерных наук Университета Ибн Зоур, оптимизирует динамические ресурсы в облачных вычислительных средах. Этот инновационный подход объединяет генетические алгоритмы (GA – Genetic Algorithm) и обучение с подкреплением (RL – Reinforcement Learning) для решения сложных задач управления энергопотреблением в контейнерных центрах обработки данных.

GA [6] представляют собой адаптивные эвристические алгоритмы поиска, основанные на принципах естественного отбора и генетики. Популяция индивидуальных решений итеративно модифицируется с целью улучшения их качества и пригодности. RL позволяет программным агентам принимать решения, взаимодействуя с окружающей средой, чтобы максимизировать совокупное вознаграждение, с акцентом на долгосрочные преимущества, что позволяет принимать оптимальные решения в сложных и динамичных средах.

Комбинируя эти два подхода, GRLA объединяет исследовательские возможности GA с навыками принятия стратегических решений RL [6]. Такая синергия повышает адаптивность системы к изменяющимся условиям, оптимизируя размещение контейнеров и снижая энергопотребление в центрах обработки данных.

Математическая модель

Основная задача алгоритма GRLA определяется целевой функцией

$F_{obj}(X)$, направленной на минимизацию общего числа активных хостов [6]:

$$F_{obj}(X) = \sum_{H=1}^n X_H$$

где X_H – статус активности хоста H ; n – общее число хостов.

Необходимо учитывать несколько ограничений: каждый контейнер должен быть назначен одному хосту, использование ресурсов на сервере не должно превышать 80% мощности. Процесс оптимизации GRLA включает инициализацию начальной популяции решений, отбор лучших особей, создание

нового потомства через генетические операции и мутации, а также динамическую корректировку операций с помощью RL на основе обратной связи по производительности [6].

Интеграция RL с этапами GA

Интеграция RL в каждый этап генетического алгоритма (ГА) улучшает его гибкость и решающие способности. RL динамически адаптирует ключевые параметры ГА, такие как интенсивность отбора, вероятность скрещивания и частоту мутаций, непосредственно в процессе выполнения. Эта адаптация позволяет алгоритму GRLA более эффективно отвечать на изменения в условиях эксплуатации данных центров и на колебания в требованиях к обработке данных [6].

Модель Hybrid Attention-enhanced Gated Recurrent Unit with Random Forest (HAGRU-RF)

Модель HAGRU-RF [13] представляет собой сложную интеграцию методов машинного обучения для повышения точности прогнозирования производства солнечной энергии. Разработанная Лабораторией компьютерных систем и зрения Университета Ибн Зоур, она сочетает возможности GRU, специализированной формы RNN, с надежной методологией ансамблевого обучения Random Forest (RF), оптимизируя ее для работы с высокоразмерными временными данными. GRU, являясь современной адаптацией классической RNN, эффективнее учитывает временные зависимости и решает проблемы исчезающего градиента, что делает ее особенно подходящей для прогнозирования временных рядов.

Алгоритм RFt надежно работает со сложными нелинейными взаимосвязями в данных. Объединяя прогнозы нескольких деревьев решений, он уменьшает переобучение и улучшает обобщающую способность модели. Метод ансамблевого обучения предполагает обучение нескольких деревьев решений на различных подмножествах данных и последующее объединение их прогнозов. Это повышает точность и устойчивость предсказаний за счет усреднения результатов всех деревьев, снижая влияние выбросов и улучшая обобщение.

Компонент RF обучается на признаках, извлеченных архитектурой HAGRU, предоставляя дополнительный уровень учета сложных нелинейных зависимостей в данных. Принцип ансамблевого обучения объединяет прогнозы нескольких деревьев решений, что уменьшает переобучение и повышает обобщающую способность модели [6].

Сравнительный анализ эффективности нейросетей

В исследовании, опубликованном в журнале *Journal of Machine and Computing* в 2023 году [13], проведен сравнительный анализ эффективности различных архитектур нейронных сетей для управления центрами обработки данных (рис. 1).

Особое внимание уделено модели HAGRU-RF. Эта модель демонстрирует высокую эффективность в снижении ошибок прогнозирования, улавливая сложную временную динамику.

На графике представлены показатели эффективности различных моделей прогнозирования: средняя абсолютная ошибка (MAE – Mean Absolute Error), средняя квадратичная ошибка (MSE – Mean Squared Error) и корень из средней квадратичной ошибки (RMSE – Root Mean Squared Error). Модель HAGRU-RF показывает наилучшие результаты с самыми низкими значениями ошибок по всем трем метрикам.

