

ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ
И ТЕЛЕКОММУНИКАЦИИ

INFORMATION
TECHNOLOGIES
AND TELECOMMUNICATIONS

УДК 004.89, 519.246.8

DOI: 10.17212/2782-2001-2025-3-113-136

Глубокое обучение для анализа многомерных временных рядов: систематизация типов данных, задач, архитектур и подходов*

А.О. ХУСНУТДИНОВ^{1,a}, В.И. ХАБАРОВ^{1,2,b}, В.С. КАРМАНОВ^{1,c}

¹ 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет

² 630049, РФ, г. Новосибирск, ул. Дуси Ковальчук, 191, Сибирский государственный университет путей сообщения

^a evolext@gmail.com ^b khabarov51@mail.ru ^c karmanov@corp.nstu.ru

В настоящее время методы анализа данных применяются к широкому спектру разнородной информации, включая временные ряды, пространственные и пространственно-временные структуры, тексты, изображения, видео-, аудиосигналы и их различные сочетания. В контексте анализа временных рядов применение моделей на основе нейронных сетей позволяет преодолеть ограничения классических методов и расширить область применения за счет учета данных разнородных источников, обучения универсальных моделей для различных задач и разработки новых постановок. Вместе с тем разнообразие прикладных областей и исходных данных обуславливает появление множества подходов к построению нейронных сетей, адаптированных под конкретную проблему, что заставляет исследователей либо разрабатывать новую конфигурацию сетей, либо проводить масштабный поиск подходящей модели.

Работа посвящена систематизации подходов к анализу многомерных временных рядов, использующих нейросетевые модели для выявления закономерностей в данных, то есть применяющих методы глубокого обучения. Рассматриваются основные аспекты, включая типы многомерных временных рядов, типы задач анализа, способы организации компонентов нейронной сети или их архитектур, а также подходы к построению моделей. Выделены типы многомерных временных рядов, такие как стандартные, пространственно-временные, регулярные и другие. Систематизированы базовые задачи анализа временных рядов, включая традиционные задачи, решаемые классическими подходами, и задачи, ставшие возможными с внедрением методов машинного обучения, а также новые подходы, разработанные благодаря глубокому обучению. Представлен обзор ключевых типов нейронных сетей, популярных архитектур, созданных на их основе, а также методов построения или обучения моделей, использующих эти архитектуры.

Настоящая работа, не ставя целью охватить все аспекты анализа многомерных временных рядов, предлагает классификацию, которая может стать основой для разработки базы знаний в этой области и выявления перспективных направлений дальнейших исследований.

* Статья получена 30 декабря 2024 г.

Ключевые слова: анализ временных рядов, многомерные временные ряды, прогнозирование, классификация, разнородные данные, машинное обучение, глубокое обучение, обучение представлениям, нейронные сети

ВВЕДЕНИЕ

Временной ряд – это последовательность показателей, упорядоченных по времени и отражающих состояние изучаемого явления [1]. Их анализ требует учета временной зависимости между наблюдениями и находит широкое применение в различных прикладных областях – от экономики и энергетики до медицины и управления транспортными потоками. Применение соответствующих методов позволяет, в частности, повышать качество продукции, ускорять процессы и улучшать точность диагностики.

С ростом объемов данных всё большую актуальность приобретает анализ *многомерных временных рядов*, получаемых с разнородных источников – датчиков, приборов учета и других устройств с требуемой периодичностью. Использование большого объема данных потенциально повышает точность получаемых решений, однако требует учета не только временной зависимости каждой переменной, но и взаимосвязей между различными характеристиками.

Подходы к анализу временных рядов можно разделить на модельные и основанные на данных [2]. Первые опираются на физические принципы и требуют формализации, задания параметров, применения численных методов. Иной подход заключается в выявлении взаимосвязей между входными и выходными данными исключительно на основе самих данных, без знания внутреннего устройства системы, при этом она рассматривается как «черный ящик». Исследуемый объект может быть слишком сложным для описания физической моделью, в то же время за счет мониторинга нескольких ключевых переменных можно выявлять неочевидные взаимосвязи.

Подходы к анализу на основе данных прошли эволюцию от классических статистических методов к методам традиционного машинного обучения (ML, Machine Learning), а затем к методам машинного обучения, использующим нейронные сети с большим количеством слоев обработки данных, которые получили название глубоких нейронных сетей (DNN, Deep Neural Networks), а сам подход к их построению стал отдельным направлением – глубоким обучением (DL, Deep Learning). Под термином «машинное обучение» понимаются алгоритмы, которые обучаются на данных и делают прогнозы или принимают решения без явного программирования; под традиционными методами обычно понимаются модели, не основанные на нейронных сетях. Развитие глубокого обучения стало возможным благодаря графическим ускорителям и тензорным процессорам, а также появлению новых методов обработки данных нейронными сетями, что значительно улучшило обработку неструктурированных данных – от текста и изображений до временных рядов – и расширило спектр решаемых задач.

Задачи анализа временных рядов определяются структурой входных и выходных данных, а также способом применения модели (к историческим или поступающим данным). Способ организации слоев или компонентов нейронной сети между собой называется *архитектурой*, и она может варьироваться в зависимости от задачи. Поиск оптимального с точки зрения метрики качества для конкретной задачи или решения становится не менее сложной проблемой,

чем само моделирование: исследователям зачастую приходится либо разрабатывать уникальные архитектуры, либо проводить детальный анализ существующих методов для выбора подходящей модели. В ряде работ рассматриваются отдельные аспекты этой проблемы: описание задач анализа [3], обзор архитектур для задач [4–6], методы построения моделей [7–9], однако выбор инструментария для конкретной прикладной проблемы зависит от совокупности каждого из этих элементов. Следует отметить, что в упомянутых работах не описываются отдельно преимущества перехода от традиционных статистических методов к современным нейросетевым подходам.

Цель настоящей работы заключается в систематизации ключевых аспектов применения моделей глубокого обучения для анализа многомерных временных рядов. В рамках этого рассматривается классификация типов данных, постановок задач, архитектур моделей и подходов к их построению. Результаты систематизации могут лечь в основу базы знаний и в дальнейшем использоваться для создания экспертной или рекомендательной системы, предназначенной для подбора подходящего инструментария в различных предметных областях.

Для достижения поставленной цели необходимо решить следующие задачи:

- предложить классификацию типов многомерных временных рядов;
- систематизировать задачи, возникающие при анализе многомерных временных рядов;
- классифицировать используемые архитектуры нейросетевых моделей и подходы к их построению.

1. ТИПЫ МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

Стандартный многомерный временной ряд определяется как последовательность из T точек данных, упорядоченных по дискретному времени: $\mathbf{X}_{1:T} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\}$, где $\mathbf{x}_t = (x_{1,t}, \dots, x_{D,t}) \in \mathbb{R}^D$, $t = 1, \dots, T$ – точка данных во времени, а $D > 1$ – размерность каждой точки данных. Элементы вектора \mathbf{x}_t отражают значения различных переменных, например, индикаторов финансовых рынков, биометрических сигналов (электрокардиограмма, артериальное давление), данные с технических датчиков.

Пространственно-временной ряд включает как временную, так и пространственные составляющие и задается как $\mathbf{X}_{1:T}^N = \{\mathbf{X}_{1:T}^{(1)}, \mathbf{X}_{1:T}^{(2)}, \dots, \mathbf{X}_{1:T}^{(N)}\}$, где каждый компонент $\mathbf{X}_{1:T}^{(n)}$ соответствует сигналам от одного из N датчиков, регистрирующих D признаков в моменты времени $t = 1, \dots, T$. В работе [9] авторы классифицируют такие временные ряды на два подтипа в зависимости от пространственных корреляций между датчиками: 1) *пространственно-временной граф*, когда пространственная корреляция этих датчиков описывается графом \mathbf{G} с матрицей смежности $A \in \mathbb{R}^{N \times N}$; 2) *пространственно-временной растр*, когда датчики распределены равномерно в виде сетки в географическом пространстве [10]. Примерами пространственно-временных данных являются метеорологические показатели (температура, давление, влажность) в различных

географических регионах, данные о потреблении энергетических ресурсов в коммунальных системах, объемы продаж в торговых сетях и другие.

Классический статистический подход к анализу временных рядов основан на предположении, что временные точки t дискретны и следуют на равном расстоянии друг за другом (равномерная, регулярная или эквидистантная сетка). Такие временные ряды называются *регулярными*. В *нерегулярных* рядах наблюдения фиксируются с переменными интервалами, что усложняет процесс моделирования. Для приведения таких данных к регулярной форме применяются методы интерполяции или биннинга (группировка временных интервалов), однако такие подходы могут привести к потере информации и искажению динамики, как отмечается в [11]. В качестве альтернативы в работе [12] предложены нейросетевые модели, которые явно учитывают неравномерность временных интервалов. В приведенных выше формулировках предполагается числовой тип данных, однако на практике встречаются категориальные (например, события и состояния), текстовые и визуальные наблюдения. Классические методы анализа требуют предварительного преобразования таких данных в числовую форму, например, с помощью порядкового или унитарного (ONE, One-Hot Encoding) кодирования [5], что ограничивает применение к неструктурированным данным. С появлением нейронных сетей стало возможным эффективно преобразовывать тексты в векторные представления – эмбединги (embedding`s) [13]. Современные архитектуры способны одновременно обрабатывать в рамках единой модели разнородные данные – тексты, изображения и числовые временные ряды [9].

