Инженерный вестник Дона, №6 (2025)
ivdon.ru/ru/magazine/archive/n6y2025/10141

Подход к моделированию кривой доходности, как многомерного временного ряда

А.А. Галимнуров

Уфимский университет наук и технологий

Аннотация: В исследовании представлен подход к моделированию многомерных временных рядов с использованием параметризации, на примере кривой доходности. Оценена эффективность добавление коэффициентов параметризации в предикаты, предложены новые функции потерь, которые акцентируют внимание на моделирование формы кривой. Были применены модели прогнозирования, включая рекуррентных нейронных сетей, Prophet и гибридные комбинации. Для автоматизации обработки и оценки данных была разработана система на базе Python. Метод повышает точность и интерпретируемость прогнозов, предлагая перспективный инструмент для финансового моделирования.
Ключевые слова: машинное обучение, финансовый инжиниринг, моделирование фондового рынка, рынок облигаций.

Введение

Прогнозирование многомерных временных рядов является одной из ключевых задач в современных исследованиях, связанных с анализом данных. Временные ряды, описываемые параметрическими кривыми, представляют собой сложные структуры, требующие применения специализированных методов для их анализа и прогнозирования. Актуальность данной темы обусловлена исследованием нового подхода к прогнозированию группы многомерных векторных рядов, описываемых параметрической кривой, а также широким спектром приложений, включая финансовые рынки [1], энергопотребление [2], погодные явления [3], медицинскую диагностику и управление транспортными потоками [4]. В последние годы наблюдается значительный прогресс в области прогнозирования временных рядов. В частности, в работе [5] предложена система нейросетевого прогнозирования, которая позволяет снизить ошибки прогнозирования на 2-30% в зависимости от метрики, что подтверждает эффективность нейросетевых подходов. В работе [6] разработан метод mDiSSiD для обнаружения аномалий в многомерных временных рядах. Этот

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

метод показал новый подход для обработки данных в условиях, потоковых данных, поступающих в режиме реального времени, что особенно важно для промышленных и медицинских приложений. В работе [7] исследовано применение глубокого обучения, включая архитектуры рекуррентных нейронных сетей, Transformer и TSMixer, для прогнозирования цен финансовых инструментов, что подчеркивает необходимость выбора оптимальных моделей в условиях высокой волатильности рынков. Кроме того, традиционные методы, такие как многомерная линейная регрессия, авторегрессия и гауссовские процессы [8], показывают ограниченную способность моделирования сложных нелинейных зависимостей между компонентами временных рядов. Современные архитектуры, включая рекуррентные нейронные сети, графовые нейронные сети и сверточные нейросети [9–10], значительно расширяют возможности прогнозирования, но требуют высоких вычислительных затрат и тщательной настройки параметров моделей [11-12].

В данном исследовании предложен подход к моделированию многомерных временных рядов, а именно использованию коэффициентов параметризации кривой, которая описывает временной ряд в заданный момент времени и предложены новые функции потерь для анализа эффективности моделирования, которые акцентируют внимание в первую очередь на форму параметрической кривой. Актуальность такого исследования обусловлена необходимостью повышения точности прогнозов, снижения вычислительных затрат и автоматизации процессов настройки моделей, а разработка новых методов и алгоритмов, учитывающих специфику многомерных данных, позволит решать сложные задачи в различных предметных областях, что делает данное направление исследований крайне важным и перспективным.

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

Теоретический анализ

В рамках описания подхода и вычислительного эксперимента рассмотрим кривую доходности – основной бенчмарк денежного рынка и показателя стоимости денег в зависимости от срока. Эта кривая формирует основу для определения безрисковой процентной ставки и широко используется в количественных исследованиях финансового рынка. Кривая доходности представляет из себя на каждый момент времени набор облигаций, как двумерный массив с показателями дюрации (приведенный срок) и доходности, как показано на рис. 1.