Особенности внедрения предлагаемых решений в дата-центры

Ниже приведен сравнительный анализ проблем и перспектив, связанных с интеграцией технологий ИИ и МО в управление подсистемами современных дата-центров. Особого внимания заслуживает внедрение RNN в сочетании с CFD для улучшения систем охлаждения. Такая интеграция требует значительных инвестиций в модернизацию инфраструктуры, включая обновление аппаратного и программного обеспечения. Несмотря на значительные затраты, ожидается, что они принесут долгосрочные выгоды, включая повышение эффективности охлаждения и снижение энергозатрат (табл. 1).

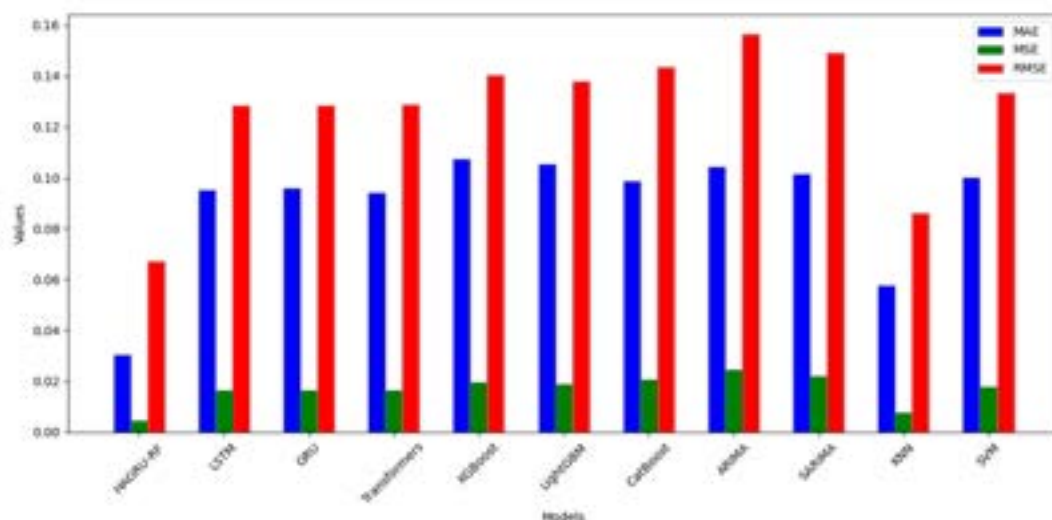


Рис. 1. Сравнение эффективности различных архитектур нейросетей [6]

Примечание: к архитектурам RNN на диаграмме относятся:

LSTM – разновидность RNN, разработанная для решения проблемы исчезающего градиента с использованием механизмов врат.

GRU – упрощенная версия LSTM, предназначенная для работы с последовательностями данных и имеющая меньше параметров.

Таблица 1

Сравнение уровней интеграции технологий ИИ и МО в инфраструктуру дата-центра

Элемент инфраструктуры	Интеграция глубокого обучения с RNN	Интеграция RNN с CFD	Модель GRLA	Модель HAGRU-RF
Система охлаждения	Капитальные изменения	Капитальные изменения	Незначительные изменения	-
Управление энергопотреблением	Незначительные изменения	-	Незначительные изменения	Капитальные изменения
Безопасность	-	-	-	-
Система хранения данных	-	-	-	-
Физическая компоновка и система распределения ресурсов	-	-	Незначительные изменения	Незначительные изменения
Управление планированием	Незначительные изменения	-	Незначительные изменения	-
Интеграция возобновляемых источников энергии	-	-	-	Капитальные изменения
Облачная архитектура	Незначительные изменения	Незначительные изменения	Незначительные изменения	Незначительные изменения
Облачные сервисы	Незначительные изменения	-	-	-
Телекоммуникационные системы	Незначительные изменения	-	Незначительные изменения	Незначительные изменения

Важно заметить, что отмечается прогресс в области интеграции возобновляемых источников энергии. Модель HAGRU-RF [6] позволяет ЦОД снизить зависимость от традиционных источников энергии и значительно уменьшить углеродный след. Внедрение этой модели требует установки нового оборудования, например, солнечных батарей или ветряных турбин, а также разработки сложных систем управления и прогнозирования, что также предполагает капитальные вложения.

Не все изменения требуют значительных инвестиций или модернизации. В таких областях, как управление энергопотреблением и планирование ресурсов, достаточно постепенных изменений, связанных с обновлением программного обеспечения и доработкой существующих систем. Эти изменения могут быть реализованы быстро и с относительно небольшими затратами, но способны значительно повысить эффективность управления ресурсами и общую энергоэффективность ЦОД.