2. ЗАДАЧИ АНАЛИЗА МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

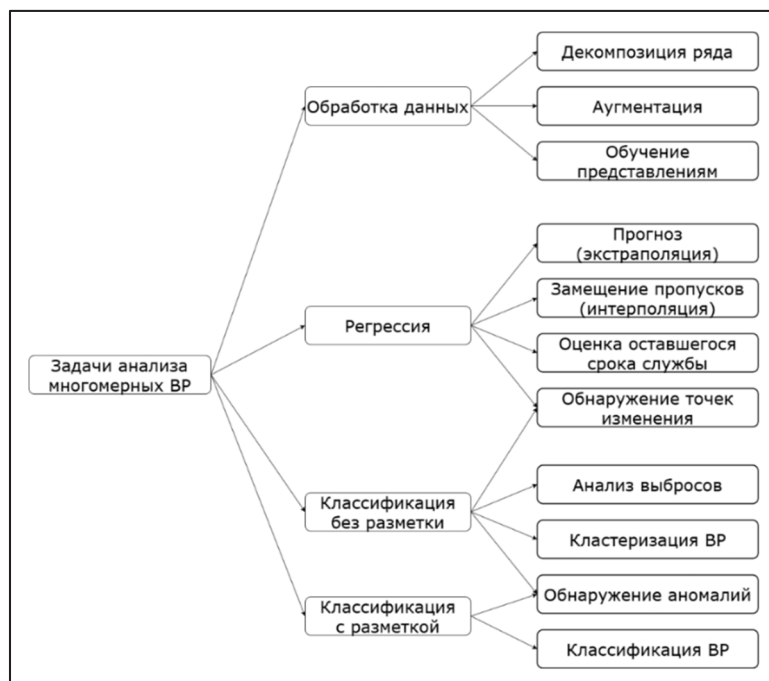
Цель анализа временных рядов – выявление внутренней структуры, закономерностей и зависимостей в данных. Модели, способные выявлять сложные зависимости, позволяют решать широкий спектр прикладных задач. В качестве варианта построения классификации предлагается разделять методы анализа на три группы: первая основывается на классических методах математической статистики; вторая объединяет подходы, сочетающие статистические методы с методами машинного обучения, появившимися до распространения нейронных сетей или не использующими их; третья группа представляет методы, возникшие благодаря применению глубоких нейронных сетей к анализу временных рядов.

К базовым задачам теории временных рядов относятся: прогнозирование (экстраполяция), замещение пропусков (интерполяция), декомпозиция временного ряда, анализ выбросов, обнаружение аномалий, выявление точек изменения (CPD, Change Point Detection), оценка оставшегося срока службы (RUL, Remaining Useful Life). Для их решения существуют методы в рамках как классического, так и нейросетевого подхода.

С развитием машинного обучения появились новые задачи анализа, в которых целевая переменная имеет дискретную природу (например, классификация и кластеризация временных рядов). Часто для их решения применяются гибридные подходы, где классические методы анализа используются для извлечения факторов, а алгоритмы машинного обучения – для прогноза целевой переменной.

Статистические подходы хотя и позволяют выявлять закономерности, но имеют ряд ограничений, связанных с отражением нелинейных связей и долгосрочных зависимостей, присутствующих в реальных данных. Кроме того, существует проблема невыполнения базовых предположений, лежащих в основе этих методов [14], для данных на практике. Многие традиционные подходы изначально были разработаны для одномерных рядов и применяются отдельно к каждой компоненте многомерных данных, что ведет к потере информации о взаимосвязях между переменными, важных, например, для энергетических, коммунальных и транспортных систем [15–17].

Применение нейронных сетей открыло новые задачи анализа временных рядов, такие как извлечение скрытых информативных признаков (RL, Representation Learning; в русскоязычной литературе часто используется перевод «обучение представлениям») [18] и аугментация данных (Data Augmentation) [19, 20]. В настоящей работе мы осознанно используем альтернативный перевод термина RL, более точно отражающий суть метода. Эта задача предполагает автоматическое извлечение информативных признаков из исходных данных и формирование компактных векторных описаний без участия человека. В свою очередь, задача Data Augmentation заключается в искусственном увеличении набора исходных наблюдений путем модификации существующих данных или генерации новых синтетических образцов, сохраняющих основные статистические и структурные свойства исходных. Нейросетевые модели интегрируют извлечение признаков и прогнозирование в единую структуру. Визуализация иерархии методов анализа временных рядов, предлагаемой авторами настоящей статьи, представлена на рисунке.



Иерархия задач анализа многомерных временных рядов

A hierarchy of tasks of multivariate time series analysis

Представленная систематизация не претендует на исчерпывающую полноту: некоторые задачи могут быть сгруппированы по общим чертам или уточнены при более глубоком анализе, однако охватываются наиболее распространенные подходы, применяемые в прикладных задачах.

2.1. ЗАДАЧА ПРОГНОЗИРОВАНИЯ

Прогнозирование (экстраполяция) временных рядов – ключевая задача анализа временных рядов, связанная с определением будущих значений на основе предыдущих наблюдений. Формально задача сводится к аппроксимации условного распределения прогнозируемых значений временного ряда:

$$p(\mathbf{X}_{t+1:t+h} | \mathbf{X}_{1:t}, \mathbf{Z}_{1:t+h}; \theta), \quad (1)$$

где \mathbf{X}_t – значения временного ряда в моменты времени t ; h – величина шага прогнозирования (горизонт планирования); \mathbf{Z}_t – значения вспомогательных (экзогенных) переменных. Здесь θ – параметры аппроксимирующей вероятностной модели:

$$\theta = \Psi(\mathbf{X}_{1:t}, \mathbf{Z}_{1:t+h}, \Phi), \quad (2)$$

где Φ – обучаемые параметры базового алгоритма прогнозирования, а $\Psi(\bullet)$ – функция, сопоставляющая обучающим данным параметры вероятностной модели.

Эта задача относится к регрессионной и не требует ручной разметки. В работе [6] прогнозные методы разделяются на вероятностные и точечные в зависимости от формы выходного результата. Вероятностные подходы оценивают всё условное распределение будущих значений (например, в виде плотности вероятности или параметров распределения), тогда как точечные выдают лишь отдельную характеристику этого распределения (например, математическое ожидание, медиану или выбранный квантиль). Оба этих типа могут быть дополнительно разделены на одношаговые и многошаговые, при этом подчеркивается, что любую одношаговую модель можно преобразовать в многошаговую.

Важное различие между подходами к прогнозированию связано с режимом их применения. Различают офлайн-прогнозирование на полном наборе накопленных данных и онлайн-прогнозирование в реальном времени по мере поступления новых наблюдений [21]. Это различие актуально для всех типов методов (как классических, так и нейросетевых) и особенно важно в задачах, требующих немедленного реагирования на изменения до наступления следующего временного шага.

К классическим методам прогнозирования многомерных временных рядов относятся векторные обобщения одномерных подходов, среди них выделяют векторную авторегрессию (VAR, Vector Autoregression), векторное скользящее среднее (VMA, Vector Moving Average) и их комбинации для стационарных (VARMA, Vector Autoregressive Moving Average) и нестационарных процессов (VARIMA, Vector Autoregressive Integrated Moving Average). Для учета экзогенных переменных применяются расширенные версии этих методов: VARMAX (Vector Autoregressive Moving Average with exogenous inputs)

и VARIMAX (интегрированная версия VARMAX) [22]. Сезонные эффекты описываются с помощью сезонного авторегрессионного интегрированного скользящего среднего (SARIMA, Seasonal Autoregressive Integrated Moving Average) и расширенной модификации с экзогенными переменными – SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous inputs). К числу других классических подходов также относятся экспоненциальное сглаживание (ES, Exponential Smoothing), нелинейная авторегрессия (NAR, Nonlinear Autoregressive), методы пространства состояний (SSM, State Space Models) [6], методы локальной аппроксимации, фильтрация Калмана (Kalman Filtering) и адаптивные полигармонические модели, нормальное времячастотное преобразование (NTFT, Normal Time-Frequency Transform) [23]. Кроме того, существуют методы, основанные на анализе многомерных зависимостей, включая методы Копул [23], сингулярный спектральный анализ (SSA, Singular Spectrum Analysis) [24], множественную линейную регрессию (MLR, Multiple Linear Regression) [6].

2.2. ЗАДАЧА ЗАМЕЩЕНИЯ ПРОПУСКОВ

Наличие пропусков в данных – постоянная и распространенная проблема в реальных прикладных задачах. Влияние отсутствующих значений на построение модели зависит от их относительной доли в наборе данных. Причинами пропусков могут быть сбои регистрирующего оборудования, отключение энергии, ошибки хранения или передачи данных [25].

Пусть многомерный временной ряд $\mathbf{X}_{1:T}$ представляется в виде матрицы признаков \mathbf{X} , где элемент x_t^d соответствует значению характеристики d в момент времени t и некоторые из элементов могут отсутствовать. Тогда задача замещения пропусков многомерного временного ряда в общем виде формулируется как восстановление матрицы \mathbf{X}^* :

$$\mathbf{X}^* = \mathbf{M}\mathbf{X} + (\mathbf{I} - \mathbf{M})\widehat{\mathbf{X}}, \quad (3)$$

где $\widehat{\mathbf{X}}$ – матрица признаков с восстановленными пропусками; \mathbf{M} – матрица-маска с элементами

$$m_{d,t} = \begin{cases} 1, & \text{если } x_t^d \text{ известно;} \\ 0 & \text{в ином случае,} \end{cases} \quad (4)$$

где $m_{d,t}$ – элемент матрицы \mathbf{M} .

В работе [26] рассматриваются три механизма возникновения пропусков, различающиеся по степени зависимости от других переменных в ряде.

Пусть Y – переменная, в которой присутствуют пропуски;

X – совокупность переменных ряда, отличных от Y (включая X_0 и X_m);

X_0 – подмножество переменных X , не содержащих пропусков (полностью наблюдаемых);

X_m – подмножество переменных X , содержащих пропущенные значения;

R – индикаторная переменная, указывающая на наличие ($R = 1$) или отсутствие ($R = 0$) значения в Y .

Выделяют три основных механизма пропусков.

1. Полностью случайные пропуски (MCAR, Missing Completely At Random). Вероятность появления пропусков не зависит ни от наблюдаемых, ни от ненаблюдаемых данных. Формально это выражается как $P(R | X) = P(R)$, где $P(R)$ – вероятность наличия пропусков, то есть пропуски распределены случайным образом и не зависят от информации, содержащейся в ряде. Этот механизм является наиболее простым для анализа.

2. Случайные пропуски (MAR, Missing At Random). Вероятность пропуска зависит только от наблюдаемых переменных, но не от переменных с пропусками. Формально $P(R | X, Y) = P(R | X_0)$ – такой механизм считается менее проблемным, поскольку позволяет учитывать структуру пропусков при правильном моделировании зависимости от X_0 .