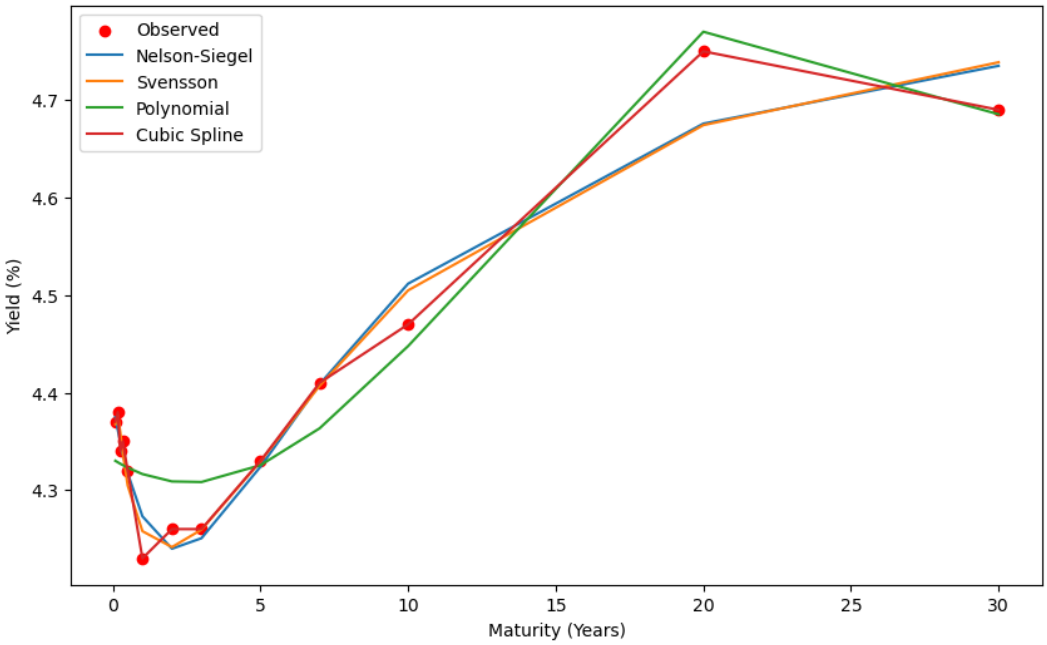


Рис.1: Визуализация исходных данных и параметризации на примере одной строки

По сути динамика кривой доходности представляет собой многомерный временной ряд. Пусть $\tau \in R^+$ — срок до погашения, а $y_\tau(t)$ — наблюдаемая доходность для данного срока на момент времени t. Целью моделирования является построение функции $f_\tau(t)$, прогнозирующей кривую доходности в будущем через h дней $f_\tau(t+h) \approx y_\tau(t+h)$.

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

В рамках данной главы предлагается оценка подхода к выбору предикатов для прогнозирования путем включения в выборку показателей, получаемых из различных параметризаций кривой.

Первым шагом предлагаемого подхода является параметризация кривой. Кривая доходности задается параметрически многими способами, но в рамках данной работы мы рассмотрим пять возможных моделей параметризации и несколько комбинаций прогнозных моделей, чтобы сравнить эффективность предложенного подхода против прогнозирования “сырых”, “допараметризированных” данных. В качестве параметризации кривой были выбраны полиномиальная аппроксимация, полиномиальные сплайны, метод главных компонент [10], а также уникальные для кривых доходностей модели серии Нельсона-Свенсона.

Пятифакторная модель Нельсона-Свенсона:

$$y(t) = \beta_0 + \beta_1 \frac{1-e^{-\tau/\lambda_1}}{\tau/\lambda_1} + \beta_2 \left(\frac{1-e^{-\tau/\lambda_1}}{\tau/\lambda_1} - e^{-\tau/\lambda_1}\right) + \beta_3 \left(\frac{1-e^{-\tau/\lambda_2}}{\tau/\lambda_2} - e^{-\tau/\lambda_2}\right) + \beta_4 e^{-\tau/\lambda_3}$$

Модель Нельсона-Сигеля с гауссовыми слагаемыми:

$$y(t) = \beta_0 + (\beta_1 + \beta_2)\frac{t}{\tau}\left(1 - \exp\left(-\frac{t}{\tau}\right)\right) - \beta_2 \exp\left(-\frac{t}{\tau}\right) + \sum_{i=1}^{9} g_i \exp\left(-\frac{(t-a)^2}{b_i^2}\right)$$

здесь $y(t)$ – доходность к погашению бумаги из набора образующего кривую доходности для срока τ, а $\lambda_1, \beta_i, a_i, b_i$ – параметры моделей типа Нельсона, которые отвечают за форму кривой.