Интеграция современных технологий ИИ и МО в работу ЦОД является значительным шагом к повышению их производительности и устойчивости. Внедрение этих технологий позволяет сократить операционные расходы и сохранить высокую степень надежности и безопасности, что особенно важно в условиях растущих требований к информационным технологиям. Все эти факторы делают ИИ не просто выгодным, а незаменимым компонентом архитектуры современного центра обработки данных.

Готовность к стратегическому внедрению систем управления, основанных на ИИ, в дата центрах

Внедрение технологий ИИ и МО в управление крупными ЦОД является важным фактором развития и повышения их эффективности. Архитектура таких ЦОД учитывает гибкость и масштабируемость, что позволяет адаптироваться к меняющимся требованиям и использовать оборудование и ПО от разных поставщиков. Финансовые ресурсы крупных ЦОД позволяют инвестировать в модернизацию, а комплексные системы безопасности способствуют интеграции дополнительных мер защиты данных и инфраструктуры. Облачные ЦОД, использующие автомати-

зацию для управления ресурсами, благодаря своей масштабируемости и гибкости, могут быстро реагировать на изменения нагрузки и требований клиентов. Эти центры находятся в авангарде технологического прогресса, что делает их подходящими для интеграции систем управления на основе нейронных сетей.

Проанализировав текущее состояние различных типов ЦОД, можно оценить их потенциал и готовность к внедрению новейших технологий ИИ. В зависимости от инфраструктурных и эксплуатационных особенностей крупные ЦОД можно разделить на три категории: гипермасштабируемые, корпоративные и колокационные дата-центры.

Гипермасштабируемые ЦОД, предназначенные для хранения больших объемов данных и предоставления облачных сервисов, уже оснащены сложными автоматизированными системами управления на базе ИИ. Эти центры часто инициируют внедрение новых технологий благодаря своей масштабируемости и инновационным инвестициям. Они обладают потенциалом для внедрения сложных решений, таких как RNN, что повышает эффективность работы и снижает энергозатраты.

Корпоративные ЦОД, обслуживающие бизнес-потребности отдельных организаций, требуют высокого уровня контроля и безопасности данных. Такие центры заинтересованы в интеграции ИИ для повышения эффективности и автоматизации, но сталкиваются с необходимостью тщательной оценки рисков и последствий нововведений. Внедрение новых технологий требует баланса между инновациями и строгими корпоративными стандартами безопасности, что может замедлить процесс интеграции и потребовать дополнительных инвестиций.

Колокационные ЦОД предоставляют инфраструктуру для размещения серверов и оборудования различных клиентов. Интеграция ИИ в таких центрах представляет собой особую проблему из-за необходимости согласовывать изменения с арендаторами и адаптироваться к их требованиям. Тем не менее, потенциал оптимизации энергопотребления и повышения производительности может стимулировать модернизацию, если удастся найти совместный подход к управлению изменениями и инвестициями.

Таблица 2

Степень готовности дата-центров к внедрению инновационных решений

Тип дата-центра	Степень готовности к внедрению ИИ и МО	Особенности
Гипермасштабируемые ЦОД	Высокая	Автоматизация, уже внедренные ИИ-решения, возможность масштабирования инноваций
Корпоративные ЦОД	Средняя	Необходимость соблюдения корпоративных стандартов безопасности, потребность в дополнительных инвестициях
Колокационные ЦОД	Низкая	Зависимость от многочисленных клиентов, сложности в координации изменений

Ниже приведена оценка степени готовности различных типов дата-центров к внедрению технологий ИИ и МО (табл. 2).

В заключение следует отметить, что внедрение ИИ и МО в ЦОД зависит от множества факторов, включая тип центра, существующую инфраструктуру и способность адаптироваться к новым технологиям. Гипермасштабируемые ЦОД находятся в более выгодном положении для внедрения новых технологий, в то время как корпоративные и колокационные ЦОД могут столкнуться с дополнительными проблемами.

Выводы

На основе анализа установлено, какие центры обработки данных лучше всего подготовлены к внедрению систем управления на базе нейронных сетей, что также подчеркивает необходимость их обширной модернизации. Внедрение таких систем значительно изменит управленческий подход в ЦОД, так как нейронные сети требуют повышенных вычислительных ресурсов и специализированных систем хранения данных, предполагая не просто обновление отдельных компонентов, но и глобальное стратегическое развитие инфраструктуры.