3. Неслучайные пропуски (MNAR, Missing Not At Random). Вероятность пропуска зависит от самой переменной Y или от переменных, которые содержат пропуски X_m . Формально $P(R | X, Y) = P(R | X_0, X_m, Y)$, что делает невозможным корректное моделирование пропусков без дополнительной информации. Такой механизм является наиболее сложным для анализа.

Как отмечено в работе [25], простое удаление отдельных компонент многомерного временного ряда, содержащих пропуски, а также исключение специфических временных точек приводит к потере значительной части данных, нарушает регулярность наблюдений и не учитывает природу возникновения пропусков.

Статистические методы замещения пропущенных значений можно условно разделить на два класса: простейшие, основанные на вычислении базовых статистических характеристик (замещение средним, медианой, модой, использование ближайшего наблюдаемого значения, линейная интерполяция [27]), и более сложные, предполагающие построение модели. Несмотря на простоту реализации, такие методы не учитывают механизм возникновения пропусков и могут приводить к искажению взаимосвязей между переменными.

В работе [28] к более сложным методам относят сплайн- и полиномиальную интерполяцию, сглаживание с использованием фильтра Калмана, множественную линейную регрессию (MLR), замещение на основе матричной факторизации (MF, Matrix Factorization), модель гауссовской смеси (GMM, Gaussian Mixture Model), метод Монте-Карло для марковских цепей (MCMC, Markov Chain Monte Carlo) [29] и анализ главных компонент (PCA, Principal Component Analysis). В дополнение к этому в работе [26] выделены также авторегрессионное интегрированное скользящее среднее (ARIMA) и алгоритм максимизации ожидания (EM, Expectation-Maximization).

2.3. ЗАДАЧА ДЕКОМПОЗИЦИИ

Декомпозиция временного ряда позволяет выделить неслучайные компоненты, отделив их от случайных и несистематических эффектов [30]. На практике ряды обычно раскладываются на такие интерпретируемые составляющие,

как тренд, сезонность и циклы. Тренд отражает долгосрочные изменения, циклический компонент – повторяющиеся колебания с переменным периодом, сезонный – регулярные шаблоны в пределах фиксированного интервала [31].

В процессе декомпозиции временного ряда выделяются его ключевые составляющие, что способствует более глубокому пониманию структуры данных, выявлению скрытых закономерностей и построению интерпретируемых моделей, поскольку каждый компонент может быть проанализирован отдельно. Классический подход к тренд-сезонной декомпозиции основан на аддитивной или мультипликативной модели, где каждая часть аппроксимируется функцией из заданного семейства [32]. STL-декомпозиция (Seasonal and Trend decomposition using Loess), основанная на сглаживании LOESS (Locally Estimated Scatterplot Smoothing) [33], особенно эффективна для работы с нерегулярной сезонностью. Простое использование скользящего среднего (MA, Moving Average) помогает выделить трендовую составляющую [22]. Преобразование Фурье позволяет анализировать временной ряд в частотной области, выделяя периодические компоненты, а вейвлет-декомпозиция фокусируется на многомасштабном анализе и выявлении закономерностей в нестационарных процессах [34]. Метод эмпирического разложения мод (EMD, Empirical Mode Decomposition) позволяет выполнять аддитивное разложение на моды различных частот [35], а анализ с использованием SSA – выделять составляющие ряда на основе внутренней структуры данных [24].

2.4. АНАЛИЗ ВЫБРОСОВ И АНОМАЛИЙ

Обнаружение выбросов – это выявление отдельных наблюдений или их групп, существенно отличающихся от общей структуры данных вследствие шума, ошибок измерения или редких событий, с целью повышения качества анализа и корректной интерпретации результатов. В работе [36] представлены статистические методы обнаружения выбросов, такие как вычисление z -оценки, основанной на стандартном отклонении, межквартильного размаха (IQR, Interquartile Range), и анализ экстремальных значений с использованием порогов, основанных на характеристиках распределения данных. В исследовании [37] рассматриваются критерии Граббса, Q -критерий Диксона и тест Рознера. Альтернативный подход основан на использовании прогнозных моделей (например, ARMA или Prophet [37]) и анализе отклонений между прогнозом и фактическими данными.

Обнаружение аномалий также связано с выявлением отклонений, однако, в отличие от выбросов, наличие аномалий может указывать на сбой, изменение характеристик системы или нарушение ее нормального функционирования. В работе [38] авторы ставят задачу обнаружения аномалий как проблему поиска наблюдений, не соответствующих ожидаемому (нормальному) поведению, то есть не соответствующих распределению наблюдений системы при нормальном режиме функционирования. Во множестве практических приложений аномалии рассматриваются как индикаторы нетипичного поведения, потенциально связанные с нежелательными последствиями и требующие внимания [39].

Задача обнаружения аномалий формально может быть представлена как бинарная классификация временного ряда (или его фрагмента) на «аномаль-

ные» и «нормальные» наблюдения, однако на практике такой подход часто оказывается неэффективным, поскольку аномалии встречаются крайне редко, что приводит к сильному дисбалансу классов и затрудняет разметку данных. В связи с этим для решения задачи обнаружения аномалий чаще применяют методы обучения без разметки (unsupervised learning) [40], что позволяет изучать структуру и распределение нормальных данных, а затем выявлять отклонения от них. Данный факт существенно отличает эту задачу от анализа выбросов, который ориентирован на выявление статистических отклонений без учета общей структуры данных.

В работе [41] описываются три типа аномалий: пунктуальные (или точечные) – отклонения в отдельных временных метках; контекстуальные – наблюдения, являющиеся аномальными только в определенном контексте; коллективные – группы наблюдений, которые в совокупности образуют аномальное поведение. К числу традиционных подходов обнаружения аномалий относят вероятностные, основанные на расстоянии между наблюдениями, а также основанные на отклонениях от прогноза [40].

2.5. ВЫЯВЛЕНИЕ ТОЧЕК ИЗМЕНЕНИЯ

Задача обнаружения точек изменения (CPD, Change Point Detection) – это задача, направленная на выявление моментов времени, когда свойства или структура временного ряда существенно меняются [42]. Такие изменения могут касаться среднего значения, дисперсии, тренда, частотных характеристик, других параметров или же указывать на «разладку» модели, когда ее способность адекватно описывать данные начинает снижаться [43]. Основная цель этого анализа – определить, когда и почему произошли изменения, чтобы лучше понять поведение системы, выявить отклонения или сегментировать временной ряд для дальнейшего изучения и других задач.

Точка времени считается точкой изменения в статистическом анализе, если распределение наблюдаемых данных до и после этой точки статистически значительно отличается. В работе [44] описаны традиционные параметрические методы, такие как CUSUM (Cumulative Sum Control Chart) и метод Ширяева – Робертса, которые требуют полного знания распределений до и после точки изменения. Кроме того, упоминается алгоритм INSPECT (Informative Sparse Projection for Estimating Change-points), использующий разреженное сингулярное разложение матрицы временного ряда (SVD, Singular Value Decomposition) для выявления изменений.

Методы статистического контроля производства (SPC, Statistical Process Control), изначально предназначенные для мониторинга и анализа производственных процессов, успешно применяются и в задаче CPD, поскольку позволяют фиксировать моменты, когда система выходит из состояния статистически управляемого процесса, что эквивалентно изменениям свойств временного ряда. Среди инструментов SPC выделяют контрольные карты Шухарта [45], которые используются для выявления значительных изменений, включая как единичные выбросы, так и устойчивые сдвиги. В свою очередь, графики экспоненциально взвешенного скользящего среднего (EWMA, Exponentially Weighted Moving Average) особенно эффективны для обнаружения медленных или постепенных изменений, таких как смена тренда или увеличение вариабельности данных.

2.6. ОЦЕНКА ОСТАВШЕГОСЯ СРОКА СЛУЖБЫ

Оценка оставшегося срока службы (RUL, Remaining Useful Life) – это задача прогнозирования времени, в течение которого объект, система или оборудование смогут продолжать функционировать в пределах допустимых характеристик до наступления отказа или необходимости ремонта. RUL играет ключевую роль в задачах прогнозирования технического состояния и планирования профилактического обслуживания, способствуя повышению надежности, снижению эксплуатационных затрат и предотвращению внезапных отказов [46].

С прикладной точки зрения задача оценки RUL тесно связана с теорией надежности, в рамках которой определяется вероятность отказа в зависимости от времени или условий эксплуатации. Методы оценки RUL применимы как для моделирования деградационных процессов (например, при снижении производительности, отсутствии профилактического обслуживания), так и для учета плановых изменений, связанных с модернизацией оборудования или внедрением новых технологий [15]. Полученные оценки позволяют прогнозировать вероятность отказа и среднее время до него, а также служат основой для принятия решений по техническому обслуживанию, планированию замены компонентов и снижению простоев [47].

С точки зрения анализа временных рядов задача RUL отличается от общей задачи прогнозирования тем, что она ориентирована на определение времени до наступления отказа или достижения предельного состояния, а не на предсказание значений временного ряда в будущем. Эта задача чаще всего формулируется как задача регрессии, где целевая переменная – это непрерывное время до отказа, но в некоторых случаях может быть представлена как задача классификации, если предсказывается принадлежность объекта к определенным категориям состояния (например, «нормальное», «в зоне риска», «требуется немедленное обслуживание»), обычно такие категории определяются на основе пороговых значений времени до отказа. Отличительной особенностью RUL является необходимость учета сложных деградационных процессов и неопределенности, что делает ее более специализированной по сравнению с общей задачей прогнозирования.

2.7. ЗАДАЧИ КЛАССИФИКАЦИИ И КЛАСТЕРИЗАЦИИ

В задаче классификации многомерных временных рядов рассматривается множество размеченных наблюдений $\{(\mathbf{X}^{(i)}, Y^{(i)})\}_{i=1}^M$, где $\mathbf{X}^{(i)}$ – многомерный временной ряд, а $Y^{(i)}$ – соответствующая ему метка класса. Целью задачи является обучение классификатора $\Psi(\mathbf{X}^{(i)}, \Phi)$ для предсказания неизвестной метки $Y^{(i)}$, где Φ – обучаемые параметры классификатора.