Вторым шагом подхода является прогнозирования параметров кривой на 75 рабочих дней (календарный квартал), при этом прогнозирование на *n* дней строится последовательно, опираясь на однодневные прогнозы. Рассмотрим следующие прогнозные модели:

1. «Долгая кратковременная память» (Long short-term memory - LSTM) это разновидность рекуррентных нейронных сетей, ввиду своей структуры

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

хорошо подходящая для задач машинного обучения в случаях, когда важные события разделены временными промежутками.

2. Transformer это модель глубоких нейронных сетей, представленная в 2017 года. Преимуществом такого типа моделей по сравнению с рекуррентными нейросетями является возможность выявлять долгосрочные зависимости, за счет нововведенного механизма самовнимания.

3. ARIMA – популярная в практическом использовании модель авторегрессии для нестационарных временных рядов.

4. Prophet – это регрессионная модель, представленная в 2017 году, созданная для прогнозирования временных рядов с учётом сезонности и событий на основе различных динамик бизнес-показателей. Общую модель можно представить как сумму функции тренда, сезонной компоненты и заданной пользователем нерегулярные события

5. Векторная модель коррекции ошибок (Vector error correction model - VECM) – регрессионная модель для прогнозирования многомерных временных рядов, когда переменные коинтегрированы, строится на основе матрицы коинтеграции.

В рамках исследования проведена оценка эффективности прогнозирования подхода на основе параметризации кривой с использованием пяти рассмотренных моделей, а также их гибридных комбинаций, объединяющие методы регрессионного анализа и машинного обучения. Так для прогнозирования трендов и сезонных компонент применяются классические регрессионные модели, такие как ARIMA и Prophet, учитывающие тренд и сезонность, а для моделирования краткосрочной динамики и нелинейных зависимостей будут использоваться нейросетевые подходы, включая LSTM и Transformer, что позволит более точно захватывать сложные паттерны в данных. Такой подход направлен на увеличение прогнозной способности и интерпретируемости данных.

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

Третий шаг заключается в выборе функции потерь. В рамках данной работы предлагается рассмотреть две новые метрики оценки эффективности и сравнить ее с существующими: квадратным корнем средквадратичной ошибки (Root mean square error - RMSE) и интегральной метрикой. Мотивация в рассмотрении новых метрик заключается в том, что для моделирования кривой временных доходностей в первую очередь важно не абсолютное сходство, а «правильное» построение формы кривой и точки изломов.

Пусть у нас есть реальный ряд $y^{true}(x)$ и прогнозируемый ряд $y^{pred}(x)$. Их кривизны:

$$k^{true}(x) = \frac{f''_{true}(x)}{(1 + (f'_{true}(x))^2)^{3/2}} \quad ; \quad k^{pred}(x) = \frac{f''_{pred}(x)}{(1 + (f'_{pred}(x))^2)^{3/2}}$$

На основе показателей кривизны можно составить метрику разницы кривизн, которая рассчитывается как средняя абсолютная разница между этими значениями на равномерно распределённых N точках области значений кривой:

$$CD(N) = Mean(|k^{true}(x_i) - k^{pred}(x_i)|)$$

Преимуществом данной метрики заключается в том, что она учитывает не только точность значений, но и сходство геометрии прогнозируемого временного ряда.

Другой метрика, которая может улучшить эффективность прогнозирования многомерных временных рядов с приоритезацией на моделирование формы кривой, может стать метрика, основанная на косинусное сходство с учетом экстремумов. Определим косинусное сходство между векторами значений экстремумов:

$$S_{cos}(X_r, X_p) = \frac{X_r \cdot X_p}{\| X_r \| \cdot \| X_p \|}$$

Общая метрика сходства экстремумов:

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

$$S_{\text{sum-cos}} = \sum_{i,j} S_{\cos}(f_r(x_i), f_p(x_j))$$

Если кривые полностью совпадают по форме экстремумов, сумма будет равна числу экстремумов. Если количество экстремумов различается, вводится штраф:

$$S_{\text{penalty}} = \lambda \mid N_r - N_p \mid$$

где N_r и N_p — количество экстремумов у реальной и прогнозируемой кривой, а λ — коэффициент штрафа. Окончательная метрика:

$$S_{\text{final}} = S_{sum-\cos} - S_{penalty}$$

Эмпирический анализ

Для оценки вышеописанного подхода и дальнейших исследований разработан блок в программном комплексе [13] на языке Python с использованием библиотек pandas, streamlit, skikit-learn, pytorch и других. Он упрощает процесс прогнозирования многомерных временных рядов за счет автоматизации получения и обработки входных данных, их дальнейшей параметризации, а также готового интерфейса для визуализации прогнозной способности модели. Разработанная программа включает несколько этапов обработки данных, прогнозирования и обратного преобразования результатов. Рассмотрим ключевые шаги в этом процессе. На вход программа получает набор данных о государственных бумагах с указанием их дюрации и доходности – красные точки, выделенные на рис.1. Следующим шагом программа модулирует кривую, основываясь на одном из описанных выше методов. Далее параметры кривой доходности, полученные на предыдущем шаге передаются в прогнозные модели как предикаты. В отличие от классического подхода, где прогнозируется сам временной ряд доходностей, в данном случае моделируются изменения параметров кривой.

На финальном этапе строится синтетический набор, аналогичный исходным данным, на основе предсказанной параметрической кривой.

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

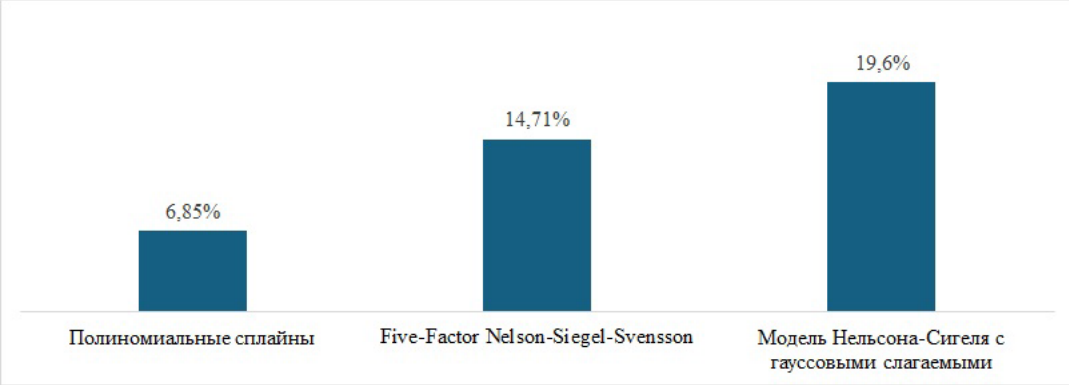
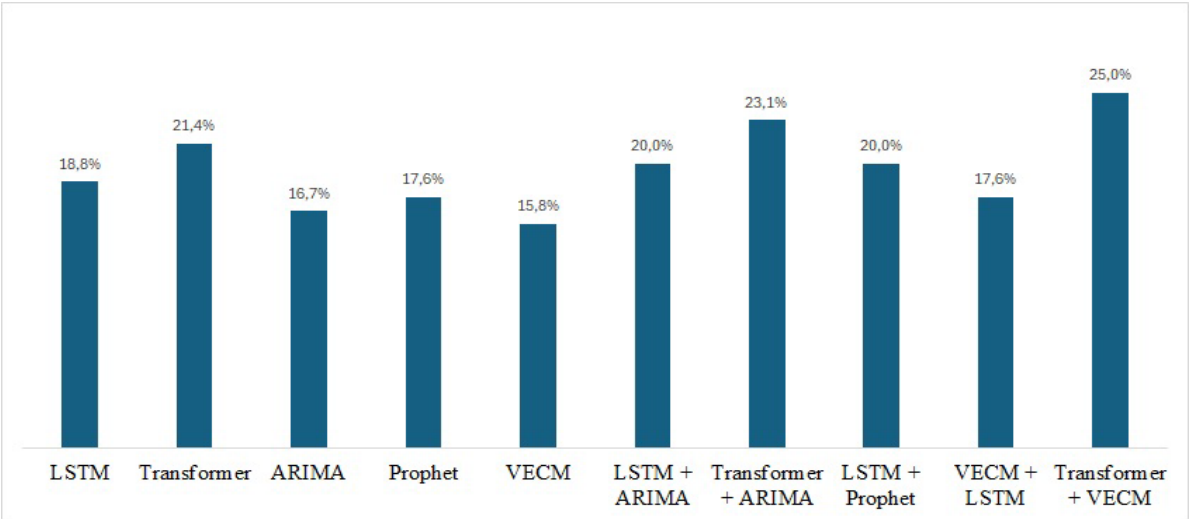


Рис. 2: Улучшение качества прогнозирования (RMSE) за счет параметризации

Результаты оценки качества прогнозирования моделей показывают, что подход с предварительной параметризацией улучшает качество прогнозирования до 19,6% (рис. 2) по сравнению с подходом без предварительной параметризации. При этом лучшие результаты демонстрирует параметризации, специально разработанные для структуры заданного многомерного временной ряда. Еще одним наблюдением, полученным из численного эксперимента, является то, что гибридные модели обеспечивают лучшую предсказательную способность по сравнению с не гибридными методами на квартальном интервале – самым длинным интервале. (рис. 3).



© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

Рис. 3: Улучшение качества прогнозирования (RMSE) с параметризацией

Для оценки эффективности функции потери, как и в работе [7] была проведена историческая валидация портфеля бумаг широкого рынка облигаций. Алгоритм оценки эффективности функции потери при этом выглядит следующим образом:

1) На вход подается широкий портфель из *n* бумаг, которые представляются в виде трехмерного массива со значениями дюраций, доходностей и весов в портфеле облигаций $(\mathrm{dur_i}, \mathrm{ytm_i}, \mathrm{w_i})_{\mathrm{i=1}}^{\mathrm{n}}$. Также на вход подаются модели LSTM + Prophet, строящие недельные прогнозы кривой и которые были обучены с использованием следующих функций потерь на 100 равномерно распределенных точках: RMSE, Интегральное сходство, средней разнице кривизн CD(N) и регуляризированной разнице косинусных сходств на экстремумах $\mathrm{S_{final}}$
2) Все бумаги портфеля разбиваются на категории - бакеты с интервалом по полгода [0, 0.5), [0.5, 1), … [5.5, 6) и для каждой категории определяется средневзвешенная по бакету дюрация $(\mathrm{t_{backet\ i}})_{i=1}^{m}$, где m кол-во бакетов.
3) Для каждой недели начиная с начала 2020 по конец 2021 года:
    a. Строим прогноз кривой на следующую неделю пятью моделями, на основе исторических данных.
    b. Для каждого бакета считаем, насколько изменяется кривая на заданной точке в бакете.

$$\forall \mathrm{i} \in [1, \ldots \mathrm{m}]\, \partial_{\mathrm{i}} = y_{pred}(t + 1_{backet_i}) - y_{true}(t_{backet_i})$$

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

c. Пропорционально ∂_i поменяй веса в портфеле w_i, при условии что рост доходности обратно пропорционален изменению цены.

d. пересчитай $(t_{\text{backet i}})_{i=1}^{m}$

e. если не последняя неделя, продолжи цикл

В результате получилось, что модель на метрике S_final была успешнее в текущем эксперименте на 3.8% (рис. 4).