Прежде всего, внедрение нейронных сетей предполагает коренные изменения в архитектуре ЦОД, требуя отказа от традиционных методов управления в пользу полной трансформации систем управления и мониторинга. Это включает в себя модернизацию серверного оборудования, систем хранения данных

и сетевой инфраструктуры для поддержки высоких требований к производительности и надежности. Во-вторых, успешная интеграция нейронных сетей зависит от стратегического планирования и управления, что требует разработки долгосрочного плана развития с учетом гибкости и масштабируемости инфраструктуры для адаптации к меняющимся данным и требованиям.

Третье ключевое направление – значительные финансовые инвестиции, необходимые для модернизации аппаратного и программного обеспечения, обучения персонала и обеспечения безопасности. Важно также учитывать возможность стратегического планирования новых ЦОД с интегрированными с самого начала системами на базе нейронных сетей, что обеспечит более высокую степень интеграции и адаптивности к новым технологиям. Это позволяет центрам обработки данных использовать оборудование различных производителей, улучшая надежность системы и оптимизируя затраты.

Эффективная архитектура должна поддерживать мультивендорность и обеспечивать гибкость для масштабирования и модернизации, что важно для поддержания стабильности системы при высоких нагрузках или в случае сбоев. Таким образом, внедрение систем управления на основе нейронных сетей в ЦОД является не только технологическим обновлением, но и стратегическим решением, требующим комплексного подхода к модернизации инфраструктуры.

Список сокращений

- ARIMA – AutoRegressive Integrated Moving Average (авторегрессионное интегрированное скользящее среднее)
 CFD – Computational Fluid Dynamics (вычислительная гидродинамика)
 DPSO – Discrete Particle Swarm Optimization (дискретная оптимизация роя частиц)
 GA – Genetic Algorithm (генетический алгоритм)
 GRLA – Genetic Reinforcement Learning Algorithm (генетический алгоритм обучения с подкреплением)
 GRU – Gated Recurrent Unit (затворный рекуррентный блок)
 HAGRU-RF – Hybrid Attention-enhanced Gated Recurrent Unit and Random Forest (гибридный рекуррентный блок с усиленным вниманием и случайным лесом)
 KNN – K-Nearest Neighbors (алгоритм k-ближайших соседей)
 LSTM – Long Short-Term Memory (долговременная краткосрочная память)
 MAE – Mean Absolute Error (средняя абсолютная ошибка)
 MSE – Mean Squared Error (среднеквадратичная ошибка)
 RF – Random Forest (случайный лес)
 RL – Reinforcement Learning (обучение с подкреплением)
 RMSE – Root Mean Squared Error (корень из среднеквадратичной ошибки)
 RNN – Recurrent Neural Network (рекуррентная нейронная сеть)
 SARIMA – Seasonal AutoRegressive Integrated Moving Average (сезонное авторегрессионное интегрированное скользящее среднее)
 SVM – Support Vector Machines (машины опорных векторов)
 XGBoost – Extreme Gradient Boosting (экстремальный градиентный бустинг)
 ИИ – искусственный интеллект
 ИТ – информационные технологии
 МО – машинное обучение
 ЦОД – центр обработки данных

Библиографический список

1. Pilz K., Heim L. Compute at Scale: A Broad Investigation into the Data Center Industry. 2023. DOI: 10.48550/arXiv.2311.02651.
2. Ahmadi A. Implementing Artificial Intelligence in IT Management: Opportunities and Challenges // Asian Journal of Computer Science and Technology. 2023. № 12(2). P. 18-23. DOI: 10.51983/ajcst-2023.12.2.3704.
3. Kulkarni A.K., Annappa B. Cost aware service broker algorithm for load balancing geo-distributed data centers in cloud, 2017 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), Kollam, India, 2017. P. 1-5. DOI: 10.1109/SPICES.2017.8091337.
4. Yang J., Xiao W., Jiang C., Hossain M.S., Muhammad G., Amin S.U. AI-Powered Green Cloud and Data Center. 2019. Vol. 7. P. 4195-4203. DOI: 10.1109/ACCESS.2018.2888976.
5. Senthilkumar G., Rajendran P., Suresh Y., Herald Anantha Rufus N., Rama Chaithanya Tanguturi, Rajdeep Singh Solanki. Computational Engineering based Approach on Artificial Intelligence and Machine Learning Driven Robust Data Centre for Safe Management // Journal of Machine and Computing. 2023. Vol. 3 (4). P. 465-474. DOI: 10.53759/7669/jmc202303038.
6. Bouaouda A., Afdel K., Abounacer R. Unveiling Genetic Reinforcement Learning (GRLA) and Hybrid Attention-Enhanced Gated Recurrent Unit with Random Forest (HAGRU-RF) for Energy-Efficient Containerized Data Centers Empowered by Solar Energy and AI. Sustainability 2024. Vol. 16. P. 4438.