Постановка задачи классификации временного ряда предполагает наличие полностью размеченных данных [48]. В отличие от этого, задача кластеризации временных рядов не требует меток классов и направлена на группировку временных рядов или их сегментов на основе их схожести. Кластеризация позволяет выделить скрытые паттерны или структуры в данных и используется, например, для сегментации наблюдений, выявления групп схожих событий

или автоматической категоризации временных рядов при отсутствии размеченных данных.

До распространения нейросетевых подходов классификация и кластеризация основывались на статистиках, метриках расстояний и классических алгоритмах машинного обучения. В [48] описано применение метода k -ближайших соседей (k -NN, k -nearest neighbors) совместно с методами учета временных деформаций, такими как Dynamic Time Warping (DTW) и евклидовым расстоянием. Подходы, основанные на преобразованиях, предполагают извлечение статистических характеристик (среднего, дисперсии, автокорреляции) [49], выполнение спектрального анализа (вейвлет- или фурье-преобразование) и выделение частотных признаков [34], что позволяет классифицировать временные ряды в преобразованном пространстве с использованием, например, метода опорных векторов (SVM, Support Vector Machine) [50].

2.8. ИЗМЕНЕНИЕ РАЗМЕРНОСТИ ВРЕМЕННЫХ РЯДОВ НЕЙРОСЕТЕВЫМИ МЕТОДАМИ

Глубокие нейронные сети, способные эффективно обрабатывать большие объемы данных, открыли новые возможности для анализа многомерных временных рядов, в том числе за счет гибкого изменения размерности. Одним из таких подходов является аугментация данных (Data Augmentation), направленная на искусственное расширение обучающей выборки путем генерации новых временных рядов или модификации существующих. Для временных данных это может включать, например, наложение шумов, временные сдвиги, масштабирование по амплитуде или комбинации сегментов. Аугментация помогает повысить обобщающую способность модели, особенно в условиях ограниченного объема исходных данных. Эффективность этого подхода ранее была продемонстрирована в таких областях, как компьютерное зрение (CV, Computer Vision) и обработка естественного языка (NLP, Natural Language Processing), где он позволил снизить чрезмерную подгонку модели и повысить устойчивость прогноза [19, 20].

Задача извлечения скрытых информативных признаков (RL, Representation Learning) направлена на автоматическое формирование компактных по размерности представлений многомерных временных рядов с устранением избыточности и шумов [18]. Возможность нейросетей преобразовывать исходные данные во внутренние представления стала основой предварительного обучения, когда модели обучаются выявлению универсальных закономерностей на больших массивах данных, а затем адаптируются к решению прикладных задач. При этом RL не противопоставляется Data Augmentation, а рассматривается как взаимодополняющий подход: в то время как аугментация увеличивает количество обучающих примеров, извлечение скрытых информативных признаков повышает их качество за счет сжатия и упорядочивания признакового пространства.

3. ОПИСАНИЕ АРХИТЕКТУР И ПОДХОДОВ К ОБУЧЕНИЮ

Архитектуры нейронных сетей, используемые в анализе многомерных временных рядов, строятся на основе нескольких фундаментальных типов. Далее перечисляются основные конструкции, лежащие в основе большинства современных решений.

Многослойный перцептрон (MLP, Multi-Layer Perceptron) – это сеть прямого распространения (FFN, Feedforward Network), включающая входной слой, один или несколько скрытых и выходной слой. Хотя архитектуру MLP традиционно считают слабо пригодной для обработки последовательностей, в работе [3] показана ее эффективность при анализе временных рядов.

Рекуррентные нейронные сети (RNN, Recurrent Neural Networks) предназначены специально для обработки последовательностей, наибольшее распространение получили сети Long Short-Term Memory (LSTM) [3], Gated Recurrent Units (GRU) [51] и двунаправленные сети BiLSTM (Bidirectional LSTM) [52]. Механизм Attention, первоначально разработанный для повышения эффективности RNN, стал ключевым элементом архитектуры Transformers [53], демонстрирующей высокую производительность за счет параллельной обработки.

Сверточные нейронные сети (CNN, Convolutional Neural Networks) эффективно извлекают локальные закономерности во временных рядах благодаря операции свертки, что особенно полезно в задачах классификации, замещения пропусков и обнаружения аномалий [3]. Графовые нейронные сети (GNN, Graph Neural Networks) учитывают сложные пространственно-временные зависимости между наблюдениями и компонентами многомерных временных рядов [54].

Отдельную группу составляют гибридные архитектуры, формирующиеся путем объединения различных базовых компонентов. Комбинации позволяют совместить сильные стороны разных конструкций: повысить производительность, улучшить способность к извлечению сложных зависимостей. Некоторые гибриды со временем оформились в самостоятельные архитектурные подходы и активно применяются в прикладных задачах.

Так, в работах [40, 55] для обнаружения аномалий используется ConvLSTM (Convolution LSTM), которая сочетает CNN с LSTM для моделирования пространственно-временных данных. В работе [56] описываются архитектуры, использующие Graph Attention Network (GAT) – объединение GNN с механизмом Attention для адаптивного взвешивания связей между узлами. В работе [57] предложена BNN-LSTM (Bayesian Neural Network LSTM), интегрирующая метод байесовского вывода в модель LSTM.

Следующую категорию составляют мета-архитектуры – обобщенные схемы, в которых базовые компоненты (например, CNN, RNN или Attention) используются как составные блоки. В отличие от гибридных архитектур, где фиксированные типы сетей соединяются для моделирования конкретных аспектов данных, мета-архитектуры задают общий принцип построения модели и допускают гибкую конфигурацию в зависимости от задачи.

К числу наиболее распространенных мета-архитектур относятся конструкции типа encoder-decoder, где сеть encoder (кодирования) извлекает признаки и формирует компактное скрытое пространство данных, а сеть decoder (декодирования) восстанавливает или генерирует данные на его основе. К ним относят автокодировщики (AE, Autoencoders) и вариационные автокодировщики (VAE, Variational Autoencoders), которые эффективно используются для обнаружения аномалий [41]. Для реализации encoder и decoder используются различные базовые сети – CNN [58], RNN [2], а также их комбинации с Attention [40, 55]. К этому типу архитектур относится и архитектура Transformers, отличающаяся механизмом self-attention и параллельной обработкой, что делает ее эффективной для длинных последовательностей в задачах анализа текста, изображений и временных рядов [3, 46].

Мета-архитектура генеративной состязательной сети (GAN, Generative Adversarial Networks), впервые представленная в [59], подобно encoder-decoder, включает генератор, создающий новые данные, и дискриминатор, оценивающий их качество. Первоначально GAN применялись в задачах компьютерного зрения, но впоследствии нашли применение и в других областях, включая обработку текстов и генерацию временных рядов преимущественно с использованием рекуррентных сетей в роли генератора и дискриминатора. В одной из первых работ по применению GAN для генерации временных последовательностей была продемонстрирована возможность синтеза музыкальных данных [60], а в [20] при построении модели дополнительно учитывались временные метки, что позволило улучшить статистические характеристики получаемых последовательностей. Позднее были разработаны модификации, такие как Wasserstein GAN (WGAN) [61], способные захватывать более длительные временные зависимости [62].

Традиционный подход к обучению нейросетевых моделей предполагает построение отдельного экземпляра модели для каждой задачи анализа временных рядов, даже если разные задачи решаются на одном наборе данных. Использование подхода Representation Learning позволяет автоматически извлекать наиболее информативные признаки, формируя универсальные и компактные описания данных, на основе которых строятся модели, способные одновременно прогнозировать несколько выходных переменных разного типа («многие-ко-многим», «многие-к-одному») [57].

Для решения новых задач на основе уже обученных моделей применяется перенос знаний (transfer learning) и тонкая настройка (fine-tuning), при которых предварительно обученные модели адаптируются к специфическим задачам с минимальными затратами времени и ресурсов. Наиболее ярким примером применения таких моделей являются большие языковые модели (LLM, Large Language Models), предварительно обученные на огромных наборах текстовых данных [9].

Предварительное обучение бывает двух типов: *контролируемое* (supervised learning), когда для каждого наблюдения требуется аннотированная метка, и *самоконтролируемое* (self-supervised learning), при котором метки генерируются автоматически без участия человека [64]. Поскольку разметка больших объемов данных является затратной и трудоемкой задачей, подходы self-supervised приобрели популярность: псевдометки получают с помощью искусственных преобразований исходных данных (маскирование элементов, добавление шума, контрастные пары и другие методы [63]). Полученные таким образом преобразования данных сохраняют важнейшие закономерности и успешно переносятся на решение новых задач.

Кроме того, авторы [64] выделяют два класса методов обучения преобразованиям данных: генеративные, основанные на моделировании исходного распределения данных и последующей генерации новых образцов, и дискриминативные, направленные на различение исходных и трансформированных данных путем прогнозирования замаскированных элементов, определения примененных преобразований или сравнения пар наблюдений.

Целевая задача не обязательно должна быть регрессионной, чтобы можно было отказаться от использования оригинальных меток или классов при применении предварительно обученной модели. В таких случаях модель может быть напрямую использована для решения задачи, что называется обучением с *нулевым выстрелом* (zero-shot learning). Однако наилучшие результаты, как

правило, достигаются при дополнительной тонкой настройке модели на небольшом наборе аннотированных данных, что позволяет повысить ее точность и адаптировать к специфическим требованиям задачи.

4. АНАЛИЗ РЕЗУЛЬТАТОВ

Обобщение и систематизация рассмотренных ранее аспектов позволяет упорядочить подходы к анализу многомерных временных рядов и облегчить выбор методов. В таблице описаны основные задачи анализа временных рядов, кратко изложена суть каждой задачи, указаны типы входных и выходных данных, а также отражена необходимость предварительной разметки наблюдений.