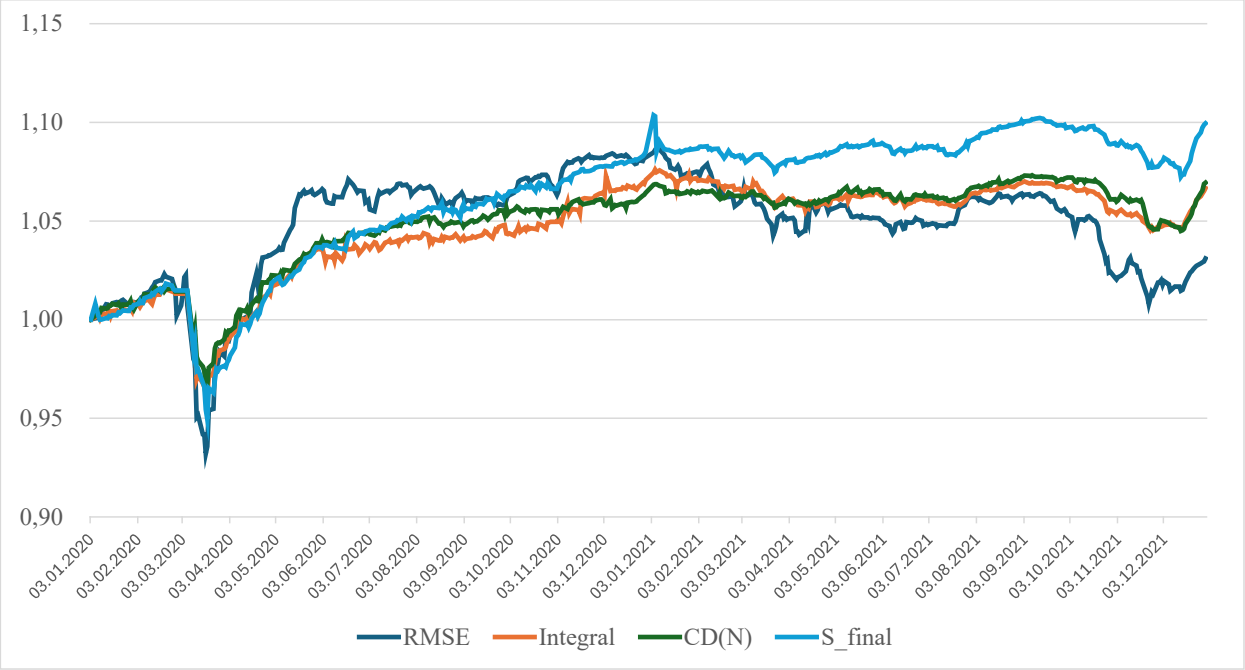


Рис. 4. Оценка функций потерь

Заключение

В данной работе продемонстрирован новый подход к моделированию многомерных временных рядов. В качестве рассмотренного ряда берется набор облигаций, который образуют кривую доходностей– ключевой индикатор денежного рынка, являющийся базисом для большинства финансовых моделирований и оценок. Основная идея подхода заключается в добавлении в предикаты коэффициентов предварительной параметризации кривой и применении новых функции оценки прогнозной эффективности.. В исследовании были рассмотрены несколько различных методов

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

параметризации, эмпирически исследована эффективность добавления этапа параметризации. Для прогнозирования параметров кривой применены различные прогностические модели включая LSTM, Prophet и другие, а также их гибридные комбинации. Также была разработана программа на языке Python, позволившая автоматизировать процесс получения, обработки и параметризации подаваемых на вход облигаций, а также автоматизировать оценки предсказательной способности заданных в программном комплексе моделей. Результаты исследования показали, что гибридные модели обеспечивают лучшую эффективность по сравнению с отдельно используемыми моделями, так комбинация Transformer с VECM показала наилучшие результаты. Были предложены новые функции потерь для данной задачи, которые на исторической валидации показали свою эффективность. Таким образом, предложенный подход к прогнозированию многомерных временных рядов на основе параметризации кривой демонстрирует преимущества по сравнению с традиционными методами моделирования кривых доходностей, он позволяет не только повысить точность прогнозов, но и улучшить интерпретацию динамики временных рядов, что делает его перспективным направлением для дальнейших исследований в области финансового моделирования и прогнозирования временных рядов.

Литература

1. Feng S., Xu C., Zuo Y., Chen G., Lin F., XiaHou J. Relation-aware dynamic attributed graph attention network for stocks recommendation.//Pattern Recognition. 2022. 121. URL: doi.org/10.1016/j.patcog.2021.108119

2. Zhicheng C., Shibo F., Zhang Z., Xiao X., Gao X., Zhao P. SDformer: Similarity-driven Discrete Transformer For Time Series Generation. // In The Thirtyeighth Annual Conference on Neural Information Processing Systems. 2024.

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

3. Jin M., Koh H. Y., Wen Q., Zambon D., Alippi C., Webb G., King I., Pan S. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. // 2023. URL: arxiv.org/abs/2307.03759

4. Praveen Kumar B., Hariharan K. Multivariate time series traffic forecast with long short-term memory based deep learning model. // Control and Computing. 2020. URL: doi.org/10.1109/PICC51425.2020.9362455

5. Милосердов Д. И. Усовершенствованная система нейросетевого прогнозирования групп временных рядов с непрерывным обучением // Информационно-управляющие системы. 2024. № 1(128). С. 20-30. URL: doi.org/10.31799/1684-8853-2024-1-20-30

6. Краева Я. А. Нейросетевой метод обнаружения аномалий в многомерных потоковых временных рядах // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2024. №4. URL: doi.org/10.14529/cmse24040

7. Зелезецкий, Д. В. Глубокое обучение в задаче прогнозирования финансовых временных рядов // Труды МФТИ. Труды Московского физико-технического института (национального исследовательского университета). 2024. Т. 16, № 3(63). С. 35-48

8. Mudelsee M. Trend analysis of climate time series: A review of methods. // Earth-Science Reviews. 2018. vol. 190. pp. 310 – 322. URL: doi:10.1016/j.earscirev.2018.12.005

9. Пашшоев Б., Петрусевич Д.А. Анализ нейросетевых моделей для прогнозирования временных рядов. // Russian Technological Journal. 2024. 12(4). С. 106–116. URL: doi.org/10.32362/2500-316X-2024-12-4-106-116.