Систематизация задач анализа многомерных временных рядов

Systematization of multivariate time series analysis tasks

№ п/п	Задача анализа	Основная идея	Вид входных данных	Вид выходных данных	Нужна разметка
1	Прогнозирование (экстраполяция)	Прогнозирование будущих значений на основе предыдущих наблюдений	Временной ряд (целиком или частично)	Временной ряд (одномерный или многомерный)	Нет
2	Замещение пропусков (экстраполяция)	Восстановление пропущенных значений на основе исторических данных	Временной ряд с пропусками	Восстановленный временной ряд	Нет
3	Декомпозиция временного ряда	Разложение временного ряда на отдельные интерпретируемые компоненты	Полный исходный временной ряд	Компоненты ряда (тренд, сезонность и др.)	Нет
4	Анализ выбросов	Выявление точек или групп точек, существенно отклоняющихся от общей структуры данных	Временной ряд (целиком или частично)	Отдельные временные точки или их группы	Нет
5	Обнаружение аномалий	Выявление наблюдений, не соответствующих нормальному поведению	Временной ряд (целиком или частично)	Бинарная метка (аномалия или норма)	Зависит от условий
6	Выявление точек изменения	Выявление точек времени, когда свойства и/или структура временного ряда существенно меняется	Полный исходный временной ряд	Одна или несколько временных точек	Нет
7	Оценка оставшегося срока службы	Прогнозирование времени до отказа наблюдаемого объекта	Полный исходный временной ряд	Прогноз остаточного ресурса (время до отказа)	Нет

Окончание таблицы

End of the Table

№ п/п	Задача анализа	Основная идея	Вид входных данных	Вид выходных данных	Нужна разметка
8	Классификация	Присвоение временному ряду (или его фрагменту) метки класса из заранее известного набора	Временной ряд или (целиком частично)	Категориальная метка	Да
9	Кластеризация	Группировка временных рядов или сегментов на основе их внутренней схожести	Временной ряд или (целиком частично)	Метка группы объектов	Нет
10	Representation learning	Автоматизированный процесс представления исходных данных в пространстве меньшей размерности	Полный исходный временной ряд	Представление данных в новом пространстве	Зависит от условий
11	Data Augmentation	Искусственное увеличение набора данных путем модификации исходных или генерации синтетических наблюдений	Полный исходный временной ряд	Синтетический временной ряд	Нет

Предложенная классификация позволяет соотнести тип данных и цель анализа с подходящими методами обработки. Архитектуры нейросетей не включены в перечень, поскольку современные модели часто объединяют несколько базовых решений, что затрудняет их однозначное структурирование. В отличие от архитектур, способ обучения модели напрямую определяется наличием или отсутствием разметки.

ЗАКЛЮЧЕНИЕ

В работе предложена классификация ключевых компонентов решений для анализа многомерных временных рядов, формируемых на основе данных. Рассмотрены типы рядов, категории задач, архитектуры нейронных сетей и подходы к обучению, что позволяет представить целостную структуру существующих методов. Предложенная классификация может стать основой для создания баз знаний, экспертных и рекомендательных систем и способствовать принятию решений и поиску перспективных направлений развития исследований.

Применение глубоких нейронных сетей стало основой перехода от гибридных методов, сочетающих подходы классической статистики и алгоритмы машинного обучения, к полностью автоматизированному процессу – от обработки сырых данных до получения прогноза. Такой подход позволил эффективно работать не только с многомерными числовыми рядами, но и с неструктурированной информацией (текстом, звуком, изображениями), преобразуя ее в пригодную для анализа форму. Кроме того, нейросетевые модели способны решать задачи с различным числом входов и выходов («многие-ко-многим», «многие-

к-одному») без необходимости перестройки архитектуры, что было недоступно для классических методов.

Несмотря на многочисленные преимущества, методы глубокого обучения имеют существенный недостаток – низкую интерпретируемость моделей. Если для статистических методов объяснение результатов было естественным, а в традиционном машинном обучении эта задача уже вызывала трудности, то с ростом сложности глубоких нейронных сетей проблема стала еще острее. Отсутствие прозрачности усложняет доверие к результатам в критически важных областях, таких как, например, производство, навигация, медицина, финансы. Эти соображения делают разработку методов улучшения интерпретируемости результатов одной из ключевых задач для дальнейшего развития методов глубокого обучения.

СПИСОК ЛИТЕРАТУРЫ

1. Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование: учебник. – Москва: Финансы и статистика, 2012. – 320 с.
2. Anomaly detection in time series data and its application to semiconductor manufacturing / R. Hwang, S. Park, Y. Bin, H.J. Hwang // IEEE Access. – 2023. – Vol. 11. – P. 130483–130490. – DOI: 10.1109/ACCESS.2023.3333247.
3. Deep time series models: A comprehensive survey and benchmark / Y. Wang, H. Wu, J. Dong, Y. Liu, M. Long, J. Wang. – URL: <https://arxiv.org/abs/2407.13278> (accessed: 09.09.2025).
4. Deep learning for multivariate time series imputation: A survey / J. Wang, W. Du, Y. Yang, L. Qian, W. Cao, K. Zhang, W. Wang, Y. Liang, Q. Wen. – URL: <https://arxiv.org/abs/2402.04059> (accessed: 09.09.2025).
5. Online model-based anomaly detection in multivariate time series: Taxonomy, survey, research challenges and future directions / L. Correia, J.-C. Goos, P. Klein, T. Bäck, A.V. Kononova // Engineering Applications of Artificial Intelligence. – 2024. – Vol. 138, pt. A. – P. 109323. – DOI: 10.1016/j.engappai.2024.109323.
6. Deep learning for time series forecasting: Tutorial and literature survey / K. Benidis, S.S. Rangapuram, V. Flunkert, Y. Wang, D. Maddix, C. Turkmen, J. Gasthaus, M. Bohlke-Schneider, D. Salinas, L. Stella, F.-X. Aubet, L. Callot, T. Januschowski // ACM Computing Surveys. – 2022. – Vol. 55 (121). – P. 1–36. – DOI: 10.1145/3533382.
7. Generative adversarial networks in time series: A survey and taxonomy / E. Brophy, Z. Wang, Q. She, T. Ward. – URL: <https://arxiv.org/abs/2107.11098> (accessed: 09.09.2025).
8. Large language models for time series: A survey / X. Zhang, R.R. Chowdhury, R.K. Gupta, J. Shang // Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24). – Jeju, South Korea, 2024. – P. 8335–8343.
9. Foundation models for time series analysis: A tutorial and survey / Y. Liang, H. Wen, Y. Nie, Y. Jiang, M. Jin, D. Song, S. Pan, Q. Wen // Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24). – New York, 2024. – P. 6555–6565. – DOI: 10.1145/3637528.3671451.
10. Bivand R.S., Pebesma E., Gómez-Rubio V. Applied spatial data analysis with R. – New York: Springer, 2013. – 405 p. – DOI: 10.1007/978-1-4614-7618-4.
11. ContiFormer: Continuous-time transformer for irregular time series modeling / Y. Chen, K. Ren, Y. Wang, Y. Fang, W. Sun, D. Li. – URL: <https://arxiv.org/abs/2402.10635> (accessed: 09.09.2025).
12. Lipton Z.C., Kale D., Wetzel R. Directly modeling missing data in sequences with RNNs: Improved classification of clinical time series // Proceedings of Machine Learning Research. – 2016. – Vol. 56. – P. 253–270.
13. Guo C., Berkahn F. Entity embeddings of categorical variables. – URL: <https://arxiv.org/abs/1604.06737> (accessed: 09.09.2025).
14. Хеннан Э. Многомерные временные ряды. – М.: Мир, 1974. – 576 с.
15. Khusnutdinov A., Karmanov V. Application of machine learning methods for predicting the dynamics of consumption of fuel and energy resources and assessing the effect of energy saving measures // AIP Conference Proceedings. – 2023. – Vol. 2999 (1). – P. 020062. – DOI: 10.1063/5.0158571.

16. Хуснутдинов А.О., Карманов В.С. Использование методов машинного обучения для прогнозирования временных рядов при планировании потребления энергетических ресурсов // Вестник Ошского государственного университета. Математика. Физика. Техника. – 2023. – № 1. – С. 220–232. – DOI: 10.52754/16948645_2023_1_220.
17. Anomaly detection in coastal wireless sensors via efficient deep sequential learning / M. Matar, T. Xia, K. Huguenard, D. Huston, S. Wshah // IEEE Access. – 2023. – Vol. 11. – P. 110260–110271. – DOI: 10.1109/ACCESS.2023.3322370.
18. Radford A., Metz L., Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. – URL: <https://arxiv.org/abs/1511.06434> (accessed: 09.09.2025).
19. Ilbert R., Hoang T.V., Zhang Z. Data augmentation for multivariate time series classification: An experimental study. – URL: <https://arxiv.org/abs/2406.06518> (accessed: 09.09.2025).
20. T-CGAN: Conditional generative adversarial network for data augmentation in noisy time series with irregular sampling / G. Ramponi, P. Protopapas, M. Brambilla, R. Janssen. – URL: <https://arxiv.org/abs/1811.08295> (accessed: 09.09.2025).
21. Aminikhanghahi S., Cook D.J. A survey of methods for time series change point detection // Knowledge and Information Systems. – 2017. – Vol. 51. – P. 339–367. – DOI: 10.1007/s10115-016-0987-z.
22. Wei W.W.S. Multivariate time series analysis and applications. – Oxford: John Wiley & Sons, 2019. – 536 p.
23. Assessment of length-of-day and universal time predictions based on the results of the second earth orientation parameters prediction comparison campaign / J. Śliwińska-Bronowicz, T. Kur, M. Wińska, et al. // Journal of Geodesy. – 2024. – Vol. 98 (22). – DOI: 10.1007/s00190-024-01824-7.
24. Модификация базового метода сингулярного спектрального анализа для повышения точности прогнозирования неравномерности вращения Земли / А.К. Гречкосеев, А.С. Толстиков, В.М. Тиссен, В.С. Карманов, А.И. Ваганова // Вычислительные технологии. – 2020. – Т. 25, № 3. – С. 54–65. – DOI: 10.25743/ICT.2020.25.3.007.
25. Umar N., Gray A. Comparing single and multiple imputation approaches for missing values in univariate and multivariate water level data // Water. – 2023. – Vol. 15 (8). – P. 1519. – DOI: 10.3390/w15081519.
26. A survey on multivariate time series imputation using adversarial learning / A. Richter, J. Ijaradar, U. Wetzker, V. Jain, A. Frotzschner // IEEE Access. – 2024. – Vol. 12. – P. 148167–148189. – DOI: 10.1109/ACCESS.2024.3473540.
27. Preptimize: Automation of time series data preprocessing and forecasting / M. Usmani, Z.A. Memon, A. Zulfiqar, R. Qureshi // Algorithms. – 2024. – Vol. 17 (8). – P. 332. – DOI: 10.3390/a17080332.
28. Scalable tensor factorizations with missing data / E. Acar, D.M. Dunlavy, T.G. Kolda, M. Mørup // Proceedings of the SDM10: 2010 SIAM International Conference on Data Mining. – Ohio, 2010. – P. 701–712.
29. Gabrielski J., Häger U. A Markov chain model for imputation of electricity consumption time series // 2023 58th International Universities Power Engineering Conference (UPEC). – IEEE, 2023. – DOI: 10.1109/UPEC57427.2023.10294910.
30. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting / H. Wu, J. Xu, J. Wang, M. Long. – URL: <https://arxiv.org/abs/2106.13008> (accessed: 09.09.2025).
31. Карманов В.С. Анализ временных рядов: конспект лекций. – Новосибирск: Изд-во НГТУ, 2007. – 28 с.
32. Taylor S.J., Letham B. Forecasting at scale // PeerJ Preprints. – 2017. – DOI: 10.7287/peerj.preprints.3190v2.
33. STL: A seasonal-trend decomposition procedure based on loess / R.B. Cleveland, W.S. Cleveland, J.E. McRae, I. Terpenning // Journal of Official Statistics. – 1990. – Vol. 6 (1). – P. 3–73.
34. Forecasting electricity consumption of electrical machines of a coal industry enterprise using the wavelet transform / V.Z. Manusov, D.V. Orlov, V.S. Karmanov, A.O. Khusnutdinov, E.S. Tretyakova, D.V. Korolev // IEEE 23rd International Conference of Young Professionals in Electron Devices and Materials (EDM), Altai, Russian Federation. – 2022. – DOI: 10.1109/EDM55285.2022.9855175.
35. Hu H., Xu W. Time series forecasting based on empirical mode decomposition and the varying-coefficient DBN-AR model // IEEE Access. – 2022. – Vol. 10. – P. 105169–105181. – DOI: 10.1109/ACCESS.2022.3210974.
36. Yaro A.S., Maly F., Prazak P. Outlier detection in time-series receive signal strength observation using Z-score method with S_n scale estimator for indoor localization // Applied Sciences. – 2023. – Vol. 13 (6). – P. 3900. – DOI: 10.3390/app13063900.

37. *Shabou S.* Time Series with R. – URL: <https://s-ai-f.github.io/Time-Series/> (accessed: 10.09.2025).
38. *Karadayı Y., Aydin M.N., Öğrenci A.S.* A hybrid deep learning framework for unsupervised anomaly detection in multivariate spatio-temporal data // *Applied Sciences*. – 2020. – Vol. 10 (15). – P. 5191. – DOI: 10.3390/app10155191.
39. *Fawwaz D.Z., Chung S.H.* Real-time and robust hydraulic system fault detection via edge computing // *Applied Sciences*. – 2020. – Vol. 10 (17). – P. 5933. – DOI: 10.3390/app10175933.
40. An attention-based ConvLSTM autoencoder with dynamic thresholding for unsupervised anomaly detection in multivariate time series / T. Tayeh, S. Aburakhia, R. Myers, A. Shami // *Machine Learning and Knowledge*. – 2022. – Vol. 4 (2). – P. 350–370. – DOI: 10.3390/make4020015.
41. *Pota M., De Pietro G., Esposito M.* Real-time anomaly detection on time series of industrial furnaces: A comparison of autoencoder architectures // *Engineering Applications of Artificial Intelligence*. – 2023. – Vol. 124. – P. 106597. – DOI: 10.1016/j.engappai.2023.106597.
42. *Aminikhanghahi S., Cook D.J.* A survey of methods for time series change point detection // *Knowledge and Information Systems*. – 2017. – Vol. 51 (2). – P. 339–367. – DOI: 10.1007/s10115-016-0987-z.
43. *Антонова Е.Н.* Последовательный алгоритм обнаружения разладки многомерных временных рядов с высокой динамикой // *Известия СПбГТИ (ТУ)*. – 2022. – № 63 (89). – С. 93–99. – DOI: 10.36807/1998-9849-2022-63-89-93-99.
44. Multivariate time series change-point detection with a novel Pearson-like scaled Bregman divergence / T. Si, Y. Wang, L. Zhang, E. Richmond, T.-H. Ahn, H. Gong // *Stats*. – 2024. – Vol. 7 (2). – P. 462–480. – DOI: 10.3390/stats7020028.
45. *Hawkins M.D., Qiu P.* The changepoint model for statistical process control // *Journal of Quality Technology*. – 2003. – Vol. 35 (4). – DOI: 10.1080/00224065.2003.11980233.
46. Shared temporal attention transformer for remaining useful lifetime estimation / G.S. Chadha, S.R.B. Shah, A. Schwung, S.X. Ding // *IEEE Access*. – 2022. – Vol. 10. – P. 74244–74258. – DOI: 10.1109/ACCESS.2022.3187702.
47. *Gómez-Pau Á., Riba J.-R., Moreno-Eguilaz M.* Time series RUL estimation of medium voltage connectors to ease predictive maintenance plans // *Applied Sciences*. – 2020. – Vol. 10 (24). – P. 9041. – DOI: 10.3390/app10249041.
48. *Rahman M.M., Farahani M.A., Wuest T.* Multivariate time-series classification of critical events from industrial drying hopper operations: A deep learning approach // *Journal of Manufacturing and Materials Processing*. – 2023. – Vol. 7 (5). – P. 164. – DOI: 10.3390/jmmp7050164.
49. *Lei Y., Wu Z.* Time series classification based on statistical features // *Wireless Communications and Networking*. – 2020. – Vol. 46. – DOI: 10.1186/s13638-020-1661-4.
50. Time series classification using support vector machine with Gaussian elastic metric kernel / D. Zhang, W. Zuo, D. Zhang, H. Zhang // *20th International Conference on Pattern Recognition*. – Istanbul, 2010. – P. 29–32. – DOI: 10.1109/ICPR.2010.16.
51. Recurrent neural networks for multivariate time series with missing values / Z. Che, S. Purushotham, K. Cho, D. Sontag, Y. Liu. – URL: <https://arxiv.org/abs/1606.01865> (accessed: 10.09.2025).
52. Remaining useful life estimation in prognostics using deep bidirectional LSTM neural network / J. Wang, G. Wen, S. Yang, Y. Liu // *Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing)*. – Chongqing, 2018. – P. 1037–1042.
53. Attention is all you need / A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin. – URL: <https://arxiv.org/abs/1706.03762> (accessed: 10.09.2025).
54. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection / M. Jin, H.Y. Koh, Q. Wen, D. Zambon, C. Alippi, G.I. Webb, I. King, S. Pan. – URL: <https://arxiv.org/abs/2307.03759> (accessed: 10.09.2025).
55. TimeTector: A twin-branch approach for unsupervised anomaly detection in livestock sensor noisy data (TT-TBAD) / J.K. Kakar, S. Hussain, S.C. Kim, H. Kim // *Sensors*. – 2024. – Vol. 24 (8). – P. 2453. – DOI: 10.3390/s24082453.
56. GTAD: Graph and temporal neural network for multivariate time series anomaly detection / S. Guan, B. Zhao, Z. Dong, M. Gao, Z. He // *Entropy*. – 2022. – Vol. 24 (6). – P. 759. – DOI: 10.3390/e24060759.
57. Long short-term memory Bayesian neural network for air pollution forecast / H.A.D. Nguyen, Q.P. Ha, H. Duc, M. Azzi, N. Jiang, X. Barthelemy, M. Riley // *IEEE Access*. – 2023. – Vol. 11. – P. 35710–35725. – DOI: 10.1109/ACCESS.2023.3265725.

58. Unsupervised anomaly detection of industrial robots using sliding-window convolutional variational autoencoder / T. Chen, X. Liu, B. Xia, W. Wang, Y. Lai // IEEE Access. – 2020. – Vol. 8. – P. 47072–47081. – DOI: 10.1109/ACCESS.2020.2977892.

59. Generative adversarial networks / I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio. – URL: <https://arxiv.org/abs/1406.2661> (accessed: 10.09.2025).

60. *Mogren O.* C-RNN-GAN: Continuous recurrent neural networks with adversarial training. – URL: <https://arxiv.org/abs/1611.09904> (accessed: 10.09.2025).

61. Conditional Sig-Wasserstein GANs for time series generation / S. Liao, H. Ni, L. Szpruch, M. Wiese, M. Sabate-Vidales, B. Xiao. – URL: <https://arxiv.org/abs/2006.05421> (accessed: 10.09.2025).

62. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks / D. Li, D. Chen, L. Shi, B. Jin, J. Goh, S.-K. Ng // Lecture Notes in Computer Science. – 2019. – Vol. 11730. – P. 703–716. – DOI: 10.1007/978-3-030-30490-4_56.

63. Self-supervised representation learning: Introduction, advances and challenges / L. Ericsson, H. Gouk, C.C. Loy, T.M. Hospedales. – URL: <https://arxiv.org/abs/2110.09327> (accessed: 10.09.2025).

64. Beyond just vision: A review on self-supervised representation learning on multimodal and temporal data / S. Deldari, H. Xue, A. Saeed, J. He, D.V. Smith, F.D. Salim. – URL: <https://arxiv.org/abs/2206.02353> (accessed: 10.09.2025).