10. Галимнуров А.А., Исмагилова А.С. Применение математического моделирования для прогнозирования спреда корпоративных облигаций // Инженерный вестник Дона. 2025. №2. URL: ivdon.ru/ru/magazine/archive/n2y2025/9813

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

11. Ехлаков Р. С., Судаков В. А., Прогнозирование стоимости котировок при помощи LSTM и GRU сетей // Препринты ИПМ им. М. В. Келдыша. 2022. 017. С. 13. URL: doi.org/10.20948/prepr-2022-17

12. Shi J., Jain M., Narasimhan G. Time Series Forecasting (TSF) Using Various Deep Learning Models. // 2022. URL: arxiv.org/abs/2204.11115v1

13. Исмагилова А.С., Галимнуров А.А. Управление портфелем на рынке облигаций на основе математического моделирования. Свидетельство о регистрации программы для ЭВМ № 2024617557. Дата регистрации: 03.04.2024. Заявка № 2024616518/69 от 17.03.2024. Опубликована: 03.04.2024. URL: fips.ru/EGD/41c7551e-ce9c-4199-b23b-6b1241beaf2d

References

1. Feng S., Xu C., Zuo Y., Chen G., Lin, F. XiaHou, J. Pattern Recognition. 2022. 121. URL: doi.org/10.1016/j.patcog.2021.108119

2. Zhicheng C., Shibo F., Zhang Z., Xiao X., Gao X., Zhao P. In The Thirtyeighth Annual Conference on Neural Information Processing Systems. 2024.

3. Jin M., Koh H. Y., Wen Q., Zambon D., Alippi C., Webb G. I., King I., Pan S. 2023. URL: arxiv.org/abs/2307.03759

4. Praveen Kumar B., Hariharan K. Control and Computing. 2020. URL: doi.org/10.1109/PICC51425.2020.9362455

5. Miloserdov D. I. Informacionno-upravljajushhie sistemy. 2024. № 1(128). pp. 20-30. URL: doi.org/10.31799/1684-8853-2024-1-20-30

6. Kraeva Ja. Vestnik JuUrGU. Serija: Vychislitelnaja matematika i informatika. 2024. №4. URL: doi.org/10.14529/cmse24040

7. Zelezeckij, D. V. TRUDY MFTI. Trudy Moskovskogo fiziko-tehnicheskogo instituta (nacionalnogo issledovatelskogo universiteta). 2024. T. 16, № 3(63). pp. 35-48

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025

8. Earth Science Reviews. 2018. vol. 190. pp. 310 – 322. doi:10.1016/j.earscirev.2018.12.005

9. Pashshoev B., Petrusevich D.A. Russian Technological Journal. 2024. 12(4). pp. 106–116. URL: doi.org/10.32362/2500-316X-2024-12-4-106-116.

10. Galimnurov A.A., Ismagilova A.S. Inzhenernyj vestnik Dona. 2025. № 2. URL: ivdon.ru/ru/magazine/archive/n2y2025/9813

11. Ehlakov R. S., Sudakov V. A. Preprinty IPM im. M. V. Keldysha. 2022. 017. pp. 13. .URL: doi.org/10.20948/prepr-2022-17

12. Shi J., Jain M., Narasimhan G. 2022. URL: arxiv.org/abs/2204.11115v1

13. Ismagilova A.S., Galimnurov A.A. Upravlenie portfelem na rynke obligacij na osnove matematicheskogo modelirovanija. Svidetel'stvo o registracii programmy dlja JeVM № 2024617557. [Portfolio management in the bond market based on mathematical modeling. Computer program registration Certificate No. 2024617557]. Data registracii: 03.04.2024. Zajavka № 2024616518/69 ot 17.03.2024. Opublikovana: 03.04.2024. URL: fips.ru/EGD/41c7551e-ce9c-4199-b23b-6b1241beaf2d

Дата поступления: 2.04.2025

Дата публикации: 25.05.2025

© Электронный научный журнал «Инженерный вестник Дона», 2007–2025