Хуснутдинов Александр Олегович, аспирант кафедры теоретической и прикладной информатики факультета прикладной математики и информатики Новосибирского государственного технического университета. Основное направление научных исследований: анализ временных данных, машинное обучение, глубокое обучение. E-mail: evolext@gmail.com

Хабаров Валерий Иванович, доктор технических наук, профессор кафедры теоретической и прикладной информатики факультета прикладной математики и информатики Новосибирского государственного технического университета. Заведующий кафедрой информационных технологий на транспорте Сибирского государственного университета путей сообщения. Основные направления научных исследований: искусственный интеллект, планирование эксперимента, статистические методы анализа данных, математическое моделирование транспортных потоков. Имеет более 150 публикаций. E-mail: khabarov51@mail.ru

Карманов Виталий Сергеевич, кандидат технических наук, доцент кафедры теоретической и прикладной информатики факультета прикладной математики и информатики Новосибирского государственного технического университета. Основное направление научных исследований: прикладная статистика; анализ данных; методы и технологии машинного обучения и искусственного интеллекта; моделирование, оптимизация, системный анализ, цифровые технологии, разработка информационных систем и программных комплексов, автоматизация промышленного производства и технологических процессов. Имеет более 80 публикаций. E-mail: karmanov@corp.nstu.ru

Khusnutdinov Alexander O., PhD student at the department of theoretical and applied informatics at the faculty of applied mathematics, Novosibirsk State Technical University. His main research fields include temporal data analysis, machine learning, and deep learning. E-mail: evolext@gmail.com

Khabarov Valery I., D.Sc. (Eng.), professor at the department of theoretical and applied informatics at the faculty of applied mathematics, Novosibirsk State Technical University. He is the head of the Department of Information Technologies in Transportation at the Siberian State Transport University. His main research fields include artificial intelligence, experiment design, statistical methods of data analysis and mathematical modeling of traffic flows. He is the author of more than 150 publications. E-mail: khabarov51@mail.ru

Karmanov Vitaly S., PhD (Eng.), associate professor at the department of theoretical and applied informatics at the faculty of applied mathematics, Novosibirsk State Technical University. His main research fields include applied statistics; data analysis; methods and technologies of machine learning and artificial intelligence; modeling, optimization, system analysis, digital technologies, development of software and information systems, automation of industrial production and technological processes. He is the author of more than 80 publications. E-mail: karmanov@corp.nstu.ru

Deep learning for multivariate time series analysis: systematization of data types, tasks, architectures and approaches*

A.O. KHUSNUTDINOV^{1,a}, V.I. KHABAROV^{1,2,b}, V.S. KARMANOV^{1,c}

¹ Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation

² Siberian State Transport University, 191 Dusi Kovalchuk Street, Novosibirsk, 630049, Russian Federation

^a evolxt@gmail.com ^b khabarov51@mail.ru ^c karmanov@corp.nstu.ru

Abstract

Data analysis methods are currently applied to a wide range of heterogeneous information, including time series, spatial and spatio-temporal structures, texts, images, video, and audio signals and their various combinations. In the context of time series analysis, the application of models based on neural networks allows overcoming the limitations of classical methods and extending the field of application by taking into account data from heterogeneous sources, training universal models for different problems and designing new formulations. Nevertheless, the diversity of application areas and source data leads to the emergence of many approaches to the construction of neural networks adapted to a specific problem, which forces researchers either to develop a new configuration of networks or to conduct a large-scale search for a suitable model.

The paper is devoted to the systematization of approaches to multivariate time series analysis using neural network models to identify patterns in data, i.e. applying deep learning methods. The main aspects including types of multivariate time series, types of analysis tasks, ways of organizing neural network components or their architectures, and approaches to model building are discussed. Types of multivariate time series such as standard, spatio-temporal, regular and others are highlighted. The basic problems of time series analysis are systematized, including traditional problems solved by classical approaches, as well as problems made possible by the introduction of machine learning methods, and new approaches developed through deep learning. This paper presents an overview of the key types of neural networks, popular architectures based on them, and methods for building or training models using these architectures.

This paper does not aim to cover all aspects of multivariate time series analysis, but proposes a classification that can serve as a basis for developing a knowledge base in this area and identifying promising areas for further research.

Keywords: time series analysis, multivariate time series, forecasting, classification, heterogeneous data, machine learning, deep learning, representation learning, neural networks

REFERENCES

1. Afanas'ev V.N., Yuzbashev M.M. *Analiz vremennykh ryadov i prognozirovanie* [Analysis of time series and forecasting]. Moscow, Finansy i statistika Publ., 2012. 320 p.
2. Hwang R., Park S., Bin Y., Hwang H.J. Anomaly detection in time series data and its application to semiconductor manufacturing. *IEEE Access*, 2023, vol. 11, pp. 130483–130490. DOI: 10.1109/ACCESS.2023.3333247.
3. Wang Y., Wu H., Dong J., Liu Y., Long M., ang J. *Deep time series models: A comprehensive survey and benchmark*. Available at: <https://arxiv.org/abs/2407.13278> (accessed 09.09.2025).
4. Wang J., Du W., Yang Y., Qian L., Cao W., Zhang K., Wang W., Liang Y., Wen Q. *Deep learning for multivariate time series imputation: A survey*. Available at: <https://arxiv.org/abs/2402.04059> (accessed 09.09.2025).
5. Correia L., Goos J.-C., Klein P., Bäck T., Kononova A.V. Online model-based anomaly detection in multivariate time series: Taxonomy, survey, research challenges and future directions. *Engineering Applications of Artificial Intelligence*, 2024, vol. 138, pt. A, p. 109323. DOI: 10.1016/j.engappai.2024.109323.
6. Benidis K., Rangapuram S.S., Flunkert V., Wang Y., Maddix D., Turkmen C., Gasthaus J., Bohlke-Schneider M., Salinas D., Stella L., Aubet F.-X., Callot L., Januschowski T. Deep learning for time series forecasting: Tutorial and literature survey. *ACM Computing Surveys*, 2022, vol. 55 (121), pp. 1–36. DOI: 10.1145/3533382.

* Received 30 December 2024.

7. Brophy E., Wang Z., She Q., Ward T. *Generative adversarial networks in time series: A survey and taxonomy*. Available at: <https://arxiv.org/abs/2107.11098> (accessed 09.09.2025).
8. Zhang X., Chowdhury R.R., Gupta R.K., Shang J. Large language models for time series: A survey. *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)*, Jeju, South Korea, 2024, pp. 8335–8343.
9. Liang Y., Wen H., Nie Y., Jiang Y., Jin M., Song D., Pan S., Wen Q. Foundation models for time series analysis: A tutorial and survey. *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*. New York, 2024, pp. 6555–6565. DOI: 10.1145/3637528.3671451.
10. Bivand R.S., Pebesma E., Gómez-Rubio V. *Applied spatial data analysis with R*. New York, Springer, 2013. 405 p. DOI: 10.1007/978-1-4614-7618-4.
11. Chen Y., Ren K., Wang Y., Fang Y., Sun W., Li D. ContiFormer: Continuous-time transformer for irregular time series modeling. Available at: <https://arxiv.org/abs/2402.10635> (accessed 09.09.2025).
12. Lipton Z.C., Kale D., Wetzel R. Directly modeling missing data in sequences with RNNs: Improved classification of clinical time series. *Proceedings of Machine Learning Research*, 2016, vol. 56, pp. 253–270.
13. Guo C., Berkahn F. *Entity embeddings of categorical variables*. Available at: <https://arxiv.org/abs/1604.06737> (accessed 09.09.2025).
14. Hannan E.J. *Mnogomernye vremennye ryady* [Multiple time series]. Moscow, Mir Publ., 1974. 576 p. (In Russian).
15. Khusnutdinov A., Karmanov V. Application of machine learning methods for predicting the dynamics of consumption of fuel and energy resources and assessing the effect of energy saving measures. *AIP Conference Proceedings*, 2023, vol. 2999 (1), p. 020062. DOI: 10.1063/5.0158571.
16. Khusnutdinov A.O., Karmanov V.S. Ispol'zovanie metodov mashinnogo obucheniya dlya prognozirovaniya vremennykh ryadov pri planirovani potrebleniya energeticheskikh resursov [Application of machine learning methods for time series forecasting in energy resource consumption planning]. *Vestnik Oshskogo gosudarstvennogo universiteta. Matematika. Fizika. Tekhnika = Journal of Osh State University. Mathematics. Physics. Technical Sciences*, 2023, no. 1, pp. 220–232. DOI: 10.52754/16948645_2023_1_220.
17. Matar M., Xia T., Huguenard K., Huston D., Wshah S. Anomaly detection in coastal wireless sensors via efficient deep sequential learning. *IEEE Access*, 2023, vol. 11, pp. 110260–110271. DOI: 10.1109/ACCESS.2023.3322370.
18. Radford A., Metz L., Chintala S. *Unsupervised representation learning with deep convolutional generative adversarial networks*. Available at: <https://arxiv.org/abs/1511.06434> (accessed 09.09.2025).
19. Ilbert R., Hoang T.V., Zhang Z. *Data augmentation for multivariate time series classification: An experimental study*. Available at: <https://arxiv.org/abs/2406.06518> (accessed 09.09.2025).
20. Ramponi G., Protopapas P., Brambilla M., Janssen R. *T-CGAN: Conditional generative adversarial network for data augmentation in noisy time series with irregular sampling*. Available at: <https://arxiv.org/abs/1811.08295> (accessed 09.09.2025).
21. Aminikhanghahi S., Cook D.J. A survey of methods for time series change point detection. *Knowledge and Information Systems*, 2017, vol. 51, pp. 339–367. DOI: 10.1007/s10115-016-0987-z.
22. Wei W.W.S. *Multivariate time series analysis and applications*. Oxford, John Wiley & Sons, 2019. 536 p.
23. Śliwińska-Bronowicz J., Kur T., Wińska M., et al. Assessment of length-of-day and universal time predictions based on the results of the second earth orientation parameters prediction comparison campaign. *Journal of Geodesy*, 2024, vol. 98 (22). DOI: 10.1007/s00190-024-01824-7.
24. Grechkoseev A.K., Tolstikov A.S., Tissen V.M., Karmanov V.S., Vaganova A.I. Modifikatsiya bazovogo metoda singulyarnogo spektral'nogo analiza dlya povysheniya tochnosti prognozirovaniya neravnomernosti vrashcheniya Zemli [The modification of the singular spectral analysis basic method for improving the accuracy of forecasting the irregularity of the Earth rotation]. *Vychislitel'nye tekhnologii = Computational Technologies*, 2020, vol. 25, no. 3, pp. 54–65. DOI: 10.25743/ICT.2020.25.3.007.
25. Umar N., Gray A. Comparing single and multiple imputation approaches for missing values in univariate and multivariate water level data. *Water*, 2023, vol. 15 (8), p. 1519. DOI: 10.3390/w15081519.
26. Richter A., Ijaradar J., Wetzker U., Jain V., Frotzsch A. A survey on multivariate time series imputation using adversarial learning. *IEEE Access*, 2024, vol. 12, pp. 148167–148189. DOI: 10.1109/ACCESS.2024.3473540.
27. Usmani M., Memon Z.A., Zulfiqar A., Qureshi R. Preptimize: Automation of time series data preprocessing and forecasting. *Algorithms*, 2024, vol. 17 (8), p. 332. DOI: 10.3390/a17080332.

28. Acar E., Dunlavy D.M., Kolda T.G., Mørup M. Scalable tensor factorizations with missing data. *Proceedings of the SDM10: 2010 SIAM International Conference on Data Mining*, Ohio, 2010, pp. 701–712.
29. Gabrielski J., Häger U. A Markov chain model for imputation of electricity consumption time series. *2023 58th International Universities Power Engineering Conference (UPEC)*. IEEE, 2023. DOI: 10.1109/UPEC57427.2023.10294910.
30. Wu H., Xu J., Wang J., Long M. *Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting*. Available at: <https://arxiv.org/abs/2106.13008> (accessed 09.09.2025).
31. Karmanov V.S. *Analiz vremennykh ryadov* [Analysis of time series]. Novosibirsk, NSTU Publ., 2007. 28 p.
32. Taylor S.J., Letham B. Forecasting at scale. *PeerJ Preprints*. 2017. DOI: 10.7287/peerj.preprints.3190v2.
33. Cleveland R.B., Cleveland W.S., McRae J.E., Terpenning I. STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 1990, vol. 6 (1), pp. 3–73.
34. Manusov V.Z., Orlov D.V., Karmanov V.S., Khusnutdinov A.O., Tretyakova E.S., Korolev D.V. Forecasting electricity consumption of electrical machines of a coal industry enterprise using the wavelet transform. *IEEE 23rd International Conference of Young Professionals in Electron Devices and Materials (EDM)*, Altai, Russian Federation, 2022. DOI: 10.1109/EDM55285.2022.9855175.
35. Hu H., Xu W. Time series forecasting based on empirical mode decomposition and the varying-coefficient DBN-AR model. *IEEE Access*, 2022, vol. 10, pp. 105169–105181. DOI: 10.1109/ACCESS.2022.3210974.
36. Yaro A.S., Maly F., Prazak P. Outlier detection in time-series receive signal strength observation using Z-score method with S_n scale estimator for indoor localization. *Applied Sciences*, 2023, vol. 13 (6), p. 3900. DOI: 10.3390/app13063900.
37. Shabou S. *Time Series with R*. Available at: <https://s-ai-f.github.io/Time-Series/> (accessed 10.09.2025).
38. Karadayı Y., Aydin M.N., Öğrenci A.S. A hybrid deep learning framework for unsupervised anomaly detection in multivariate spatio-temporal data. *Applied Sciences*, 2020, vol. 10 (15), p. 5191. DOI: 10.3390/app10155191.
39. Fawwaz D.Z., Chung S.-H. Real-time and robust hydraulic system fault detection via edge computing. *Applied Sciences*, 2020, vol. 10 (17), p. 5933. DOI: 10.3390/app10175933.
40. Tayeh T., Aburakhia S., Myers R., Shami A. An attention-based ConvLSTM autoencoder with dynamic thresholding for unsupervised anomaly detection in multivariate time series. *Machine Learning and Knowledge*, 2022, vol. 4 (2), pp. 350–370. DOI: 10.3390/make4020015.
41. Pota M., De Pietro G., Esposito M. Real-time anomaly detection on time series of industrial furnaces: A comparison of autoencoder architectures. *Engineering Applications of Artificial Intelligence*, 2023, vol. 124, p. 106597. DOI: 10.1016/j.engappai.2023.106597.
42. Aminikhanghahi S. and Cook D. J. A Survey of Methods for Time Series Change Point Detection // *Knowl Inf Syst.* – 2017. – Vol. 51, N 2. – P. 339–367. – DOI: 10.1007/s10115-016-0987-z.
43. Antonova E.N. Posledovatel'nyi algoritm obnaruzheniya razladki mnogomernykh vremennykh ryadov s vysokoi dinamikoı [Sequential algorithm for disorder detection in multivariate time series]. *Izvestiya Sankt-Peterburgskogo gosudarstvennogo tekhnologicheskogo instituta = Bulletin of the Saint Petersburg State Institute of Technology*, 2022, no. 63 (89), pp. 93–99. DOI: 10.36807/1998-9849-2022-63-89-93-99.
44. Si T., Wang Y., Zhang L., Richmond E., Ahn T.-H., Gong H. Multivariate time series change-point detection with a novel Pearson-like scaled Bregman divergence. *Stats*, 2024, vol. 7 (2), pp. 462–480. DOI: 10.3390/stats7020028.
45. Hawkins M.D., Qiu P. The changepoint model for statistical process control. *Journal of Quality Technology*, 2003, vol. 35 (4). DOI: 10.1080/00224065.2003.11980233.
46. Chadha G.S., Shah S.R.B., Schwung A., Ding S.X. Shared temporal attention transformer for remaining useful lifetime estimation. *IEEE Access*, 2022, vol. 10, pp. 74244–74258. DOI: 10.1109/ACCESS.2022.3187702.
47. Gómez-Pau Á., Riba J.-R., Moreno-Eguilaz M. Time series RUL estimation of medium voltage connectors to ease predictive maintenance plans. *Applied Sciences*, 2020, vol. 10 (24), p. 9041. DOI: 10.3390/app10249041.
48. Rahmani M.M., Farahani M.A., Wuest T. Multivariate time-series classification of critical events from industrial drying hopper operations: A deep learning approach. *Journal of Manufacturing and Materials Processing*, 2023, vol. 7 (5), p. 164. DOI: 10.3390/jmmp7050164.
49. Lei Y., Wu Z. Time series classification based on statistical features. *Wireless Communications and Networking*, 2020, vol. 46. DOI: 10.1186/s13638-020-1661-4.

50. Zhang D., Zuo W., Zhang D., Zhang H. Time series classification using support vector machine with Gaussian elastic metric kernel. *20th International Conference on Pattern Recognition*, Istanbul, 2010, pp. 29–32. DOI: 10.1109/ICPR.2010.16.
51. Che Z., Purushotham S., Cho K., Sontag D., Liu Y. *Recurrent neural networks for multivariate time series with missing values*. Available at: <https://arxiv.org/abs/1606.01865> (accessed 10.09.2025).
52. Wang J., Wen G., Yang S., Liu Y. Remaining useful life estimation in prognostics using deep bidirectional LSTM neural network. *Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing)*, Chongqing, 2018, pp. 1037–1042.
53. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. *Attention is all you need*. Available at: <https://arxiv.org/abs/1706.03762> (accessed 10.09.2025).
54. Jin M., Koh H.Y., Wen Q., Zambon D., Alippi C., Webb G.I., King I., Pan S. *A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection*. Available at: <https://arxiv.org/abs/2307.03759> (accessed 10.09.2025).
55. Kakar J.K., Hussain S., Kim S.C., Kim H. TimeTector: A twin-branch approach for unsupervised anomaly detection in livestock sensor noisy data (TT-TBAD). *Sensors*, 2024, vol. 24 (8), p. 2453. DOI: 10.3390/s24082453.
56. Guan S., Zhao B., Dong Z., Gao M., He Z. GTAD: Graph and temporal neural network for multivariate time series anomaly detection. *Entropy*, 2022, vol. 24 (6), p. 759. DOI: 10.3390/e24060759.
57. Nguyen H.A.D., Ha Q.P., Duc H., Azzi M., Jiang N., Barthelemy X., Riley M. Long short-term memory Bayesian neural network for air pollution forecast. *IEEE Access*, 2023, vol. 11, pp. 35710–35725. DOI: 10.1109/ACCESS.2023.3265725.
58. Chen T., Liu X., Xia B., Wang W., Lai Y. Unsupervised anomaly detection of industrial robots using sliding-window convolutional variational autoencoder. *IEEE Access*, 2020, vol. 8, pp. 47072–47081. DOI: 10.1109/ACCESS.2020.2977892.
59. Goodfellow I.J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio, Y. *Generative adversarial networks*. Available at: <https://arxiv.org/abs/1406.2661> (accessed 10.09.2025).
60. Mogren O. *C-RNN-GAN: Continuous recurrent neural networks with adversarial training*. Available at: <https://arxiv.org/abs/1611.09904> (accessed 10.09.2025).
61. Liao S., Ni H., Szpruch L., Wiese M., Sabate-Vidales M., Xiao B. *Conditional Sig-Wasserstein GANs for time series generation*. Available at: <https://arxiv.org/abs/2006.05421> (accessed 10.09.2025).
62. Li D., Chen D., Jin B., Shi L., Goh J., Ng S.-K. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. *Lecture Notes in Computer Science*, 2019, vol. 11730, pp. 703–716. DOI: 10.1007/978-3-030-30490-4_56.
63. Ericsson L., Gouk H., Loy C.C., Hospedales T.M. *Self-supervised representation learning: introduction, advances and challenges*. Available at: <https://arxiv.org/abs/2110.09327> (accessed 10.09.2025).
64. Deldari S., Xue H., Saeed A., He J., Smith D.V., Salim F.D. *Beyond just vision: A review on self-supervised representation learning on multimodal and temporal data*. Available at: <https://arxiv.org/abs/2206.02353> (accessed 10.09.2025).

Для цитирования:

Хуснутдинов А.О., Хабаров В.И., Карманов В.С. Глубокое обучение для анализа многомерных временных рядов: систематизация типов данных, задач, архитектур и подходов // Системы анализа и обработки данных. – 2025. – № 3 (99). – С. 113–136. – DOI: 10.17212/2782-2001-2025-3-113-136.

For citation:

Khusnutdinov A.O., Khabarov V.I., Karmanov V.S. Glubokoe obuchenie dlya analiza mnogomernykh vremennykh ryadov: sistematizatsiya tipov dannykh, zadach, arkhitektur i podkhodov [Deep learning for multivariate time series analysis: systematization of data types, tasks, architectures and approaches]. *Sistemy analiza i obrabotki dannykh = Analysis and Data Processing Systems*, 2025, no. 3 (99), pp. 113–136. DOI: 10.17212/2782-2001-2025-3-113-136.